



Thuanny Andrade Borges

Caracterização funcional de genes importantes no metabolismo de
xilose das leveduras *Rhodotorula dairenensis* e *Pseudozyma*
brasiliensis sp. nov.

Functional characterization of important genes in xylose metabolism of
the yeasts *Rhodotorula dairenensis* and *Pseudozyma brasiliensis* sp.
nov.

Campinas 2014



UNIVERSIDADE ESTADUAL DE CAMPINAS
Instituto de Biologia



THUANNY ANDRADE BORGES

Caracterização funcional de genes importantes no metabolismo de xilose das leveduras *Rhodotorula dairenensis* e *Pseudozyma brasiliensis* sp. nov.

Functional characterization of important genes in xylose metabolism of the yeasts *Pseudozyma brasiliensis* sp. nov. and *Rhodotorula dairenensis*

Dissertação apresentada ao
Instituto de Biologia da Universidade Estadual de
Campinas como parte dos requisitos exigidos
para obtenção do Título de Mestre em
Genética e Biologia Molecular na área de Microbiologia

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

Orientador: Dr Gustavo H. Goldman

Co-orientadora: Dra. Juliana V. de Castro Oliveira

Este exemplar corresponde à versão final da dissertação
defendida pela aluna Thuanny Andrade Borges e orientada pelo
Prof. Dr. Gustavo H. Goldman

Campinas, 2014

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

B644c Borges, Thuanny Andrade, 1989-
Caracterização funcional de genes importantes no metabolismo de xilose das leveduras *Rhodotorula dairenensis* e *Pseudozyma brasiliensis* sp. nov. / Thuanny Andrade Borges. – Campinas, SP : [s.n.], 2014.

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

1. Leveduras. 2. Xilose. 3. Etanol. 4. Insetos. I. Goldman, Gustavo Henrique. II. Oliveira, Juliana Velasco de Castro. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Functional characterization of important genes in xylose metabolism of the yeasts *Rhodotorula dairenensis* and *Pseudozyma brasiliensis* sp. nov.

Palavras-chave em inglês:

Yeasts

Xylose

Ethanol

Insects

Área de concentração: Microbiologia

Titulação: Mestra em Genética e Biologia Molecular

Banca examinadora:

Gustavo Henrique Goldman [Orientador]

Suzan Pantaroto de Vasconcelos

Rosana Goldbeck

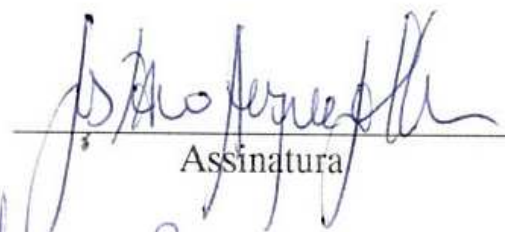
Data de defesa: 11-04-2014

Programa de Pós-Graduação: Genética e Biologia Molecular

Campinas, 11 de abril de 2014

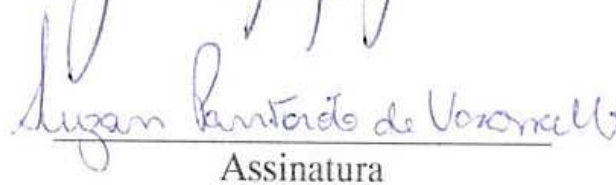
BANCA EXAMINADORA

Prof. Dr. Gustavo Henrique Goldman (Orientador)



Assinatura

Profa. Dra. Suzan Pantaroro De Vasconcelos



Assinatura

Dra. Rosana Goldbeck



Assinatura

Prof. Dr. Marcelo Brocchi

Assinatura

Dra. Thabata Maria Alvares

Assinatura

DEDICATÓRIA

Especialmente à Deus, aos meus pais Margarete e Antonio
e ao meu noivo Diogo.

AGRADECIMENTOS

Ao Prof. Dr. Gustavo H. Goldman pela oportunidade em ser sua aluna de Mestrado, por me oferecer um projeto tão interessante e por todo o direcionamento.

À Dra. Juliana Velasco de Castro Oliveira, minha co-orientadora, agradeço aos inúmeros ensinamentos e dicas que foram fundamentais no meu processo de aprendizado e crescimento bem como no andamento do projeto.

Ao Dr. André Ricardo de Lima Damásio, pela paciência e ensinamentos.

Ao Dr. Diego Mauricio Riano-Pachon, pelo suporte em bioinformática.

Aos Profs. Drs. Carlos Augusto Rosa e Ednildo Machado, pelas parcerias.

À Universidade Estadual de Campinas, através do Instituto de Biologia e a pós-graduação em Genética e Biologia Molecular.

Ao Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE) pela oportunidade, por aceitar meu projeto de pesquisa e pelo suporte técnico.

À FAPESP, pelo suporte financeiro.

À todos os amigos e equipe do PPB do CTBE, muito obrigada por serem sempre tão gentis e atenciosos e pelos ótimos momentos que passamos juntos.

Em especial, à Aline Tieppo, pela amizade, paciência, dedicação, suporte, mas principalmente pelas inúmeras conversas, conselhos e almoços, muito obrigada, sentirei muita falta. À Lívia e a Ana obrigada pela amizade, lealdade, companheirismo e pelas conversas, risadas e choros, vocês se tornaram grandes amigas.

À todos da “ilha 4” e à Rebequinha, obrigada pelas conversas e momentos agradáveis. Em especial, ao Gustavo Borin pelas muitas risadas, ajudas e por me aguentar “maritacando” ao lado da sua bancada.

Ao João, Fernanda (loira) e à Thabata por me fazerem sentir acolhida em uma ilha que não fosse a minha.

À Dra. Patrícia, por todas as conversas, desabafos e por me fazer ser uma pessoa mais confiante e forte.

Em especial à Carolzinha, obrigada por me aguentar na TPM, pela parceria, conversas, jantas, risadas e desabafos.

Ao meu noivo Diogo, por ser meu melhor amigo, por aguentar todas as minhas reclamações e choros, por fazer eu me sentir a pessoa mais inteligente do mundo, pelos conselhos e broncas, por fazer o “fardo” ficar mais leve, pelo carinho, amor e cumplicidade.

Aos meus pais Margarete e Antonio e à minha Irma Isabela, por sempre estarem ao meu lado e por terem me ensinado os verdadeiros valores da vida. Pelo amor incondicional, todo apoio, base e ajuda. Amo muito vocês.

Agradeço principalmente a Deus pela minha vida e por sempre me dar forças para alcançar e vencer os meus objetivos, por eu ter motivos para celebrar e pela capacidade de aprender cada dia mais.

RESUMO

A completa utilização da biomassa lignocelulósica é um dos pré-requisitos para tornar o processo do etanol de segunda geração economicamente competitivo. Entretanto, a fermentação dos açúcares disponíveis na biomassa celulósica apresenta um desafio único em decorrência da presença de outros açúcares, além da glicose, tais como xilose e arabinose, os quais não são fermentáveis por *Saccharomyces cerevisiae*, a principal levedura utilizada na indústria. Em leveduras, a conversão de D-xilose a D-xilulose ocorre em duas etapas: inicialmente D-xilose é reduzida a D-xilitol pela enzima xilose redutase (XR); em seguida, D-xilitol é oxidado a D-xilulose pela enzima xilitol desidrogenase (XDH). Em *S. cerevisiae*, essas duas enzimas possuem cofatores distintos, provocando assim um desequilíbrio redox na célula, impedindo a utilização anaeróbica de pentoses. Uma alternativa para que *S. cerevisiae* possa produzir etanol através destes açúcares é modificá-la geneticamente com genes provenientes de micro-organismos que naturalmente realizem esta conversão. Dessa forma, este projeto isolou e caracterizou as leveduras *Rhodotorula dairenensis* TAB01 e *Pseudozyma brasiliensis* sp.nov. GHG001, a partir do intestino de insetos parasitas da cana-de-açúcar, através da análise de expressão dos genes *xyl1* e *xyl2* que codificam, respectivamente, as enzimas XR e XDH. Adicionalmente, foi avaliada a hiper-expressão destes na cepa industrial de *S. cerevisiae* PE-2 e na cepa PE-2 super-expressando o gene *XKS1* de *S. cerevisiae* (PE-XKS), o qual codifica a enzima xilulose quinase (XK). *P. brasiliensis* GHG001 e *R. dairenensis* TAB01 mostraram-se promissoras para o estudo. Estas duas espécies são ótimas assimiladoras de xilose e seus genes de XR e XDH foram expressos heterologicamente em *S. cerevisiae*, gerando as cepas PE Rd, PE Pb, PE-XKS Rd e PE-XKS Pb, que foram testadas quanto a capacidade de utilizar xilose para crescimento. Porém, essas cepas recombinantes não conseguiram utilizar xilose como única fonte de carbono, o que evidencia o quão importante é o desbalanço de cofatores das enzimas no metabolismo de xilose. Além disso,

devido a uma notável importância biotecnológica, também foi caracterizada uma enzima xilanolítica produzida por *P. brasiliensis* GHG001 a qual apresentou uma alta atividade xilanolítica. Essa xilanase, pertencente à família GH11, apresentou pH e temperatura ótimos iguais a 4 e 55°C, respectivamente, sendo sua estrutura secundária constituída de folhas beta. A sua atividade enzimática é influenciada por alguns íons como Ca^{2+} e os resultados de sua cinética indicaram um comportamento sigmoidal característico de enzimas alostéricas. Usando xilano como substrato, esta xilanase produz xilooligossacarídeos que podem ser usados industrialmente como prébióticos, que junto a outras aplicabilidades desta enzima na indústria, demonstram seu alto potencial biotecnológico.

ABSTRACT

Full utilization of lignocellulosic biomass is one of the prerequisites to make the process of second generation ethanol economically competitive. However, the fermentation of sugars available in cellulosic biomass presents a unique challenge due to the presence of other sugars besides glucose, such as xylose and arabinose, which are not fermentable by *Saccharomyces cerevisiae*, the main yeast used in the industry. In yeasts, the conversion of D-xylose to D-xylulose occurs in two steps: first D-xylose is reduced to xylitol by the enzyme D-xylose reductase (XR), and then D-xylitol is oxidized to D-xylulose by the enzyme xylitol dehydrogenase (XDH). In *S. cerevisiae*, these two enzymes have different cofactors, thereby causing a redox imbalance in the cell, preventing the use of pentoses anaerobically. An alternative to *S. cerevisiae* can produce ethanol by these sugars is genetically modifying it with genes derived from microorganisms that naturally carry out this conversion. Thus, this project characterized the yeasts *Rhodotorula dairenensis* TAB01 and *Pseudozyma brasiliensis* GHG001, isolated from the gut of insects pests of sugarcane, through the analysis of genes expression of *xy11* and *xy12* which encoding respectively, the XR and XDH enzymes and measured the enzymatic activity of the same. Additionally, was evaluated the overexpression of these in *S. cerevisiae* industrial strains PE-2 and PE-2 XKS1 overexpressing *xks1* gene of *S. cerevisiae*, which encodes the enzyme xylulokinase (XK). *P. brasiliensis* GHG001 and *R. dairenensis* TAB01 shown promise for this study as they are great xylose assimilators and their assimilating xylose genes are expressed in the presence of this sugar. XR and XDH were heterologously expressed in *S. cerevisiae*, generating strains Rd PE, PE Pb, PE-XKS Rd and PE-XKS Pb, which were tested for the ability to utilize xylose for growth. However, these recombinant strains were unable to utilize xylose as sole carbon source, which shows how important is the imbalance of cofactors of the metabolize xylose enzymes. Moreover, due to a remarkable biotechnological importance, has also been characterized a xylanolytic enzyme produced by *P.*

brasiliensis GHG001 which showed a high activity, yet unreported by other eukaryotic microorganism. This xylanase belonging to family GH11, showed pH and temperature optima equal to 4 and 55°C, respectively, and its secondary structure consists of β -sheets. This enzyme has influence of some ions such as Ca^{2+} and showed a sigmoidal kinetic behavior, characteristic of allosteric enzymes. From xylan, the xylanase produces xylooligosaccharides which can be used industrially as prebiotics, which together with other applicability of this enzyme in the industry, demonstrate their high biotechnological potential.

Sumário:

1.	Introdução.....	21
2.	Revisão Bibliográfica.....	23
2.1.	Os combustíveis fósseis e os biocombustíveis	23
2.2.	Etanol de Primeira e Segunda Geração	24
2.3.	Biomassa lignocelulósica e a parede celular	27
2.4.	Utilização de pentoses por micro-organismos	30
2.4.1.	A levedura <i>Saccharomyces cerevisiae</i>	32
2.4.2.	As cepas recombinantes	35
2.4.3.	A Cepa industrial Pedra-2	37
2.4.4.	A levedura <i>Rhodotorula dairenensis</i> TAB01	39
2.4.5.	A levedura <i>Pseudozyma brasiliensis</i> sp. nov. GH001.....	41
3.	Objetivos.....	43
3.1.	Objetivo geral	43
3.2.	Objetivos específicos	43
4.	Metodologia.....	45
4.1.	Obtenção dos isolados.....	45
4.2.	Identificação dos isolados	46
4.2.1.	Extração de DNA <i>in house</i>	46
4.2.2.	Amplificação e sequenciamento da região ITS2	47
4.2.3.	Construção de árvores filogenéticas através das regiões ribossomais ITS, 5.8S e D1/D2	48
4.3.	Perfil de crescimento das leveduras em xilose	50
4.4.	Caracterização molecular dos genes da <i>Rhodotorula dairenensis</i> TAB01 e da <i>Pseudozyma brasiliensis</i> sp. nov. GHG001.....	51
4.4.1.	Extração de DNA.....	51
4.4.2.	Isolamento do RNA e síntese de cDNA	51
4.4.3.	Identificação molecular dos genes da xilose redutase (xyl1) e xilitol desidrogenase (xyl2) da <i>Rhodotorula dairenensis</i> TAB01.....	52
4.4.4.	Purificação, ligação, clonagem e sequenciamento	53
4.4.5.	Identificação molecular dos genes da xilose redutase (xyl1) e da xilitol desidrogenase (xyl2) de <i>Pseudozyma brasiliensis</i> sp. nov. GHG001	54
4.5.	RT-PCR em Tempo Real (qRT-PCR).....	56

4.6.	Atividade enzimática de XR e XDH endógenos.....	59
4.7.	Clonagem de <i>xyl1</i> e <i>xyl2</i> no vetor de expressão heteróloga pSP-G1	60
4.7.1.	Clonagem dos genes da <i>R. dairenensis</i> TAB01.....	62
4.7.2.	Clonagem dos genes de <i>P. brasiliensis</i> GHG001	62
4.8.	Construção das cepas de <i>S. cerevisiae</i> recombinantes	63
4.8.1.	Transformação por acetato de lítio	64
4.9.	Caracterização das cepas de <i>S. cerevisiae</i> PE-2 recombinantes.....	66
4.9.1.	Perfil de crescimento das cepas de <i>S. cerevisiae</i> PE-2 recombinantes	66
4.9.2.	Atividade enzimática de XR e XDH das cepas de <i>S. cerevisiae</i> recombinantes.....	67
4.10.	Cepas de <i>S. cerevisiae</i> PE-2 XKS1 recombinantes.....	67
4.10.1.	Perfil de crescimento das cepas de <i>S. cerevisiae</i> PE-2 XKS1 recombinantes	67
5.	Resultados e Discussão.....	69
5.1.	A Identificação dos isolados.....	69
5.1.1.	A Identificação das leveduras através da região ITS2	69
5.1.2.	<i>Rhodotorula dairenensis</i> TAB01	70
5.1.3.	<i>Pseudozyma brasiliensis</i> sp.nov. GHG001	72
5.2.	Perfil de crescimento das leveduras	75
5.3.	A caracterização molecular dos genes <i>xyl1</i> e <i>xyl2</i> de <i>R. dairenensis</i> TAB01	77
5.4.	A expressão dos genes <i>xyl1</i> e <i>xyl2</i> da <i>R. dairenensis</i> TAB01.....	80
5.5.	A atividade enzimática da XR e XDH endógenas da <i>R. dairenensis</i> TAB01	82
5.6.	A identificação molecular dos genes <i>xyl1</i> e <i>xyl2</i> da <i>Pseudozyma brasiliensis</i> GHG001	84
5.7.	Análise da expressão gênica de <i>xyl1</i> e <i>xyl2</i> de <i>P. brasiliensis</i> GHG001	90
5.8.	A atividade enzimática da XR e XDH endógenas da <i>P. brasiliensis</i> GHG001...91	
5.9.	A clonagem dos genes <i>xyl1</i> e <i>xyl2</i> no vetor pSP-G1 e construção das cepas de <i>S. cerevisiae</i> recombinantes.....	92
5.9.1.	A atividade enzimática de XR e XDH das cepas PE-2 recombinantes ..95	
5.9.2.	O perfil de crescimento das cepas PE-2 recombinantes.....	98

5.9.3. O perfil de crescimento das cepas PE-XKS1 recombinantes	99
6. Artigo: <i>Pseudozyma brasiliensis</i> sp. nov., a novel xylanolytic Ustilaginomycetous yeast species isolated from an insect pest of sugarcane roots.....	103
7. Artigo: Draft Genome Sequence of <i>Pseudozyma brasiliensis</i> sp. nov. Strain GHG001, a High Producer of Endo-1,4-Xylanase Isolated from an Insect Pest of Sugarcane.....	129
8. Artigo: Biochemical characterization of an endoxylanase from <i>Pseudozyma brasiliensis</i> sp. nov. strain GHG001 isolated from the intestinal tract of <i>Chrysomelidae</i> larvae associated to sugarcane roots.....	133
9. Considerações finais.....	141
10. Referências.....	143

Lista de abreviações:

cDNA – DNA complementar

CTBE – Laboratório Nacional de Ciência e Tecnologia do Bioetanol

dNTP – Desoxirribonucleotídeos fosfatados

DNA – Ácido desorribonucléico

h – Hora

ITS – espaçador transcrito interno

LSU – Sub-unidade maior do ribossomo

min – Minuto

MM – Meio mínimo

NAD⁺ - Nicotinamida adenina dinucleotídeo

NADH – Nicotinamida adenina dinucleotídeo reduzido

NADPH – Nicotinamida adenina dinucleotídeo fosfato

NEB – *New England Biolabs*

nm – Nanômetro

nt – Nucleotídeo

Oligo – Oligonucleotídeo

OD – Densidade Ótica

pb – Par de bases

PCR – Reação em Cadeia da Polimerase

qRT-PCR – RT-PCR em Tempo Real

RNA – Ácido ribonucléico

RPE – Enzima ribulose 5-fosfato 3-epimerase

RPI – Enzima ribulose 5-fosfato isomerase

rpm – Rotações por minuto

s – Segundo

SSU – Sub-unidade menor do ribossomo

TAE – Tris-Acetato-EDTA

TAL – Enzima transaldolase

tef – Gene do fator de elongação

TKL – Enzima transquetolase

T_m – Temperatura de *Melting*

U – Unidade

µg – Micrograma

YPER – Reagente de extração de proteína de leveduras

1. Introdução

A queima de combustíveis fósseis é um grande contribuinte para o aumento do nível de CO₂ na atmosfera, que está diretamente associada ao aquecimento global observado nas últimas décadas. Os efeitos adversos de gases de efeito estufa (GEE) no meio ambiente, em conjunto com declínio das reservas de petróleo e a contínua volatilidade dos mercados do mesmo, têm levado as nações do mundo todo a adotarem políticas que promovam o uso de fontes renováveis de energia (ARGUESO et al., 2009; NAIK et al., 2010). Uma das principais fontes de energia renovável pensando-se em biocombustíveis trata-se da conversão de hidratos de carbono de origem vegetal em bioetanol através de fermentação microbiana. Nesse contexto, indústrias dos EUA e Brasil têm desenvolvido processos que usam o amido de milho e a sacarose da cana-de-açúcar, respectivamente, para produzir bioetanol. Como resultado, esses dois países lideram atualmente o *ranking* de produção mundial de etanol (SOUZA et al., 2012). Entretanto, por ser derivada, principalmente, de culturas agrícolas, a produção de biocombustíveis por essas tecnologias (Primeira Geração) causa preocupação com relação ao possível impacto na biodiversidade local, origem da matéria-prima e uso da terra, que têm como consequência a competição com culturas alimentares (NAIK et al., 2010). Assim, para acompanhar a demanda crescente pelo combustível etanol faz-se necessário o desenvolvimento de novas tecnologias (DIAS, et al., 2012; HAHN-HAGERDAL, et al., 2006), surgindo então, os biocombustíveis de Segunda Geração, que são gerados a partir de materiais lignocelulósicos, tais como resíduos agrícolas, florestais ou industriais (RODRIGUES, 2011). O bagaço da cana-de-açúcar é o principal resíduo do agronegócio brasileiro, pois uma tonelada de cana-de-açúcar moída resulta em cerca de 250 kg desse resíduo (SOUZA, 2013). Desta forma, embora uma parte do bagaço seja utilizada pelas usinas para a co-geração de energia elétrica, o uso de aproximadamente 25% para a produção do etanol lignocelulósico (Segunda Geração), já representaria um aumento considerável na produção mundial desse biocombustível (LIMA, 2011).

No entanto, desafios biotecnológicos devem ser superados para garantir que o etanol lignocelulósico seja produzido de forma comercialmente viável, dentre eles destaca-se a completa utilização de todos os tipos de açúcares que compõem a celulose e a hemicelulose (HAHN-HAGERDAL et al., 2007). A fermentação desses açúcares representa um desafio único, pois além dos polissacarídeos da parede celular formarem uma arquitetura extremamente complexa, recalcitrante, a levedura *Saccharomyces cerevisiae*, o principal micro-organismo utilizado na produção de etanol, é incapaz de naturalmente fermentar pentoses como a xilose, o segundo açúcar mais abundante da biomassa lignocelulósica (ALVIRA et al., 2010). Em consequência disso, a *S. cerevisiae* tem sido extensivamente engenheirada geneticamente (CAI et al., 2012) com o intuito de gerar um micro-organismo capaz de converter eficientemente pentoses a etanol em condições industriais (HAHN-HAGERDAL et al., 2007). Neste sentido, com o objetivo de gerar uma cepa de *S. cerevisiae* que fermente xilose a etanol eficientemente, este trabalho propôs a expressão heteróloga dos genes *xyl1* e *xyl2* da via da redução/oxidação do metabolismo de xilose na cepa industrial *Saccharomyces cerevisiae* Pedra-2 (PE-2). Esses genes foram obtidos a partir das leveduras assimiladoras de xilose, *Rhodotorula dairenensis* TABO1 e *Pseudozyma brasiliensis* GHG001, isoladas do trato intestinal de insetos parasitas da cana-de-açúcar. Além disso, este trabalho também caracterizou bioquimicamente uma enzima xilanase produzida pela *P. brasiliensis* GHG001.

2. Revisão Bibliográfica

2.1. Os combustíveis fósseis e os biocombustíveis

A crescente preocupação com as mudanças climáticas, a necessidade de redução de emissão de gases efeito estufa (evidenciada desde o Protocolo de Kyoto em 1997), a busca pela sustentabilidade, associada ao aumento contínuo dos preços dos combustíveis fósseis fazem com que a procura por fontes de energia mais limpas seja uma medida essencial e inevitável (BUCKERIDGE et al., 2010). Entre as opções mais viáveis, o bioetanol se destaca como um biocombustível de referência por se tratar de uma fonte inesgotável de energia, considerada de baixo impacto ambiental (ARGUESO et al., 2009), sendo uma alternativa à gasolina (BABRZADEH et al., 2012).

As culturas energéticas (culturas onde sua biomassa pode servir para a obtenção de energia) são capazes de compensar as emissões de gases do efeito-estufa através da conversão do dióxido de carbono atmosférico em carbono orgânico na biomassa e no solo. Em teoria, a mesma quantidade de carbono que é liberado quando o bioetanol é queimado, é absorvida da atmosfera pela colheita do próximo ano através da fotossíntese, completando assim o ciclo do carbono (BABRZADEH et al., 2012). Em fevereiro de 2010, a *Environmental Protection Agency* (EPA) dos Estados Unidos, designou o etanol de cana-de-açúcar como um biocombustível avançado, capaz de reduzir consideravelmente as emissões de gases do efeito estufa (GEE) em comparação com a gasolina (RODRIGUES, 2011). A utilidade das plataformas de produção de biocombustíveis depende tanto sua viabilidade econômica quanto ambiental, sendo assim, o etanol de cana brasileiro tem se mostrado plenamente competitivo com a gasolina devido a várias conquistas biotecnológicas (BABRZADEH et al., 2012).

O bioetanol é produzido por meio da fermentação de carboidratos derivados de matérias primas agrícolas, principalmente do amido de milho, muito utilizado pelas indústrias nos EUA, e da sacarose da cana-de-açúcar, pelas indústrias no

Brasil (ARGUESO et al., 2009). Como resultado, esses dois países são atualmente os dois maiores produtores de bioetanol do mundo, compartilhando cerca de 70% do mercado mundial, sendo que os EUA ocupam o primeiro lugar no *ranking* (ARGUESO et al., 2009; SOUZA et al., 2012).

A cana-de-açúcar foi escolhida como substrato para a produção de etanol devido a várias razões, incluindo sua ótima adaptação aos solos e as condições climáticas brasileiros (SOCCOL et al., 2010). O Brasil é o maior produtor de cana do mundo, com uma produção equivalente a 600 toneladas por ano (BASSO et al., 2013), e o maior produtor mundial de etanol produzido a partir da fermentação do seu caldo (SOCCOL et al., 2010). Apesar dos Estados Unidos serem os maiores produtores mundiais de etanol utilizando o milho como matéria-prima, o sistema brasileiro de agroenergia é o mais eficiente, sendo considerado mais vantajoso por diversas razões, entre elas: (1) em um ciclo de vida as emissões de GEE são reduzidas em 84% (em volume), enquanto o etanol de milho reduz as emissões em 30%; (2) o custo de produção é 60% menor do que o custo de produção a partir do milho, e (3) o rendimento em litros de etanol por hectare (6470 L/ha) é significativamente maior do que o rendimento do etanol de milho (4180 L/ha) (GOLDEMBERG; GUARDABASSI, 2009). Segundo a União da Indústria de Cana-de-Açúcar (Unica), na Safra 2012/2013 foram produzidos no país 21,36 bilhões de litros deste combustível e a estimativa para a safra de 2013/2014 é de 25,37 bilhões de litros, 17,21% superior à safra anterior.

2.2. *Etanol de Primeira e Segunda Geração*

Atualmente, toda a produção mundial dos biocombustíveis se baseia nas tecnologias de primeira geração (1G), ou seja, produção de etanol principalmente a partir da sacarose da cana-de-açúcar e do amido do milho (SIMS et al., 2010). Como desvantagem, por serem produzidos a partir de matérias-primas utilizadas também como alimentos para humanos e animais, a produção deste combustível

competem com as indústrias de alimentos pelo uso da biomassa e terras agrícolas, o que gera questionamentos éticos, políticos e ambientais (RODRIGUES, 2011).

O processo de produção de etanol no Brasil geralmente utiliza alta densidade de células de levedura em um modo semi-contínuo de batelada-alimentada que também é conhecida como processo *Melle-Boinot*, adotado em 75% das destilarias brasileiras. São fermentados caldos contendo elevadas concentrações de açúcar (caldo de cana e/ou melaço diluído), produzindo altas concentrações deste combustível, com alto rendimento (90-92% do máximo teórico) e produtividade (cada ciclo fermentativo dura de 6-10 h) (BABRZADEH et al., 2012; BASSO, BASSO, ROCHA, 2011).

A fermentação do caldo ou melaço é realizada com a levedura *Saccharomyces cerevisiae* a partir de uma concentração inicial de sacarose em torno de 200 g/L em dornas de 100 a 1000 mil litros de volume total. Após cada ciclo obtém-se um teor alcoólico inferior a 100 g/L que é enviado para a unidade de destilação, resultando no etanol absoluto (95,5%) ou o anidro com teor de álcool de no mínimo 99,7%. A levedura é separada do fermentado (chamado de vinho) por centrifugação e reciclada para uma nova fermentação após o tratamento com ácido sulfúrico concentrado, neutralização e determinação de contaminação microbiológica. A contaminação por leveduras selvagens e bactérias é elevada e há necessidade de combatê-las constantemente com antibióticos antes da reutilização em novo ciclo fermentativo. Uma parte do processo é impulsionada pela energia obtida com a queima do bagaço da cana que alimenta as caldeiras e gera eletricidade (RODRIGUES, 2011).

Embora a produção mundial de etanol em 2012 tenha sido de 85,2 bilhões de litros (dados da F. O. Licht, divulgados pela *Global Renewable Fuels Alliance*), está cada vez mais evidente que o etanol de primeira geração (1G) não será suficiente para atender às demandas futuras por biocombustíveis (SOUZA et al., 2012). Para contornar este problema, a produção de etanol de segunda geração (2G) vem ganhando crescente interesse. Estes combustíveis são produzidos a partir de uma ampla variedade de fontes não alimentícias, e entre eles está

incluída a utilização de materiais lignocelulósicos, tais como resíduos agrícolas, florestais e industriais (RODRIGUES, 2011).

Essas tecnologias utilizam os materiais lignocelulósicos como matéria prima uma vez que estes são mais baratos, abundantes e podem ser produzidos nas mais variadas condições de solo e clima (LEITE, LEAL, 2007). Neste sentido, resíduos, rejeitos ou produtos não comestíveis como lascas de madeira provenientes de árvores (eucaliptos) e biomassa de gramíneas (palha de milho e bagaço de cana-de-açúcar) estão entre as matérias-primas consideradas promissoras para a produção de bioetanol 2G (RODRIGUES, 2011; SOUZA et al., 2012).

No processo de produção de etanol 1G cerca de 92% do bagaço é usado para gerar energia. Soccol e colaboradores (2010) relataram que, se os 8% não utilizados no processo fossem convertidos em etanol, poderia ser obtido um aumento na produção de 2200 litros (L) de etanol por hectare e uma redução do uso da terra de 29%.

No processo de obtenção do etanol 2G, o objetivo é “desmontar” a parede celular para utilizar os polissacarídeos como fonte de açúcares fermentáveis. No entanto, a estrutura da parede é muito complexa e este processo de desmonte deve ser minucioso para preservar intactos os monossacarídeos que serão utilizados na fermentação (SOUZA et al., 2012). Esse tipo de processo envolve passos físicos, químicos e/ou biológicos, e se inicia com um pré-tratamento (físico e/ou químico) da biomassa, seguido do tratamento com hidrolases glicolíticas e proteínas acessórias para a geração de monossacarídeos (hexoses e pentoses). Por fim, como na produção do combustível 1G, ocorre a fermentação dos açúcares a etanol e recuperação dos produtos e sub-produtos (STEPHEN; COLUMBIA, 2012).

Entretanto, até o momento, embora “plantas-pilotos” já estejam sendo desenvolvidas, a produção de bioetanol 2G ainda não é economicamente viável (NAIK et al., 2010; DIAS et al., 2011). Isto se deve ao fato que várias barreiras técnicas precisam ser superadas para que seu potencial seja realizado. Dentre

elas, destaca-se a redução e eliminação de químicos utilizados no pré-tratamento, diminuição do custo e quantidade de enzimas empregadas na etapa de hidrólise e a necessidade de um micro-organismo eficiente que seja capaz de fermentar todos os açúcares disponíveis (ALVIRA et al., 2010; KIM et al., 2013). Neste sentido, a fermentação de pentoses representa um grande desafio e sua eficiente e rápida utilização é um pré-requisito para a produção de biocombustíveis, a partir de biomassa renovável, de forma sustentável e econômica (KIM et al., 2013).

Nos últimos anos, em laboratórios do mundo todo, intensos esforços vêm sendo feitos para desenvolver tecnologias eficientes para o pré-tratamento da biomassa e de enzimas a fim de aprimorar a sacarificação da celulose/hemicelulose, bem como para o desenvolvimento de micro-organismos capazes de fermentar ambos os açúcares (pentoses e hexoses) liberados do material lignocelulósico (SOUZA et al., 2012; SOCCOL et al., 2010).

2.3. *Biomassa lignocelulósica e a parede celular*

A biomassa vegetal é responsável por cerca de 60% da biomassa no mundo, sendo que 90% desse material é composto de material lignocelulósico e os 10% restantes por materiais oleaginosos e outros (TENGARDY E SZACKS, 2002). A parede celular da biomassa lignocelulósica é composta por uma mistura de polissacarídeos, proteínas, compostos fenólicos e sais minerais (BUCKERIDGE et al., 2010). Os três componentes principais são a celulose, a hemicelulose e a lignina, que são unidas entre si por ligações covalentes, formando uma rede complexa resistente a ataques microbianos (Figura 1) (AMORIM et al., 2011; BASSO et al., 2013).

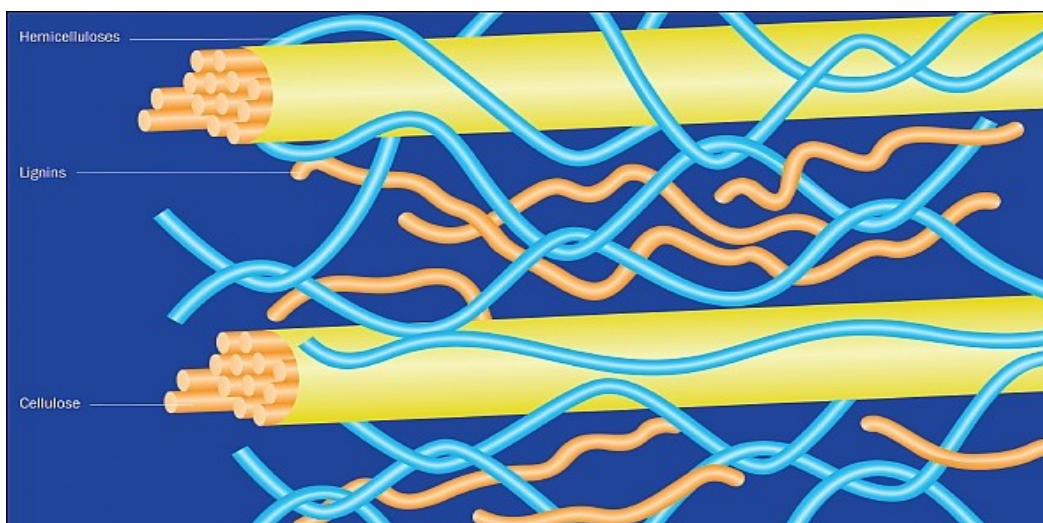


Figura 1 – Esquema da parede celular vegetal. Disponível em: <http://www.scidacreview.org/0905/html/biofuel.htm>

A celulose é um polissacarídeo composto por microfibrilas formadas por moléculas de glicose (C6) unidas por ligação β -1,4. A união de microfibrilas forma uma estrutura linear e semicristalina. A cristalinidade confere resistência à hidrólise devido à ausência de água na estrutura e as fortes ligações entre as cadeias de glicose, que impedem a ação de hidrolases nas ligações β -1,4 (BASSO et al., 2013). A hemicelulose é um componente relativamente amorfo, constituída por xilose, arabinose, galactose, manose e outros açúcares. A hemicelulose pode ligar-se às microfibrilas de celulose por pontes de hidrogênio, formando uma proteção que impede o contato entre uma microfibrila e outra, formando uma rede coesiva. O xiloglucano é a principal hemicelulose em muitas paredes celulares primárias. No entanto, na parede celular secundária, a qual predomina na biomassa vegetal, as hemiceluloses são tipicamente xilanos e arabinoxilanos (BASSO et al., 2013). Por fim, a lignina é essencialmente o cimento que propicia a rigidez estrutural das plantas, sendo formada por uma rede polimérica tridimensional de unidades de metoxilas, arilpropanos e hidroxifenóis (RODRIGUES, 2011).

Assim, dois terços do conteúdo energético presente nos carboidratos da cana-de-açúcar são armazenados nos polissacarídeos localizados nas paredes celulares das folhas e do colmo, sendo que apenas a celulose e hemicelulose podem ser utilizadas como matérias-primas para a produção de etanol 2G (SOUZA et al., 2012). Desta forma, o principal componente do hidrolisado lignocelulósico é a glicose e o segundo açúcar mais abundante é a D-xilose, sendo que esta última é particularmente elevada em gramíneas e madeira. Sendo assim, a conversão da biomassa em energia não é economicamente viável a não ser que a hemicelulose seja utilizada em adição à celulose (JEFFRIES, 2006; HECTOR et al., 2011).

O bagaço da cana-de-açúcar é o resíduo fibroso resultante da moagem dos colmos e da extração do caldo da cana-de-açúcar (PANDEY, 2000). No Brasil, o bagaço representa um dos principais materiais lignocelulósicos a ser considerado para a produção do etanol de segunda geração, visto que a cada tonelada de cana moída, são produzidos 250 kg deste resíduo (KÄMPER et al., 2006). Além disso, o bagaço é considerado um rico reservatório de energia solar devido ao maior rendimento massa/área de cultivo de cana-de-açúcar comparado às outras culturas (cerca de 80 t/ha em comparação à 1 t/ha de trigo e 2 t/ha de outras gramíneas) (BASSO et al., 2013). O bagaço apresenta em sua parede celular, aproximadamente, 4-9% de cinzas, 19-24% de lignina, 32-42% de celulose e 27-32% de hemicelulose (SOCCOL et al., 2010), sendo que esta última apresenta elevado teor de D-xilose, correspondendo em até 80% do total de açúcar nesta fração (MARTON et al., 2004). Devido ao seu baixo conteúdo de cinzas, o bagaço oferece várias vantagens em relação aos outros resíduos de culturas quando utilizados em bioprocessamento.

Desta forma, embora uma parte do bagaço seja utilizada pelas usinas para geração de energia elétrica, se pequenas frações do mesmo fossem utilizadas para a produção do etanol lignocelulósico já representaria um aumento considerável na produção mundial deste biocombustível (PACHECO, 2011). Entretanto, como já dito anteriormente, para que o etanol lignocelulósico seja

produzido de uma forma economicamente viável, todos os tipos de açúcares derivados da biomassa (hexoses e pentoses) devem ser convertidos em etanol, fazendo necessária a busca de micro-organismos que realizem eficientemente esta conversão em condições industriais (RAGAUSKAS et al., 2006).

2.4. *Utilização de pentoses por micro-organismos*

Inúmeros micro-organismos são capazes de metabolizar pentoses e análises das vias metabólicas envolvidas neste processo revelaram duas formas distintas de utilização destes açúcares (Figura 2). A primeira é a via de isomerização, que ocorre principalmente em bactérias, onde a xilose isomerase (XI) converte D-xilose em D-xilulose, sem a utilização de cofatores (MATSUSHIKA et al., 2009).

Em leveduras e demais organismos eucariotos, a conversão de D-xilose a D-xilulose ocorre por uma via de redução/oxidação, em duas etapas: inicialmente, D-xilose é reduzida a D-xilitol, pela enzima xilose redutase (XR) NADPH ou NADH-dependente; em seguida, D-xilitol é oxidado a D-xilulose, pela enzima xilitol desidrogenase (XDH) NAD^+ -dependente (MATSUSHIKA et al., 2009).

Após a produção de D-xilulose por uma destas vias, ela é fosforilada pela enzima xilulose quinase (XK) a D-xilulose 5-P que em seguida é convertida pela via das pentoses a gliceraldeído-3P e frutose-6P, ingressando na via da glicólise, chegando a piruvato que é convertido a etanol e vários outros produtos (MATSUSHIKA et al., 2009).

Porém, as leveduras apresentam vantagens sobre as bactérias em fermentação industriais como o tamanho das células, as paredes celulares mais espessas, o melhor crescimento em pHs baixos, os requisitos nutricionais menos rigorosos, e uma maior resistência à contaminação (JEFFRIES, 2006).

Via da redução-oxidação

Via da isomerização

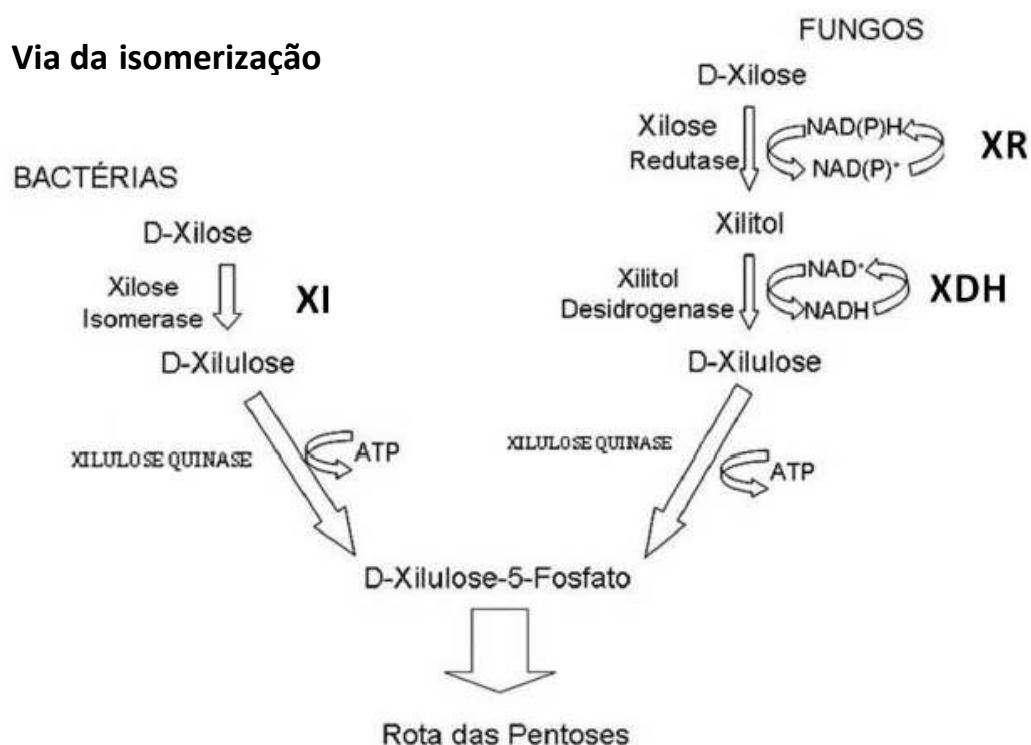


Figura 2 – Vias de utilização de xilose por diferentes micro-organismos. Adaptado de Hahn-Hägerdal et al., 2004.

A maioria das xiloses redutases já identificadas apresenta uma dependência estrita ao NADPH como cofator, enquanto que a xilitol desidrogenase é específica para o cofator NAD⁺. Esse desbalanço de cofatores pode levar ao excesso de NADH e a depleção de NADPH. Aerobicamente, o excesso de NADH pode ser reoxidado pela mitocôndria, mas em condições anaeróbicas, o excesso de NADH é acumulado e a utilização de xilose fica mais lenta e em alguns casos cessa (HECTOR et al., 2011). Acredita-se que o desbalanço de cofatores seja um dos principais responsáveis pela fermentação ineficiente anaeróbia de xilose, o que resulta em um acúmulo de xilitol e baixa produtividade de etanol (CAI et al., 2012).

2.4.1. A levedura *Saccharomyces cerevisiae*

Saccharomyces cerevisiae é a principal levedura utilizada na produção industrial de etanol devido à excelente capacidade de fermentar glicose. Além disso, possui alta tolerância ao etanol, resistência e robustez aos diferentes pHs e temperaturas do processo, bem como aos inibidores presentes nas altas concentrações de melaço (CAI et al., 2012; BETTIGA et al., 2008; MATSUSHIKA et al., 2009), além de não exibir muitas das limitações fermentativas presentes nas bactérias (LIN E TANAKA., 2006). Assim, *S. cerevisiae* é considerada o micro-organismo com maior potencial na produção de etanol a partir de biomassa, principalmente no que diz respeito à eficiência e robustez (LEANDRO et al., 2008). Entretanto, ela não é capaz de metabolizar eficientemente pentoses, o que para a produção do etanol 2G representaria deixar frações de 20-30% do hidrolisado lignocelulósico inutilizáveis (CAI et al., 2012), uma vez que, como dito anteriormente, ele contém grandes quantidades de xilose e arabinose (LEANDRO et al., 2008).

A incapacidade da *S. cerevisiae* em fermentar xilose é atribuída a alguns fatores como: 1) a impossibilidade dessa levedura converter eficientemente esse substrato em xilulose, uma vez que apresenta baixos níveis de expressão dos genes de utilização de xilose e baixas atividades das enzimas correspondentes a esses genes (gene *xy1*: xilose redutase – XR e gene *xy2*: xilitol desidrogenase - XDH) (RAGAUSKAS et al., 2006; MATSUSHIKA et al., 2009); 2) o desbalanço redox causado pelos cofatores dessas enzimas; 3) baixa atividade da enzima xilulo-quinase (XK); 4) baixas atividades das enzimas ribulose 5-fosfato 3-epimerase (RPE), ribulose 5-fosfato isomerase (RPI), transquetolase (TKL) e transaldolase (TAL) da via de pentose fosfato (CAI et al., 2012); 5) não apresenta um eficiente transportador de pentoses, uma vez que, a absorção de xilose ocorre unicamente através dos transportadores de hexoses que aceitam xilose como substrato, porém com baixíssima afinidade. Isso faz com que a absorção de xilose seja baixa e dependente da concentração de glicose presente no ambiente

(LEANDRO et al., 2008). Assim, segundo Cai et al. (2012), os esforços para que a *S. cerevisiae* seja capaz de fermentar xilose podem ser divididos em seis aspectos ao longo de toda a via metabólica de xilose (Figura 3):

- a) Buscar um transportador para xilose mais eficiente (em laranja);
- b) Interligar a xilose intracelular à xilulose, através da expressão heteróloga dos genes da xilose redutase e xilitol desidrogenase, ou através do gene da xilose isomerase (em verde);
- c) Atenuar o desbalanço redox das enzimas XR e XDH mudando as especificidades de seus cofatores (em amarelo);
- d) Fortalecer o fluxo da xilulose à xilulose-5-fosfato através da expressão do gene endógeno da xilulo-quinase (XK) (em azul);
- e) Aumentar a expressão das enzimas da via de pentose fosfato (em roxo) e;
- f) Engenheirar a célula por inteiro (em cinza).

Assim, as pesquisas têm focado na engenharia genética da *S. cerevisiae* com a finalidade de torná-la apta a co-fermentar hexoses e pentoses derivadas de matérias-primas lignocelulósicas (LEANDRO et al., 2008).

2.4.2. As cepas recombinantes

Alguns micro-organismos são naturalmente capazes de assimilar xilose como as leveduras *Scheffersomyces stipitis* (*Pichia stipitis*), *Pachysolen tannophilus* e *Candida shehatae*, que metabolizam esse açúcar principalmente pela via da redução/oxidação (HO et al., 1998). A *S. stipitis* aparece como a única dentre as leveduras assimiladoras de xilose, sob condições de limitação de oxigênio, com capacidade em fermentar xilose para produzir etanol. A capacidade de *S. stipitis* em fermentar a xilose é atribuída principalmente à capacidade de sua XR utilizar NADH quase tão bem quanto NADPH, evitando, assim, o desbalanço de cofatores (HECTOR et al., 2011).

Desse modo, uma das propostas que a engenharia genética e metabólica tem buscado é a melhoria das atividades fermentativas de micro-organismos que naturalmente metabolizem xilose, como é o caso da *S. stipitis* (JEFFRIES, 2006). Entretanto, embora a *S. stipitis* consiga fermentar xilose em etanol com rendimentos teóricos sob condições anaeróbias, ela requer oxigênio para a utilização eficiente de xilose, e apresenta uma taxa de consumo de açúcar mais lenta em comparação com *S. cerevisiae* (MATSUSHIKA et al., 2009). Adicionalmente, outro problema é que a maioria das leveduras naturalmente fermentadoras de xilose excretam quantidades substanciais do sub-produto xilitol e sua produção reduz o rendimento de etanol (MATSUSHIKA et al., 2009).

Além disso, sabe-se que micro-organismos nativos que fermentam xilose a etanol não possuem a robustez necessária para sobreviver às condições industriais estressantes deste processo, tais como a alta concentração de etanol, elevada temperatura, estresse osmótico (devido às altas concentrações de açúcares e sais), acidez, contaminação bacteriana e reciclo celular além de ser um ambiente estritamente anaeróbio assim, esses micro-organismos não

conseguem fermentar xilose a etanol eficientemente (BASSO et al., 2008; KIM et al., 2013).

Sendo assim, atualmente, a maioria dos esforços visam desenvolver uma *S. cerevisiae* que fermente xilose, engenheirando-o geneticamente com genes da via da redução/oxidação ou da via da isomerização. A via da xilose isomerase tem sido considerada como uma abordagem promissora para utilização de xilose em *S. cerevisiae* devido a ter uma única etapa de conversão de xilose a xilulose, sem a produção de xilitol e uso de cofatores. No entanto, poucas XIs heterólogas são funcionalmente expressas em *S. cerevisiae* (CAI et al., 2012). Já a via da redução/oxidação apresenta a vantagem de ser termodinamicamente mais vantajosa quando comparada com a via da isomerização, nos quesitos taxa de consumo de xilose e produção de etanol, mas apresenta dois gargalos intrínsecos: a produção de xilitol e o desbalanço de cofatores.

Assim, devido às diferenças nas propriedades bioquímicas e origens distintas das vias (XI ou XR/XDH) ainda é difícil determinar qual delas é mais adequada para ser introduzida em *S. cerevisiae* (KIM et al., 2013). Deste modo, as pesquisas têm focado no engenheiramento da *S. cerevisiae* com genes da *S. stipitis* (xilose redutase e xilitol desidrogenase) ou a partir do fungo filamentoso *Piromyces* sp. (xilose isomerase) (LEANDRO et al., 2008). Segundo Kim e colaboradores (2013) a superexpressão do gene endógeno da XK (*xks1*) da *S. cerevisiae*, o qual naturalmente tem baixa expressão, ou a expressão heteróloga do gene XK (por exemplo *xy13* de *S. Stipitis*) são necessários para facilitar o metabolismo de xilose através da via das pentoses fosfato.. Assim, a via XR/XDH/XK, XK é um como um passo limitante e as cepas com superexpressão tanto de XR/XDH ou de XI tiveram redução significativa de xilitol quando um gene adicional de XK foi integrado (KIM et al., 2013). No entanto, foi sugerido que a expressão demasiada dos genes XK poderia causar uma inibição do crescimento em xilose ou uma redução na taxa de consumo de xilose. Assim como estes trabalhos, inúmeros outros vêm sendo realizados, mas o rendimento e a produção

de etanol por estas cepas engenheiradas ainda são muito inferiores quando comparadas à fermentação de glicose (CAI et al., 2012).

Desta forma, é fundamental a descoberta de genes da via metabólica da xilose mais eficientes, como por exemplo, os isolados de leveduras selvagens fermentadoras de xilose. A identificação destes genes pode criar a oportunidade do melhoramento genético da assimilação de xilose em *S. cerevisiae* através da introdução dos mesmos sob o controle de sinais regulatórios apropriados.

Estas leveduras selvagens podem ser isoladas de nichos ecológicos enriquecidos para o seu crescimento, tais como o trato intestinal de insetos herbívoros, uma vez que a microbiota intestinal é importante para que estes insetos utilizem recursos nutricionais que sem eles seriam inacessíveis (SUH et al., 2005; SUH et al., 2003). Como o intestino humano e o rúmen do gado, o intestino dos insetos hospeda um diverso grupo de micro-organismos, e a comunidade microbiana realiza funções simbióticas importantes, incluindo a desconstrução da biomassa para os insetos herbívoros (SHI et al., 2010).

Neste sentido, em um trabalho prévio realizado em nosso grupo, foi identificado diversos fungos e leveduras presentes no trato intestinal de insetos que interagem com a cana-de-açúcar (Projeto de auxílio à pesquisa FAPESP 10/51324-2 “Identificação de leveduras fermentadoras de xilose isoladas a partir do trato intestinal de insetos que parasitam a cana-de-açúcar”). Entre estes micro-organismos, duas leveduras, *Rhodotorula dairenensis* TAB01 e uma levedura do gênero *Pseudozyma* sp., a qual posteriormente verificamos tratar-se de uma nova espécie que foi denominada *Pseudozyma brasiliensis* sp.nov., foram selecionadas para o presente estudo.

2.4.3. A Cepa industrial Pedra-2

Cepas de *S. cerevisiae* utilizadas nas fermentações industriais, as chamadas linhagens industriais, foram e são muitas vezes comparadas com as linhagens laboratoriais (domésticas) em termos de eficiência de utilização de

açúcar e tolerância ao estresse que o processo fermentativo impõe (KIM et al., 2013). Linhagens laboratoriais são úteis para estudar a genética molecular de leveduras, enquanto linhagens industriais são geralmente selecionadas para terem um ótimo desempenho sob condições industriais (MATSUSHIKA et al., 2009). Segundo Matsushika e colaboradores (2009), as linhagens industriais têm se mostrado mais tolerantes ao hidrolisado que as linhagens laboratoriais, além de serem superiores na produção de etanol.

No processo de produção de etanol brasileiro, as células de levedura são recicladas e utilizadas na fermentação subsequente durante os 6-9 meses de safra (ARGUESO et al., 2009). Assim, como já citado anteriormente, uma cepa de levedura ideal para a produção de etanol deve crescer neste ambiente hostil com estresses bióticos e abióticos. Dessa forma, algumas cepas de leveduras altamente produtivas tendem a dominar o fermentador durante toda a temporada de produção, permitindo fermentações eficientes e estáveis (BABRZADEH et al., 2012).

Na década de 1980 cepas de leveduras eram melhoradas geneticamente com o intuito de aumentar a eficiência da fermentação do etanol de primeira geração, mas no ambiente industrial elas eram incapazes de competir com as leveduras contaminantes selvagens que eram mais robustas (ARGUESO et al., 2009). Então, na década de 1990, uma estratégia foi adotada: a seleção, dentre as leveduras selvagens contaminantes, das cepas que combinavam alta eficiência de fermentação com persistência prolongada no sistema. Um destes isolados mais bem sucedidos é a cepa Pedra-2 (PE-2), que atualmente é utilizada por 30% das destilarias brasileiras, gerando 10% do bioetanol mundial (ARGUESO et al., 2009). Características como dominância sobre outras linhagens e permanência no processo durante a época de colheita da cana foram levados em consideração quanto à escolha dessa cepa, que hoje já tem seu genoma sequenciado (ARGUESO et al., 2009; AMORIM et al., 2011). Recentemente, no trabalho de Brown e colaboradores (2013) foi obtida informação sobre a expressão dos genes (transcriptoma) da PE-2 sob condições que mimetizavam a fermentação industrial

a fim de entender como essa linhagem responde ao ambiente estressante e alternante do processo fermentativo. Recentemente, foi construída uma cepa de PE-2 com os alelos para o gene *URA3* deletados, fazendo com que ela seja passível de receber genes exógenos. Desse modo, esta estirpe pode ser utilizada para melhorias genéticas através da tecnologia do DNA recombinante (AMORIM et al., 2011).

2.4.4. A levedura *Rhodotorula dairenensis* TAB01

A *R. dairenensis* foi elevada ao nível de espécie há pouco tempo através de análises da região D1/D2 do 26S rRNA, da região ITS e *fingerprints* com micro e mini-satélites (GADANHO; SAMPAIO, 2002). Inicialmente esta espécie era considerada como *Rhodotorula glutinis*, entretanto Gadanho e colaboradores (2002) mostraram que *R. dairenensis* é mais próxima filogeneticamente da *R. mucilaginosa* do que da própria *R. glutinis* (Figura 4).

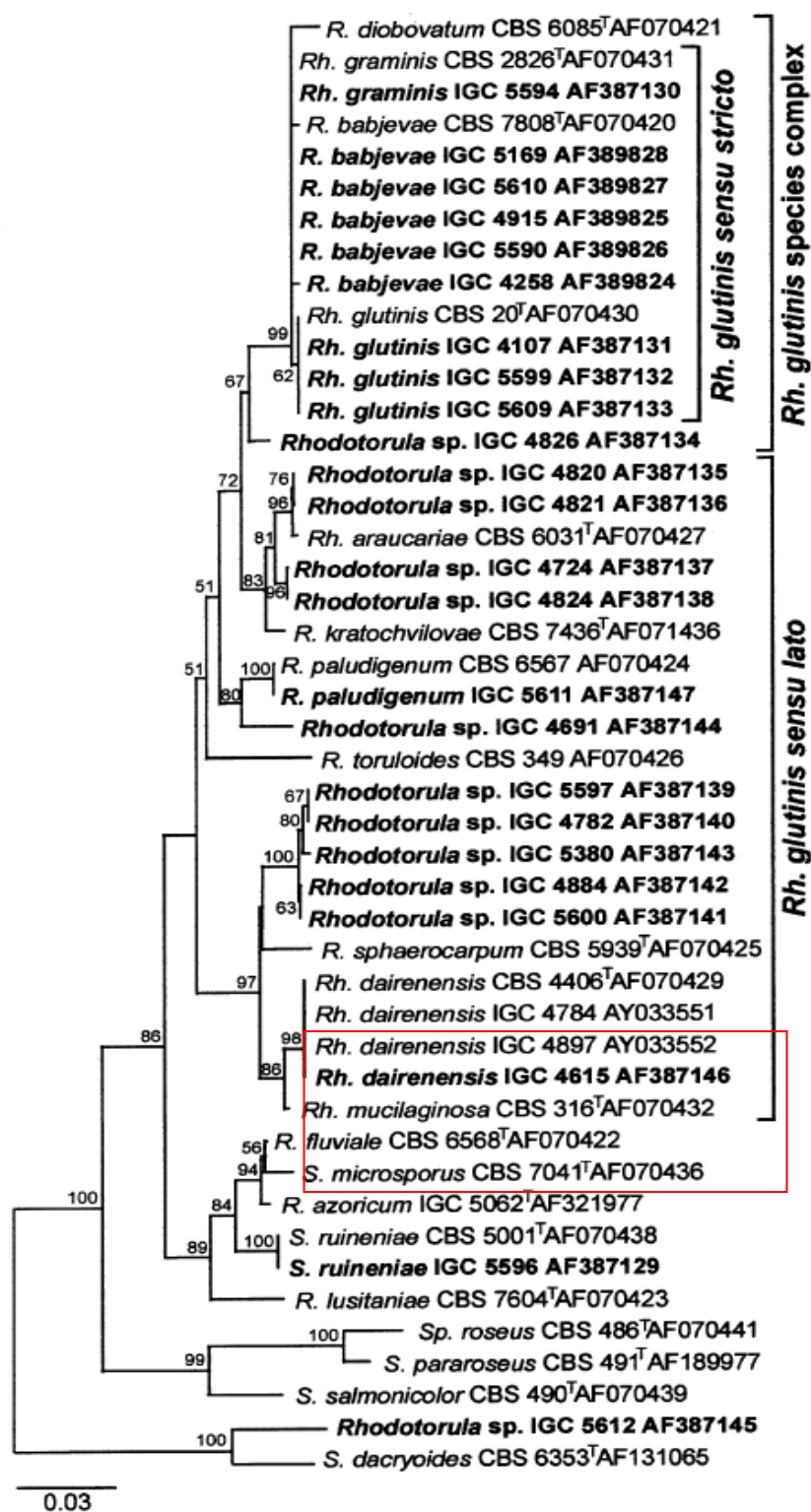


Figura 4 – Árvore filogenética do gênero *Rhodotorula* obtida por análises da região D1/D2 do 26S rDNA (GADANHO; SAMPAIO, 2002).

Nutricionalmente, a *R. dairenensis* se difere da *R. mucilaginosa* pela utilização de nitrato e nitrito, e da *R. glutinis* por sua capacidade de crescer sem vitaminas bem como a 35°C. A espécie pode ser encontrada em diferentes *habitats*, sendo que já foram isoladas de ambientes aquáticos, ar, plantas e solo (GADANHO; SAMPAIO, 2002). Em nosso trabalho a levedura foi isolada do trato intestinal da larva do inseto *Migdolus fryanus*, um besouro parasita da cana-de-açúcar.

Como característica de importância biotecnológica, a *R. dairenensis* produz trealose, um dissacarídeo de glicose (α -D-glicopiranosil-(1 $\rightarrow$ 1)- α -D-glicopiranosida) encontrado em leveduras, micro-organismos termofílicos, algumas plantas e insetos. A trealose é um açúcar não redutor, que além de ser fonte de energia, protege contra estresses ambientais como altas temperaturas, estresse osmótico, escassez de nutrientes e congelamento. Muitos estudos têm correlacionado os níveis de trealose intracelular com o aumento da viabilidade celular, visto que nas camadas lipídicas, protege as membranas contra danos e, na estabilização de proteínas, previne o processo de desnaturação. Essas propriedades sugerem sua ampla aplicação em diversas áreas incluindo a indústria de cosméticos, farmacêutica e agro-alimentar, sendo que as aplicações mais interessantes desse dissacarídeo são na estabilização de vacinas, preservação de enzimas e na estabilização de órgãos para transplante (COLLA et al., 2008). Ainda, recentemente foi caracterizada bioquimicamente uma enzima frutofuranosidase produzida pela *R. dairenensis*. (GUTIÉRREZ-ALONSO et al., 2009).

2.4.5. A levedura *Pseudozyma brasiliensis* sp. nov. GH001

P. brasiliensis sp. nov. é uma levedura basidiomiceto pertencente a ordem Ustilaginales, isolada do trato intestinal de um besouro parasita da cana-de-açúcar pertencente à família *Chrysomelidae*, e que foi descoberta como uma nova espécie neste trabalho (vide Item 5). Notavelmente, membros do gênero

Pseudozyma spp. são capazes de produzir biomoléculas biotecnologicamente interessantes como lipases (BUSSAMARA et al., 2012; BUSSAMARA et al., 2010), xilanases (ADSUL et al. 2009), e biosurfactantes como o lipídio manosil eritritol (KONISHI et al., 2011; MORITA et al., 2010; MORITA et al., 2009).

A *P. brasiliensis* sp. nov. GHG001 teve seu genoma sequenciado e publicado pelo nosso grupo (OLIVEIRA et al., 2013) (vide Anexo I) e tem um grande potencial biotecnológico, visto que é capaz de produzir xilanase (BORGES et al., 2013) (vide Anexo II), propriedade muito comum em vários fungos filamentosos como, *Aspergillus sp*; *Fusarium sp*, *Melanocarpus sp* e *Penicillium sp.*, porém não vista frequentemente em leveduras (ADSUL et al., 2009).

3. Objetivos

3.1. Objetivo geral

Este projeto teve como objetivo a caracterização de genes da via das pentoses de *R. dairenensis* TAB01 e *P. brasiliensis* sp. nov. GHG001., visando a superexpressão destes em cepa industrial de *S. cerevisiae* PE-2 e ainda, objetivou caracterizar bioquimicamente uma enzima xilanase produzida pela *P. brasiliensis* sp. nov. GHG001.

3.2. Objetivos específicos

- a) Identificar os genes *xyl1* e *xyl2* das leveduras;
- b) Caracterizar as leveduras em termos da atividade das enzimas XR e XDH, bem como a expressão de seus respectivos genes (*xyl1* e *xyl2*) em xilose e glicose;
- c) Introduzir os genes envolvidos na assimilação de xilose e superexpressá-los em *S. cerevisiae* (cepa industrial PE-2 e doméstica Sc9721), com sinais regulatórios apropriados;
- d) Analisar as cepas recombinantes quanto às suas capacidades de assimilar/fermentar xilose;
- e) Caracterizar a xilanase produzida pela *P. brasiliensis* sp. nov. GHG001.

4. Metodologia

4.1. Obtenção dos isolados

Os conteúdos intestinais dos insetos colonizadores da cana-de-açúcar crisomelídeo, *Diatrea saccharalis*, *Scaptocoris castanea*, *Migdolus fryanus*, larva de arame, da larva pão de galinha foram inoculados em meio mínimo (MM) líquido (pH 6,5) com 5% de bagaço explodido (Tabela 1) ou 1% de xilose, suplementado com um coquetel de antibióticos (100 µg/mL de ampicilina, 34 µg/mL de kanamicina, 20 µg/mL de tetraciclina e 35 µg/mL de cloranfenicol). O inóculo foi incubado à 30°C sob agitação de 130 rpm para enriquecimento dos micro-organismos. Semanalmente, foram realizadas transferências das culturas para meios frescos.

Tabela 1 – Meio Mínimo (MM).

Reagente	Quantidade
Solução de Sais	50 mL
Fonte de carbono (xilose ou bagaço)	10 ou 50 g
Solução de elementos traços	1 mL
Água destilada q.s.p	Para volume final de 1000 µL

Após a 2ª transferência foi feito o plaqueamento das culturas mistas (crescida com bagaço ou xilose) em MM com 1% de xilose, acrescido de 2 % de agar. As colônias de levedura que cresceram foram re-plaqueadas até que se obtivesse um isolado puro. Estes isolados foram inoculados em meio YPD líquido (Tabela 2) e incubados sob agitação rigorosa a 30°C por 3 dias.

Tabela 2 – Meio YPD.

Reagente	Quantidade
Peptona	20 g
Extrato de levedura	10 g
Dextrose	40 g
Água destilada q.s.p	Para volume final de 1000 µL

4.2. Identificação dos isolados

4.2.1. Extração de DNA *in house*

Para a extração de DNA genômico das leveduras, o inóculo foi centrifugado a 2.500 g por 5 minutos e o *pellet* foi ressuspensionado em 1 mL de tampão de extração de DNA (Tabela 3). A esta suspensão foi adicionado o mesmo volume de pérolas de vidro estéreis, tratadas com ácido nítrico para retirada de impurezas. Os tubos foram acoplados em agitador mecânico e a ruptura celular procedeu-se por 10 minutos. Transferiu-se o sobrenadante (sem as pérolas de vidro) para novos tubos e foi adicionado o mesmo volume de uma mistura fenol-clorofórmio (1:1). Novamente o material foi agitado em agitador mecânico por 10 minutos e em seguida centrifugado a 20.800 g por 20 minutos. À fase aquosa contendo o DNA genômico foi adicionado 0,6 volume de isopropanol para precipitação do DNA, seguido de centrifugação a 20.800g por 20 minutos. O sobrenadante foi descartado e o DNA foi lavado com 150 µL de etanol 70% (v/v) gelado. O volume foi novamente centrifugado a 20.800 g por 5 minutos, o sobrenadante foi descartado e o *pellet* foi seco à temperatura ambiente. Por fim, o DNA foi ressuspensionado em 50 µL de água estéril + 1 µL de RNase (10 mg/mL) e incubado a 37°C por 12 horas (SAMBROOK; RUSSEL, 2001).

Tabela 3 – Tampão de extração de DNA genômico.

Reagente	Quantidade
Tris HCl pH 8,5	31, 52 g
NaCl	14,61 g
EDTA	7, 30 g
SDS	5 g
Água destilada q.s.p	Para volume final de 1000 µL

4.2.2. Amplificação e sequenciamento da região ITS2

Para a identificação dos isolados, primeiramente foi realizada a amplificação por PCR da região espaçadora interna transcrita 2 (ITS2) do DNA ribossomal. As reações de PCR estão descritas na tabela 4, sendo utilizados os oligonucleotídeos ITS3 (5'- GCATCGATGAAGAACGCAGC -3') e ITS4 (5'- TCCTCCGCTTATTGATATGC- 3'). A reação de amplificação ocorreu em termociclador com a seguinte ciclagem: o DNA molde foi inicialmente desnaturado a 94°C por 3 minutos, seguido de 35 ciclos de desnaturação a 94°C por 30 segundos, anelamento dos oligonucleotídeos a 56°C por 30 segundos e extensão das novas moléculas de DNA a 72°C por 1 minuto. Por fim, um ciclo de extensão final a 72°C por 10 minutos. Os fragmentos foram analisados em eletroforese em gel de agarose 1% em tampão TAE 1X, a 90 V por 45 minutos. Após a corrida eletroforética, o gel foi corado em solução de brometo de etídeo (0,1 mg/mL) por 15 minutos e visualizado sob luz ultravioleta no fotodocumentador Gel Logic 2200 Pro (Carestream). A determinação do tamanho dos fragmentos foi feita comparando as bandas obtidas com o *ladder* de 1 Kb (Fermentas). Os produtos das reações de PCR foram purificados com o kit *GFX PCR DNA and Gel Band Purification* (GE Healthcare), segundo as recomendações do fabricante, e sequenciados utilizando o kit *Big Dye® Terminator v3.1 Cycle Sequencing* (Applied Biosystems/Life Technologies), seguindo o protocolo do fabricante, no sequenciador Applied Biosystems® 3500xL Genetic Analyzers do Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE). As sequências obtidas

foram comparadas com as depositadas no banco de dados do *National Center for Biotechnology Information* (NCBI), através do programa Blast (Basic Local Alignment Search Tool). Os parâmetros *e-value* e identidade forneceram o grau de confiabilidade da identificação.

Tabela 4 – Reação de PCR.

Reagentes	Quantidade
H ₂ O milliQ autoclavada q.s.p	20 µL
Buffer KCl (10x)	2,0 µL
MgCl ₂ (25mM)	1,6 µL
dNTP (10mM)	0,8 µL
Primer Fw (10 pmol/µL)	0,4 µL
Primer Rv (10 pmol/µL)	0,4 µL
Taq DNA polimerase (5 U/ µL)	0,1 µL
DNA molde	1 µL

4.2.3. Construção de árvores filogenéticas através das regiões ribossomais ITS, 5.8S e D1/D2

Objetivando uma identificação dos isolados mais curada e confiável, foi inferida a árvore filogenética com a sequência das regiões ribossomais ITS, 5.8S e D1/D2 das leveduras *Rhodotorula dairenensis* TAB01 e *Pseudozyma brasiliensis* GHG001. Para isso, foi realizada a extração do DNAs (vide 3.1) e a amplificação das regiões ribossomais 28S (LSU, mais especificamente a região D1/D2) e das regiões ITS1, 5.8S e ITS2. Essas regiões vêm sendo muito utilizadas na diferenciação de fungos através de análises de PCR e a região ITS vem se mostrando como um *barcode maker* (MARTIN; RYGIEWICZ, 2005; SHOCH et al., 2012). A região ITS inclui as regiões ITS1 e ITS2, separadas pelo gene 5.8S, e está situado entre os genes 18S (SSU) e 28S (LSU) na unidade de repetição nrDNA (Figura 5). Os oligonucleotídeos utilizados estão descritos na tabela 5.

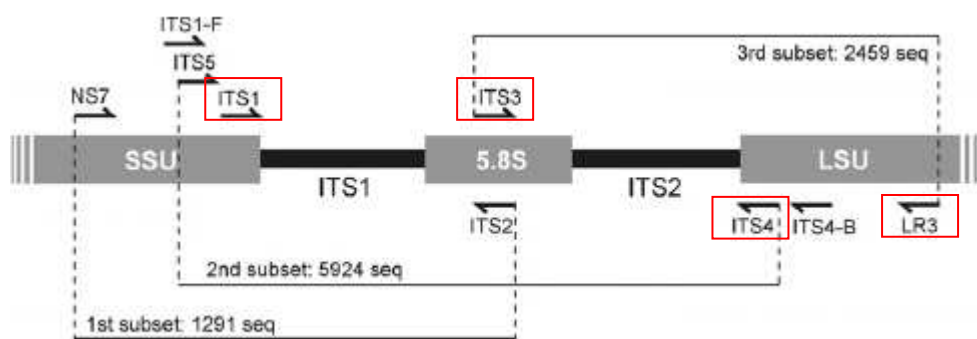


Figura 5 – Esquema mostrando a localização de alguns oligonucleotídeos utilizados para amplificar as regiões ITS e D1/D2 (BELLEMAIN et al., 2010).

Tabela 5 – Lista de oligonucleotídeos utilizados na identificação das leveduras.

Oligonucleotídeo	Sequência (5'-3')
ITS1	TCCGTAGGTGAACCTGCGG
NSI1	GATTGAATGGCTTAGTGAGG
ITS4	TCCTCCGCTTATTGATATGC
ITS3	GCATCGATGAAGAACGCAGC
LR3	GGTCCGTGTTTCAAGACGG

A PCR foi feita de acordo com a tabela 4, com a seguinte ciclagem: um ciclo inicial de 94°C/3 min de desnaturação, seguido de 35 ciclos de 94°C/30 segundos, anelamento – temperatura variável (vide tabela 6)/30 segundos e extensão – 72°C/ tempo variável (vide tabela 6); e extensão final de 72°C/10 min.

Tabela 6 – Combinações de oligonucleotídeos e as respectivas temperaturas de anelamento e tempos de extensão.

Oligonucleotídeos	Temperatura de anelamento	Tempo de extensão
ITS3-LR3	59°C	2 min
ITS1-ITS4	52°C	1 min
NSI1-ITS4	55°C	1,5 min

Depois de a amplificação ter sido confirmada por eletroforese em gel de agarose 1%, os produtos de PCR foram purificados com o kit *GFX PCR DNA and*

Gel Band Purification (GE Healthcare Life Sciences) segundo o protocolo sugerido pelo fabricante, e posteriormente sequenciados utilizando o kit de reação *Big Dye® Terminator v3.1 Cycle Sequencing* (Applied Biosystems/Life Technologies). As amostras foram analisadas no sequenciador Applied Biosystems® 3500xL Genetic Analyzer (Applied Biosystems/Life Technologies). As sequências obtidas foram trimadas por qualidade (phred score 20) utilizando o programa phred (EWING et al., 1998). Os *contigs* foram montados utilizando o programa *online* CAP3 (HUANG; MADAN, 1999). A partir dos *contigs*, as sequências de ITS1 e ITS2 foram extraídas utilizando ITSx V 1.0.3 (BENGTTSSON-PALME et al., 2013), assim como o posicionamento da região 5.8S. Para cada região foi realizado um alinhamento múltiplo estruturado, utilizando MXSCARNA V 2 (TABEL et al., 2008). Os alinhamentos múltiplos foram inicialmente separados da estrutura secundária, e concatenados em um único arquivo utilizando FASconCAT (KÜCK; MEUSEMANN, 2010). A inferência das filogenias particionadas foi realizada utilizando RAxML V 7.5.7 (STAMATAKIS, 2012), e a construção da árvore foi realizada utilizando FigTree V 1.4 (RAMBAUT, 2007). As espécies tipo utilizadas para a construção das árvores filogenéticas foram retiradas do banco de dados do CBS-KNAW *Fungal Biodiversity Centre*. A taxa de substituição foi obtida usando o gamma-modelo de Yang (YANG, 1996). Valores de *bootstrap* foram computados acima de 1000 repetições.

Para confirmar a identificação da espécie, foi verificado o número de *gaps* entre a nossa sequência e a sequência reportada na literatura através do programa Water (RICE et al., 2000).

4.3. Perfil de crescimento das leveduras em xilose

Para verificar o crescimento das leveduras de interesse na presença de xilose foi realizado ensaio de curva de crescimento com a *Rhodotorula dairenensis* TAB01 e *Pseudozyma brasiliensis* sp. nov. GHG001, em meio de cultura mínimo (0,67% YNB) com diferentes fontes e concentrações de carbono (1% glicose e

0,5%, 1% e 5% xilose). Como controle, também foi realizado o mesmo ensaio com as leveduras *S. cerevisiae* 9721 (controle negativo) e *Pichia stipitis* NRRL-Y7124 (controle positivo). 10^6 leveduras/mL foram inoculadas em 12,5 mL de cada meio, em triplicata, em erlenmeyer de 125 mL, e incubadas à 30°C a 250 rpm. A cada 24 horas foi coletado 1 mL de cada cultura para determinação espectrofotométrica da densidade óptica (OD) utilizando frações de absorbância no comprimento de onda de 595 nm. Foram coletados os pontos 24, 48, 72 e 96 horas.

4.4. Caracterização molecular dos genes da *Rhodotorula dairenensis* TAB01 e da *Pseudozyma brasiliensis* sp. nov. GHG001

4.4.1. Extração de DNA

Primeiramente, foi feito um pré-inóculo da levedura em meio YPD (Tabela 2), a 30°C, sob agitação de 200 rpm. Após 24 horas de crescimento, o *pellet* da levedura foi coletado e realizada a extração de DNA genômico com o kit “DNeasy Plant Mini” (Qiagen), com a adaptação de que a lise celular foi feita com pérolas de vidro em *vortex* durante dez minutos. O restante do procedimento foi realizado segundo o protocolo sugerido pelo fabricante.

4.4.2. Isolamento do RNA e síntese de cDNA

Para o isolamento do RNA, a levedura foi inoculada em uma concentração de 10^6 leveduras/mL em 130 mL dos meios e condições utilizadas para a curva de crescimento (vide 3.2). Após cada coleta a amostra foi centrifugada por 10 minutos a 4000 rpm e o *pellet* foi guardado a -80°C até o momento da extração de RNA. O RNA total foi extraído com kit “RNeasy Plant Mini” (Qiagen), segundo o protocolo “Purification of Total RNA from Yeast” pelo método de lise mecânica com a seguinte adaptação: a lise foi feita com pérolas de vidro em *vortex* durante quinze minutos, alternando a cada 3 minutos no *vortex*, 2 minutos no gelo. Para

quantificação do RNA foi utilizado o espectrofotômetro Nanodrop (Thermo Scientific) e para verificar a integridade do RNA, o 2100 BioAnalyzer (Agilent). Em seguida, foi realizado o tratamento com DNase com o kit *TURBO DNA-free* (Life Technologies), a fim de se eliminar qualquer resíduo de DNA na amostra. Após confirmar a ausência de DNA através de PCR utilizando como molde o RNA (vide protocolo de PCR 3.1.1), foi realizada a síntese de cDNA (partindo da mesma quantidade de RNA) utilizando a *SuperScript II Reverse Transcriptase* (Life Technologies), de acordo com as recomendações do fabricante.

4.4.3. Identificação molecular dos genes da xilose redutase (*xyl1*) e xilitol desidrogenase (*xyl2*) da *Rhodotorula dairenensis* TAB01

Como as sequências de *xyl1* e *xyl2* da *R. dairenensis* TAB01 não eram conhecidas (ou pelo menos não estavam disponíveis nos bancos de dados), foram desenhados oligonucleotídeos degenerados para estes genes baseados nas sequências da *Rhodotorula graminis* e *Rhodotorula mucilaginosa*, obtidas no banco de dados do *National Center for Biotechnology Information* (NCBI - <http://blast.ncbi.nlm.nih.gov/Blast.cgi>) (Tabela 7). Após um alinhamento dessas sequências realizada no programa *ClustalW* (LARKIN et al., 2007), foram desenhados oligonucleotídeos para as regiões mais conservadas. Para o gene *xyl2*, os oligonucleotídeos não funcionaram adequadamente, então foram utilizados os descritos por Xu e colaboradores (2011) o qual estudou o metabolismo de xilose das cepas *R. graminis* e *R. mucilaginosa*.

Utilizando o cDNA sintetizado, foi feita a reação de PCR (vide 3.3.1) variando a concentração de MgCl₂ em 1,5 mM, 3 mM e 4,5 mM e as temperaturas de anelamento em 50°C, 53°C e 57°C, para o XR e, 50°C e 52°C, para o XDH. Em seguida, foi realizada a eletroforese, purificação e sequenciamento dos produtos de PCR (vide 3.3.4) que tinham o tamanho esperado, e as sequências foram comparadas com as depositadas no NCBI.

Uma vez que foi obtida uma região de cada gene de *R. dairenensis* TAB01, foi utilizada a técnica de RACE para amplificar as bordas 5' e 3' dos mesmos, utilizando o kit *SMARTer RACE cDNA Amplification* (Clontech), segundo protocolo do fabricante. Os oligonucleotídeos desenhados e utilizados estão descritos na Tabela 7. Os amplificadores obtidos foram confirmados por sequenciamento (vide 3.3.4), e as sequências completas das regiões codantes de cada gene foram montadas utilizando o programa *Geneious* (<http://www.geneious.com>).

Tabela 7 – Lista de oligonucleotídeos utilizados na identificação dos genes *xyl1* e *xyl2* de *R. dairenensis* TAB01.

Oligonucleotídeo	Sequência do oligo (5'-3')	Referência
Xyl1- Fw	AGTACCTCGACCTSTACCTCAT	Este estudo
Xyl1 – Rv	KMGAYGAYCTTGAAGTTCGAKWYGAT	Este estudo
Xyl2 – Fw	AACGTGGTKGTSTTTGGYG	Este estudo
Xyl2 – Rv	CKGTGMGWGATGAGCGAYTTGAGG	Este estudo
Xyl2 – Fw	CCMATGGTCYTSGGNCACGA	(XU et al., 2011)
Xyl2 – Rv	CCGACVGGVCCDGCDCCAAAGAC	(XU et al., 2011)
RACE_xyl1_3'Fw	CCCGACCGTCAAGAAGATCGCCGAGG	Este estudo
RACE_xyl1_5'Rv	GGGAAGAGGTTTCGAGTGGGGGTCTG	Este estudo
RACE_xyl2_3'Fw	CTCGCTGGCCGCTATGTCTCCC	Este estudo
RACE_xyl2_5'Rv	CGCTCGTAGTGACCACCCTTGCACTCG	Este estudo

4.4.4. Purificação, ligação, clonagem e sequenciamento

A região do gel que continha o fragmento com o tamanho do amplificado esperado foi excisada e submetida à purificação utilizando o kit *GFX PCR DNA and Gel Band Purification* (GE Healthcare). Os produtos amplificados foram clonados no vetor pGEM®-T Easy (Promega) (Figura 6) e transformados por eletroporação (Multiporador® system, Eppendorf) em células de *E. coli* DH5α. Colônias brancas foram selecionadas de placas de LB (Tabela 8) com ampicilina (100 µg/mL) e procedeu-se uma reação de PCR de colônia com os

oligonucleotídeos universais M13-Fw (5'- GACCGGCAGCAAAATG -3') e M13-Rv (5'- CAGGAAACAGCTATGAC -3') para verificar a efetividade da clonagem. As colônias que tiveram a confirmação de que seus plasmídios possuíam o fragmento de interesse foram inoculadas em 3 mL de LB + ampicilina (100 mg/mL) e incubadas sob agitação (~200 rpm) a 37°C *overnight*. Em seguida os DNAs plasmidiais dos clones positivos foram extraídos utilizando-se o kit *QIAprep Spin Miniprep* (Qiagen) e os insertos foram sequenciados com os primers universais M13 utilizando o kit de reação *Big Dye® Terminator v3.1 Cycle Sequencing* (Applied Biosystems/Life Technologies) no sequenciador Applied Biosystems® 3500xL Genetic Analyzer (Applied Biosystems/Life Technologies), afim de obter sequência com qualidade desde o início do gene (onde está o ATG).

Tabela 8 – Meio LB.

Reagente	Quantidade
NaCl	10 g
Triptona	10 g
Extrato de levedura	5 g
Ágar	12 g
Água destilada q.s.p	Para volume final de 1000 µL

4.4.5. Identificação molecular dos genes da xilose redutase (*xyl1*) e da xilitol desidrogenase (*xyl2*) de *Pseudozyma brasiliensis* sp. nov. GHG001

Do mesmo modo que para a *R. dairenensis* TAB01, foi realizado o alinhamento global do gene *xyl1* de leveduras que tinham suas sequências depositadas no banco de dados do NCBI, sendo elas *S. stipitis*, *S. cerevisiae*, *Ogataea siamensis*, *R. graminis*, *R. mucilaginosa*, *Kluyveromyces lactis*, *Pichia guilliermondii*, *Spathaspora passalidarum* e *Pachysolen tannophilus*. Para isso, foi desenhado um par de oligonucleotídeos degenerados baseados em regiões conservadas dos genes dessas leveduras. O mesmo foi feito para o gene *xyl2*,

sendo que nesse caso o alinhamento e desenho dos oligonucleotídeos foram baseados nas leveduras *S. stipitis*, *S. cerevisiae*, *R. mucilaginosa*, *R. graminis*, *Arxula adeninivorans*, *Cryptococcus neoformans* e *Pichia angusta*.

Tabela 9 – Oligonucleotídeos usados para amplificação de *xyl1* e *xyl2* da *P. brasiliensis* GHG001.

Oligonucleotídeos	Sequência (5'-3')
Universal xyl1-Fw	GABTAYBTHGACHTVTTYTNATBCA
Universal xyl1-Rv	TRHGGRTGRTGTTTCRAYTTG
Universal xyl2-Fw	CCHMTSGTBYTVGGHCAYGA
Universal xyl2-Rv	ACNGANADKGGCTCNABNAGN

Foi feita a extração do DNA genômico (vide 3.5.1) da *P. brasiliensis* GHG001 e posteriormente, foi realizada a reação de PCR (vide 3.3.1) com diferentes concentrações de MgCl₂ (1,5 mM; 3,0 mM e 4,5 mM) e utilizando *enhancers* (5% DMSO e 1% Triton X-100) nas temperaturas de anelamento 45°C, 47°C, 50°C, 52°C e 55°C. Além disso, os oligonucleotídeos degenerados *xyl1* e *xyl2* desenhados para *R. dairenensis* TAB01 (Tabela 7) também foram testados em *P. brasiliensis* GHG001.

Devido ao insucesso das amplificações utilizando esses oligonucleotídeos, optou-se por fazer a busca dos genes de interesse no genoma da *P. brasiliensis* GHG001 que havia sido recentemente sequenciado (vide item 6). Para o gene *xyl2* foi realizada uma busca no genoma da *P. brasiliensis* GHG001 a partir de sequências conhecidas desse gene de outras leveduras. Foi obtida apenas uma região de DNA similar ao gene de interesse, para o qual foi desenhado oligonucleotídeos para amplificação das bordas 5' e 3' do mesmo (Tabela 10), utilizando a técnica de RACE com o kit *SMARTer RACE cDNA Amplification* (Clontech), segundo protocolo do fabricante. Os produtos amplificados foram purificados, clonados no vetor pGEM-T e sequenciados conforme o item 4.3.4, a fim de se obter uma sequência com qualidade desde o ATG do gene.

Tabela 10 – Oligonucleotídeos utilizados no RACE para identificação do gene *xyl2* da *P. brasiliensis* GHG001.

Oligonucleotídeo	Sequência (5'-3')
RACE_Pb_xyl2_3'Fw	TTCGTCCAGGTCGGCATGGGCAG
RACE_Pb_xyl2_5'Rv	TCAAGCAGACGGGGCAGGAGCGACA

Como a xilose redutase apresenta alto similaridade com outras redutases, não obtivemos sucesso com a mesma abordagem utilizada com o *xyl2*, então foi realizada uma busca deste gene nos proteomas de dez leveduras (*Pseudozyma antarctica* T-34, *Pseudozyma hubeiensis* SY62, *Saccharomyces cerevisiae* M3707-1, *S. cerevisiae* S288C, *Rhodotorula graminis*, *Rhodotorula glutinis* strain ATCC 204091, *Pichia stipitis*, *Pichia pastoris*, *Kluyveromyces lactis*, *Spathaspora passalidarum* NRRL Y-27907) e do fungo *Ustilago maydis* nos bancos de dados do JGI (GRIGORIEV et al., 2012) e do UniProt (www.uniprot.org). Um HMM (Modelo Oculto de Markov) específico para busca de domínios de aldeído redutases foi obtido do banco de famílias de proteínas PFAM (PUNTA et al., 2012). Assim, a busca foi realizada nos proteomas com o HMM (PF00248), utilizando *hmmsearch* do HMMER3 (EDDY, 1998), e as sequências de proteínas completas referentes a esses domínios encontrados foram alinhadas com MAFFT (KATO; TOH, 2008). O alinhamento foi utilizado como entrada para o ProtTest (ABASCAL, ZARDOYA, POSADA, 2005), o qual estima o melhor modelo evolutivo a ser utilizado. O alinhamento em fasta foi então convertido em phylip, para a inferência da árvore filogenética com RAxML (<http://www.exelixis-lab.org/>). Em seguida, foi verificado o clado onde as xilose redutases conhecidas estavam presentes, e onde a sequência da proteína em *P. brasiliensis* sp. nov. GHG001 deveria estar inserida.

4.5. RT-PCR em Tempo Real (qRT-PCR)

Para avaliar o perfil da expressão dos genes *xyl1* e *xyl2* de *R. dairenensis* TAB01 e *P. brasiliensis* GHG001, quando crescidos em xilose e em glicose,

realizou-se o PCR em Tempo Real (qRT-PCR). O qRT-PCR é uma técnica extremamente sensível que permite a detecção em tempo real da amplificação e uma medição precisa (GACHON; MINGAM; CHARRIER, 2004). O sistema utilizado do tipo Sybr® Green (Life Technologies) consiste em um corante fluorescente intercalar que se liga a dupla fita de DNA. O acúmulo do produto de PCR ao longo dos ciclos de amplificação resulta no aumento da intensidade da fluorescência através da qual é possível quantificar a geração dos *amplicons*. No PCR quantitativo utilizando o corante Sybr® Green, a especificidade do produto da reação de PCR pode ser verificada plotando o sinal da fluorescência em função da temperatura para gerar uma curva de *melting* do *amplicon*. Um pico característico da temperatura de *melting* do DNA amplificado irá distingui-lo dos artefatos de amplificação que se formam geralmente em temperaturas mais baixas (HUGGETT; BUSTIN 2011).

Os oligonucleotídeos para *xyl1* e *xyl2* de ambas as leveduras foram desenhados utilizando o programa *Primer Express* (Tabela 10), com os seguintes parâmetros: tamanho do oligonucleotídeo de 18-24 nt; conteúdo G-C de 40-60%; Tm 50-60°C e tamanho do *amplicon* 80-200 pb.

Para normalizar a quantidade de RNA utilizada na qRT-PCR é necessário um gene de referência (*house keeping gene*), e para isso foi escolhido o gene do fator de alongação (*tef*) (NARDI et al., 2010). O gene de referência é o controle interno da reação de qRT-PCR e deve apresentar um nível de expressão que não seja afetada pelos fatores experimentais, com variação mínima na sua expressão em diferentes estados fisiológicos do organismo. O gene de referência deve, por sua vez, corrigir a variabilidade resultante das imperfeições da tecnologia utilizada e também dos procedimentos preparatórios, garantindo que qualquer variação na quantidade cDNA utilizada seja relativa na mesma proporção tanto nas reações de qRT-PCR dos genes alvo como na reação do gene de referência (KOZERA; RAPACZ, 2013). Como a sequência do *tef* não era conhecido para a *R. dairenensis*, foram utilizados os oligonucleotídeos desenhados para sequência do gene da *R. mucilaginosa* (oligonucleotídeos 150F e 977R). Foi feita então a PCR

(vide 3.3.1) do DNA e cDNA da *R. dairenensis* TAB01 utilizando esses oligonucleotídeos e o produto foi sequenciado. Utilizando como molde a sequência de *tef* obtida para *R. dairenensis* TAB01, foram desenhados novos oligonucleotídeos para os ensaios de PCR em tempo real, com o mesmo programa e parâmetros utilizados para os genes *xyl* (Tabela 11). Para obtenção do gene *tef* da *P. brasiliensis* GHG001 foi realizado um *megablast* contra o genoma da própria levedura.

As reações de PCR em tempo real foram realizadas utilizando o fluoróforo *SYBR-Green* (Life Technologies), seguindo as recomendações do fabricante. Para a *R. dairenensis* TAB01, foi utilizado como molde o cDNA (vide 3.3.2) diluído 100x para o gene *xyl1* e *tef*, e o cDNA não diluído para *xyl2*. Para a *P. brasiliensis* GHG001, foi utilizado como molde o cDNA (vide 3.3.2) diluído 100x para ambos os genes (*xyl1*, *xyl2* e *tef*).

As amplificações foram feitas no equipamento “Applied Biosystems ViiA™ 7” com a ciclagem: 95°C – 10 min e 40 ciclos de 95° - 15 s e; 60°C – 1 min. Após os 40 ciclos, é realizada a curva de *melting* utilizando a seguinte ciclagem: 95°C – 15 s / 60°C – 1 min / 95°C – 15 s.

Tabela 11 – Oligonucleotídeos utilizados na qRT-PCR.

Oligonucleotídeo	Sequencia do oligo (5'-3')
<i>R. mucilaginosa</i>	
150F	GTCATCGTCGCCATCAACAAGATG
977R	ATGAACTTGGGGAGGTCCTCGA
<i>R. dairenensis</i> TAB01	
RT_Xyl1-Fw	AGCCGCTTGTCAACCAGATC
RT_Xyl1-Rv	AATGTTCTTCTCCTTGTGGTGCTT
RT_Xyl2-Fw	CAAGGTGGGTAAGAATGTCAAGA
RT_Xyl2-Rv	AGCAGGATCGGCACGACTT
RT_tef-Fw	TCAAGAAGGTCGGCTTCAA
RT_tef-Rv	CGCCTCGATCATGTTGTCA
<i>P. brasiliensis</i> GHG001	
q-Xyl1-Fw	CGAATTTCAACCACCGATGAC
q-Xyl1-Rv	GTGAGCCTCGATCTGGTTGA
q-Xyl2-Fw	ATGGTGTGGGTACGAATC
q-Xyl2-Rv	ACACGATCCCCGACCTTGT
q-tef-Fw	GACAACCATACCGCCCTTGA
q-tef-Rv	ATGTCTACAAGATCGGTGGTATCG

4.6. Atividade enzimática de XR e XDH endógenos

Foram inoculadas 10^6 leveduras/mL em 50 mL dos meios com xilose utilizados para curva de crescimento, tamponado com 50 mM de hidrogenofterato de potássio (pH 5,5) e incubadas a 30°C/250 rpm, por 48 horas (fase exponencial de crescimento). As células foram então centrifugadas, lavadas com água estéril e lisadas com o reagente Y-PER “Yeast Protein Extraction Reagent” (Thermo Scientific) para a obtenção dos extratos celulares, seguindo recomendação do fabricante. Como controle utilizou-se a *S. stipitis* crescida por 24 horas, em meio 5% xilose.

As atividades de XR e XDH foram quantificadas segundo Rizzi e colaboradores (1989), com algumas modificações. Para a medição das atividades enzimáticas foram utilizadas as soluções descritas na Tabela 12. A atividade da

XR foi aferida através do decréscimo da absorbância a 340 nm (em espectrofotômetro Evolution 60, Thermo Scientific), devido à oxidação do NADPH ou NADH utilizando a xilose como substrato limitante, e a atividade da XDH pelo aumento da absorbância a 340 nm devido à redução do NAD⁺ e uso de xilitol como substrato limitante.

Tabela 12 – Componentes das reações para medição das atividades enzimáticas de XR e XDH.

Reagentes	XR (μL)	XDH (μL)
Tampão trietanolamina 0,5M pH 5,5*	100	100
NADPH ou NADH 20 mM**	10	-
NAD ⁺ 30 mM**	-	10
Xilose 3,5 M*	100	-
Xilitol 3 M*	-	100
Extrato celular**	Variável	Variável
Água*	Para completar 1 mL de reação	Para completar 1 mL de reação

* Mantidos a 30°C; ** Mantidos no gelo antes de iniciar a reação.

As atividades específicas foram calculadas em U.mg proteína⁻¹, e uma unidade (U) de atividade enzimática é definida como a quantidade de enzima que catalisa a redução/oxidação de 1mmol de coenzima por minuto. As análises de proteínas totais dos extratos celulares foram realizadas pelo método de Bradford (1976).

4.7. Clonagem de *xyI1* e *xyI2* no vetor de expressão heteróloga pSP-G1

Uma vez obtida a sequência dos genes e o mesmos terem sido clonados no vetor pGEM®-T Easy (vide 3.5.4), foi iniciada a clonagem no vetor de expressão pSP-G1 (Figura 6), desenvolvido por Partow e colaboradores (2010). O vetor pSP-G1 é específico para *S. cerevisiae* e foi escolhido por possuir

promotores bidirecionais fortes (TEF1 e PGK1), que possibilitam a expressão simultânea de dois genes diferentes em um único vetor, portanto, um sistema mais apropriado para a expressão de vias metabólicas. Para a clonagem do gene inteiro, foram desenhados oligonucleotídeos com sítios de restrição de enzima presentes em cada sítio múltiplo de clonagem (MSC) (Tabela 13).

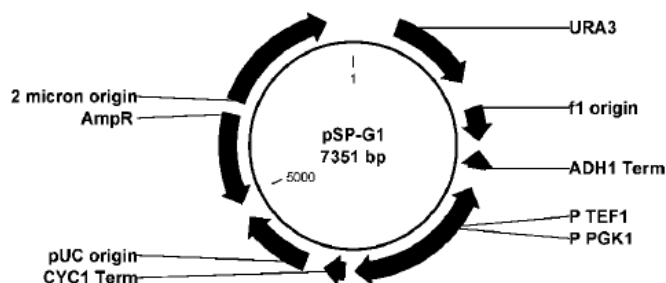


Figura 6 – Vetor pSP-G1 (PARTOW et al., 2010).

Tabela 13 – Oligonucleotídeos utilizados na clonagem dos genes *xyI* e *xyI2*.

Oligonucleotídeo	Sequencia (5'-3')	MSC e promotor	Enzima
<i>Rhodotorula dairenensis</i> TAB01			
Clo_xyI1_Fw	TAT GGATCCT ATG GGGACCACAGTCTCG	MSC2 – pGK1	BamH I
Clo_xyI1_Rv	CAT GCTAGCT TTACTGGATCTTGACCTGG TAC	MSC2 – pGK1	Nhe I
Clo_xyI2_Fw	AGG GCGGCCGCA ATG GCCAGCAGCAAC ACC	MSC1 – pTEF1	Not I
Clo_xyI2_Rv	ACC AGATCT TTACGCAGTGTCTGGACTCG GT	MSC1 – pTEF1	Bgl II
<i>Pseudozyma brasiliensis</i> GHG001			
Clo_xyI1_Fw	TAT GGATCCG ATG TCGCTCAACAAGACT C	MSC2 – pGK1	BamH I
Clo_xyI1_Rv	CAT GCTAGCT TAGTTGCCACGGTTGACC TT	MSC2 – pGK1	Nhe I
Clo_xyI2_Fw	AGG GCGGCCGCA ATG CCGCTCCAGTC GCGCCCAA	MSC1 – pTEF1	NotI
Clo_xyI2_Rv	ATT TTAATTAA CTACTGCTGGCCCTCGCC CTGGCT	MSC1 – pTEF1	PacI

*Em vermelho, o sítio de restrição da enzima, em negrito o códon de iniciação.

4.7.1. Clonagem dos genes da *R. dairenensis* TAB01

Inicialmente, tanto o vetor pSP-G1 quanto o vetor *xyl1*-pGEM-T, foram digeridos utilizando 1 µg de cada vetor, com 1 µL das enzimas BamHI (10U) e NheI (10U) (Fermentas), 5 µL *Buffer* Tango (10x), em um volume final de 50 µL (em duplicata). A reação foi incubada a 37°C por 2 horas e inativada a 65°C por 20 minutos.

As regiões do gel com os fragmentos de interesse foram excisadas e purificadas com o kit *GFX PCR DNA and Gel Band Purification* (GE Healthcare). Para promover a ligação entre as extremidades coesivas do inserto e do vetor pSP-G1, foi utilizada a enzima T4 DNA ligase (Promega) e a razão molar de DNA 1:1 (vetor: inserto). A reação foi incubada a temperatura ambiente por 2 horas e em seguida a ligação foi transformada em células de *E. coli* DH5α. Para verificar a efetividade da ligação, foi feita PCR de colônias selecionadas de placas com LB + Ampicilina (100 µg/mL) com os oligonucleotídeos de clonagem utilizados para a amplificação dos genes (Tabela 13). Dos clones positivos foram extraídos os DNAs plasmidiais com o kit *QIAprep Spin Miniprep* (Qiagen), que foram sequenciados com o objetivo de conferir o sucesso da clonagem no gene vetor pSP-G1.

Tendo sido confirmado o sucesso da clonagem do gene *xyl1* no vetor pSP-G1, foi realizada a clonagem do gene *xyl2*. Para o gene *xyl2* o processo foi o mesmo, com a diferença que as digestões dos vetores *xyl2*-pGEM-T e *xyl1*-pSP-G1 foram realizadas com as enzimas NotI e BglII (Fermentas) e o *Buffer* O. O plasmídeo contendo os genes *xyl1* e *xyl2* de *R. dairenensis* TAB01 foi denominado pSP-G1-Rd.

4.7.2. Clonagem dos genes de *P. brasiliensis* GHG001

Do mesmo modo que a *R. dairenensis* TAB01, tanto o vetor pSP-G1 quanto o vetor *xyl1*-pGEM-T, foram digeridos, utilizando 1 µg de cada vetor, com 1 µL das

enzimas BamHI (10U) e NheI (10U) (Fermentas), 5 µL *Buffer* Tango (10x), em um volume final de 50 µL (em duplicata). A reação foi incubada a 37°C por 2 horas e inativada a 65°C por 20 minutos. Os procedimentos de ligação e confirmação da clonagem foram realizados como para os genes de *R. dairenensis* (item 3.7.1).

Para o *xyl2* da levedura *P. brasiliensis* GHG001 foi realizada uma digestão sequencial, pois não havia um único *Buffer* ideal para as duas enzimas, NotI (Fermentas) e PaeI (NEB). Assim, primeiramente fez-se a digestão com a enzima PaeI cujo *Buffer* continha menor quantidade de sais (NEBuffer 1) e em seguida foi feita uma segunda digestão com a enzima NotI e o *Buffer* O (Fermentas). O plasmídio contendo os genes da *P. brasiliensis* GHG001 foi denominado pSP-G1-Pb.

4.8. Construção das cepas de *S. cerevisiae* recombinantes

Para a construção de uma cepa industrial recombinante foram utilizadas as cepas industriais brasileiras PE-2 e a PE-2 XKS1, cepas que contem o gene *ura3* deletado o qual será utilizado como marcador de seleção. Além disso, a PE-2 XKS1 apresenta o gene *xks1*, que codifica a enzima xilose-quinase, hiper-expresso através da substituição do promotor (*Xks1p*) pelo promotor forte *Hxt7p*, em um de seus alelos. As cepas e os plasmídios utilizados estão listados na tabela 14. Os plasmídios centroméricos pRH 209, pRH 274 e p416, gentilmente cedidos, foram usados como controle. O p416, como um controle negativo por ser um plasmídio vazio, o pRH 209 por conter os genes *xyl1* e *xyl2* da *P. stipitis* e, o pRH 274 por apresentar a mesma construção do pRH 209 e, adicionalmente, o gene *xks1* hiper-expresso.

Tabela 14 – Plasmídios e cepas utilizadas neste estudo.

Cepa	Genótipo relevante	Referência
PE-2	Δ URA3	*
PE-2 XKS1	Δ URA3, P _{HXT7} -XKS1	**
Plasmídeo	Descrição	Referência
pSP-G1-Rd	pSP-G1+P _{PGK1} -RdXYL1-T _{PGK1} ;P _{TEF1} -RdXYL2-T _{TEF1}	Este estudo
pSP-G1-Pb	pSP-G1+P _{PGK1} -PbXYL1-T _{PGK1} ;P _{TEF1} -PbXYL2-T _{TEF1}	Este estudo
pSP-G1	pSP-G1 contendo os promotores pTEF1 e pPGK1	(PARTOW et al., 2010)
pRH 209	pRS416+(P _{ADH1} -PsXYL2-T _{ADH1})-(P _{PGK1} -PsXYL1-T _{PGK1})	***
pRH 274	pRS416+P _{PGK1} -PsXYL1-T _{PGK1} ;P _{ADH1} -PsXYL2-T _{ADH1} ;P _{HXT7} -XKS1-T _{HXT7}	(HECTOR et al., 2011)
p416	pRS416 contendo promotor MET 25 e terminador CYC1	****

Cepas gentilmente cedidas pelo: *Dr. Gonçalo Amarante Guimarães Pereira (Unicamp); **Laboratório de Biologia Molecular do Prof. Gustavo H. Goldman da FCFRP da USP (Ribeirão Preto); Plasmídeo gentilmente cedido pelo: *** Pesquisador Ronald E. Hector (United States Departament of Agriculture, EUA). ****Pesquisador Britta Herzog (Institut Microbiologie and Genetik, Gottigen).

4.8.1. Transformação por acetato de lítio

A cepa PE-2 foi transformada com todos os plasmídeos da tabela 14, enquanto que a PE-2 XKS1 foi transformada apenas com os plasmídeos pSP-G1-Rd, pSP-G1-Pb, pRH274 e pSP-G1. Todas as transformações foram realizadas através do método de acetato de lítio, proposto por Gietz et al. (1992), com algumas modificações. As leveduras foram inoculadas em 5 mL de meio líquido YPD (Tabela 2) e incubadas sob agitação a 30°C durante aproximadamente 16 horas. Em seguida, foi medida a OD em espectrofotômetro com comprimento de onda de 600 nm e calculado o volume necessário para que um novo inóculo começasse a uma OD=0,1. O pré-inóculo foi transferido então, para 200 mL de meio líquido YPD e mantido a 30°C por cerca de 4 horas sob agitação de 200 rpm para crescimento exponencial, até atingir OD=0,4. Após atingir a OD desejada, a

cultura foi centrifugada a 5000g por 5 minutos e o sobrenadante foi descartado. As células foram ressuspensas em 20 mL de água estéril e centrifugadas a 5000 g por 5 minutos. O sobrenadante foi descartado e as células foram ressuspensas em 1 mL de solução 1xTE/LiAc (100 µL TE 10x; 100 µL LiAc 10x e 800 µL água). Em um tubo de 1,5 mL foi adicionado 1 µg do DNA de interesse e 100 µg de DNA de esperma de salmão (Promega) previamente desnaturado (10 minutos a 100°C). Posteriormente, a este tubo foi adicionado 100 µL das células e 600 µL de 40% PEG/1x LiAc (800 µL PEG50%; 100 µL LiAc 10x e 100 µL água). Após homogeneização, as células foram incubadas a 30°C por 30 minutos sob agitação de 200 rpm. Em seguida, foram adicionados 70 µL de DMSO (Merck) e o volume foi misturado por inversão. As células foram aquecidas a 42°C por 15 minutos em banho-maria e transferidas imediatamente para o no gelo por 2 minutos. Foram adicionados 700 µL de água, o volume foi homogeneizado por inversão e centrifugado por 30 segundos a 13000 rpm. O sobrenadante foi descartado e as células foram ressuspensas em 1 mL água estéril. Novamente foram centrifugadas por 30 segundos a 13000 rpm, o sobrenadante foi descartado e as células foram ressuspensas com 60 µL de água. Por fim, as células foram plaqueadas em meio sólido YNB ura⁻ (Tabela 15) e incubadas a 30°C por 2 a 4 dias para crescimento.

Tabela 15 – Meio YNB ura⁻.

Reagentes	Quantidade
YNB	7 g
Glicose	20 g
Ágar	20 g
Água destilada q.s.p	Para volume final de 1000 µL

As cepas PE-2 e PE-2 XKS1 transformadas foram denominadas de acordo com a tabela 16.

Tabela 16 – Nomenclatura adotada pelas cepas PE-2 recombinantes.

Nome da cepa recombinante	Plasmídio inserido
PE Rd	pSP-G1 Rd (genes da <i>R. dairenensis</i>)
PE Pb	pSP-G1 Pb (genes da <i>P. brasiliensis</i>)
PE pSP-G1	pSP-G1 vazio
PE 209	pRH209
PE 274	pRH274
PE 416	p416
PE-XKS1 Rd	pSP-G1 Rd
PE-XKS1 Pb	pSP-G1 Pb
PE-XKS1 pSP-G1	pSP-G1
PE-XKS1 274	pRH274

Para confirmar se as transformações tinha sido efetiva, foi extraído o DNA total de algumas das colônias crescidas em meio YNB ura^- e esse foi transformado em bactéria por eletroporação (vide 3.3.4). Se as colônias selecionadas possuísem os respectivos vetores, as bactérias seriam transformadas com sucesso (confirmação por miniprep e sequenciamento). Também foi feito PCR do DNA total da levedura, utilizando os oligos de clonagem (Tabela 13). Como controle negativo foram utilizadas colônias que tinham o vetor pSP-G1 vazio.

4.9. Caracterização das cepas de *S. cerevisiae* PE-2 recombinantes

4.9.1. Perfil de crescimento das cepas de *S. cerevisiae* PE-2 recombinantes

Para verificar o crescimento das cepas de *S. cerevisiae* PE-2 na presença de xilose e glicose, foi realizado ensaio de curva de crescimento com as cepas PE Rd, PE Pb, PE pSP-G1, PE 274, PE 209 e PE 416. O ensaio foi realizado em meio de cultura mínimo (0,67% YNB) com diferentes fontes e concentrações de carbono (1% glicose e 0,5%, 1% e 5% xilose) tamponados com 50 mM de hidrogenofterato de potássio (Sigma-Aldrich), pH 5,5. O experimento foi realizado em *erlenmeyers*

de 125 mL com 12,5 mL dos respectivos meio de cultura, onde foram inoculadas 10^6 leveduras/mL, em triplicata. As leveduras cresceram a 30°C, a 200 rpm. A cada 24 horas foi coletado 1 mL de cada cultura para determinação espectrofotométrica da densidade óptica (OD) à 595nm (Tecan Infinite M200), até 144h pós inóculo.

4.9.2. Atividade enzimática de XR e XDH das cepas de *S. cerevisiae* recombinantes

Também foram determinadas as atividades das enzimas XR e XDH (descrita em 4.5) das cepas PE Rd, PE Pb, PE 274 e PE 209, crescida em meio 0,67% de YNB enriquecido com 2% de glicose após 6h de cultivo. Como controles, foram utilizadas as cepa PE 416 e PE pSP-G1 (controles negativos), e a *S. stipitis* (controle positivo).

4.10. Cepas de *S. cerevisiae* PE-2 XKS1 recombinantes

4.10.1. Perfil de crescimento das cepas de *S. cerevisiae* PE-2 XKS1 recombinantes

Com o objetivo de verificar se houve alguma diferença e um possível aumento na assimilação de xilose devido à superexpressão de *xks1*, foi realizado ensaio de curva de crescimento, em triplicata, com as cepas de *S. cerevisiae* PE-2 XKS1 recombinantes na presença de xilose e glicose. As cepas analisadas foram a PE-XKS1 Rd e a PE-XKS1 Pb. Como controle positivo foi utilizada a cepa PE-XKS1 274 e negativo, a PE-XKS1 pSP-G1. O ensaio foi realizado da mesma maneira que no item 3.9.1.

5. Resultados e Discussão

5.1. A Identificação dos isolados

5.1.1. A Identificação das leveduras através da região ITS2

Inicialmente, foi feita uma busca por leveduras no trato intestinal de larvas de insetos parasitas da cana-de-açúcar por acreditar que, se esses insetos perfuram e se alimentam do colmo da cana e eles provavelmente têm uma microbiota capaz de metabolizar eficientemente a hemicelulose. Suh et al. (2005) afirmam que os fungos presentes no intestino são essenciais para a nutrição de vários insetos e seus resultados revelam que o intestino de insetos abrigam uma espantosa diversidade de leveduras não descritas. Segundo Suh et al. (2003), há uma contribuição das enzimas dos micro-organismos para a sobrevivência de vários insetos, como é o caso de cupins que utilizam fungos simbioses intestinais para quebrar seus alimentos celulósicos. Dessa forma, foram isolados 25 micro-organismos, sendo duas leveduras e o restante, fungos filamentosos. Foi feita a identificação dos isolados de leveduras ao nível de gênero ou espécie, através da amplificação e o sequenciamento da região espaçadora interna transcrita 2 (ITS2) do DNA ribossomal, uma região altamente conservada e com grande probabilidade de identificação bem sucedida para uma ampla gama de fungos (SCHOCH et al. 2012). Foi verificado que o isolado proveniente do inseto *Migdolus fryanus* tratava-se da espécie *Rhodotorula dairenensis* e que o isolado proveniente do inseto Crisomelídeo tratava-se de uma espécie pertencente ao gênero *Pseudozyma*. Entretanto, para uma identificação ao nível de espécie mais confiável e curada, optou-se por uma análise completa das regiões ribossomais ITS1, 5.8S, ITS2 e D1/D2 (28S).

5.1.2. *Rhodotorula dairenensis* TAB01

Para confirmar a identificação da espécie *Rhodotorula dairenensis* foi realizado um estudo da posição filogenética dessa levedura em relação a outras espécies tipo, utilizando os *amplicons* das regiões ITS1, 5.8S, ITS2 e D1/D2 do 28S. Essas regiões vêm sendo muito utilizadas na diferenciação de fungos através de análises de PCR (MARTIN; RYGIEWICZ, 2005). Sonnenberg et al. (2007) avaliaram a utilidade da subunidade ribossomal maior (LSU) como um marcador potencial na identificação de espécies. Os genes ribossomais do LSU são considerados altamente conservados, porém na realidade são constituídos por uma mistura de regiões conservadas e divergentes. Essas regiões divergentes são chamadas de regiões “D”. Assim, foi evidenciado que as divergências da região D1/D2 permitem a sua utilização como marcadores de DNA em uma ampla variedade de metazoários (SONNENBERG et al., 2007). Já a região ITS do DNA nuclear é o marcador de DNA preferido tanto para a identificação de um táxon único quanto para modelos ambientais mistos (BELLEMAIN et al., 2010).

Assim, as sequências obtidas de cada região foram concatenadas e após o alinhamento, foi inferida uma árvore filogenética (Figura 7). É possível observar que o isolado inicialmente identificado como *R. dairenensis* se encontra no mesmo ramo que a espécie tipo *R. dairenensis* CBS 4406. Porém, a identificação de uma espécie não pode ser confirmada somente pela árvore filogenética e sim, pelo número de diferenças em relação à espécie tipo (FELL et al., 2000). Então, comparamos as sequências dessas regiões da espécie tipo *R. dairenensis* CBS 4406 com as sequências do nosso isolado através do *software Water*. Para as regiões ITS1, 5.8S e ITS2 foi observado 100% de similaridade e identidade entre as sequências, sem a presença de *gaps*. Já para a região 28S (D1/D2), houve 99,8% de identidade e similaridade entre as sequências, com apenas 1 *gap*. Portanto, é possível afirmar que o isolado antes identificado como *R. darienensis* trata-se realmente dessa espécie, a qual a denominamos *R. dairenensis* TAB01.

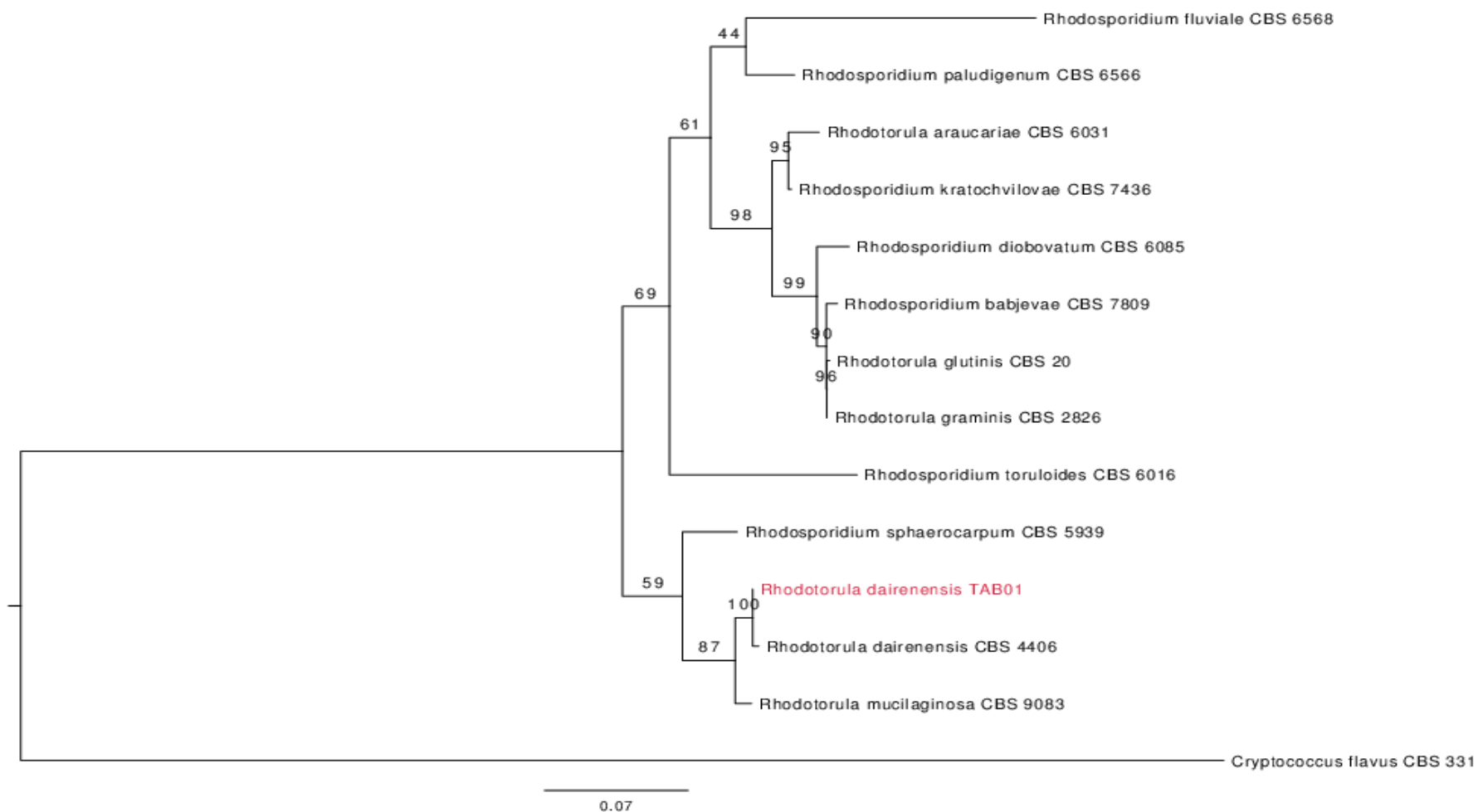


Figura 7 – Árvore filogenética da *R. dairenensis* TAB01 obtida por análises das regiões ITS e D1/D2 do 28S rDNA. Pela árvore, é possível observar sua posição filogenética em relação às demais espécies, a qual se localiza no mesmo ramo da espécie *R. dairenensis* CBS 4406 e filogeneticamente próximo à espécie *R. mucilaginosa*.

Esse resultado corrobora o trabalho de Gadanho e Sampaio (2002), confirmando que a *R. dairenensis* é filogeneticamente próxima à *R. mucilaginosa*, uma levedura basidiomiceto estudada principalmente por conta de seu metabolismo de xilose e produção de etanol (XU et al., 2011; BURA et al., 2012). Ainda, segundo Nagahama, Hamamoto e Horikoshi (2006), não é fácil diferenciar a *R. dairenensis* da *R. pacifica* em termos de diferenças fenotípicas. Pode-se apontar como diferença que a ramnose e galactitol são pouco assimiladas como fonte de carbono pela *R. pacifica*, contudo não são assimilados pela *R. dairenensis*.

5.1.3. *A Pseudozyma brasiliensis* sp.nov. GHG001

A similaridade de sequências da região ITS2 de um segundo isolado, utilizando o MEGABLAST, revelou que o mesmo pertencia ao gênero *Pseudozyma*. As regiões ITS, 5.8S e D1/D2 foram amplificadas por PCR e a partir delas, uma árvore filogenética foi gerada utilizando espécies tipos deste gênero (Figura 8). Pela árvore é possível visualizar que a *Pseudozyma* isolada neste trabalho formou um ramo com a *P. vetiver* DMKU-LV99 e a *P. fusiformata* JCM 3931, sendo mais relacionada filogeneticamente com a primeira, a qual foi recentemente isolada na Tailândia associada ao capim vetiver e a folhas de cana de açúcar (CHAMNANPA et al., 2013). A distância entre os ramos destas cepas sugeriu que ela poderia se tratar de uma nova espécie. Em uma análise das diferenças entre as sequências, foi observado que as principais diferenças encontram-se na região ITS. A *Pseudozyma* isolada apresentou 19 substituições comparadas à *P. vetiver* e 40 substituições comparadas à *P. fusiformata*. Na região D1/D2, a *Pseudozyma* isolada apresentou 3 substituições em relação à *P. vetiver* e 14 em relação à *P. fusiformata*. Fell et al (2000) sugerem que cepas diferindo por 2 ou mais nucleotídeos na região D1/D2 representam diferentes táxons. Assim, o número de substituições encontrado mostra que este isolado é uma nova espécie no gênero, que foi denominada *P. brasiliensis* sp.nov. GHG001.

A *P. brasiliensis* sp.nov. GHG001 consegue crescer em meio contendo xilano como única fonte de carbono, produzindo elevados níveis de uma endoxilanase, a qual apresenta alta atividade específica em comparação com outras xilanases de qualquer eucarioto já estudado, o que demonstra o elevado potencial biotecnológico dessa levedura. A descrição morfológica e testes fisiológicos foram realizados e descritos em um artigo do nosso grupo o qual foi submetido e está em fase final de avaliação (Oliveira et. al., manuscrito submetido para publicação). Dentre eles, vale ressaltar que a *P. brasiliensis* sp.nov. GHG001 cresce em xilose o que é relevante para este estudo, e diferentemente da *P. vetiver* ela cresce em inulina, galactose, arabinose, ribose, etanol, glicerol, eritritol, ribitol e ácido glucônico (Oliveira et. al, manuscrito submetido para publicação).

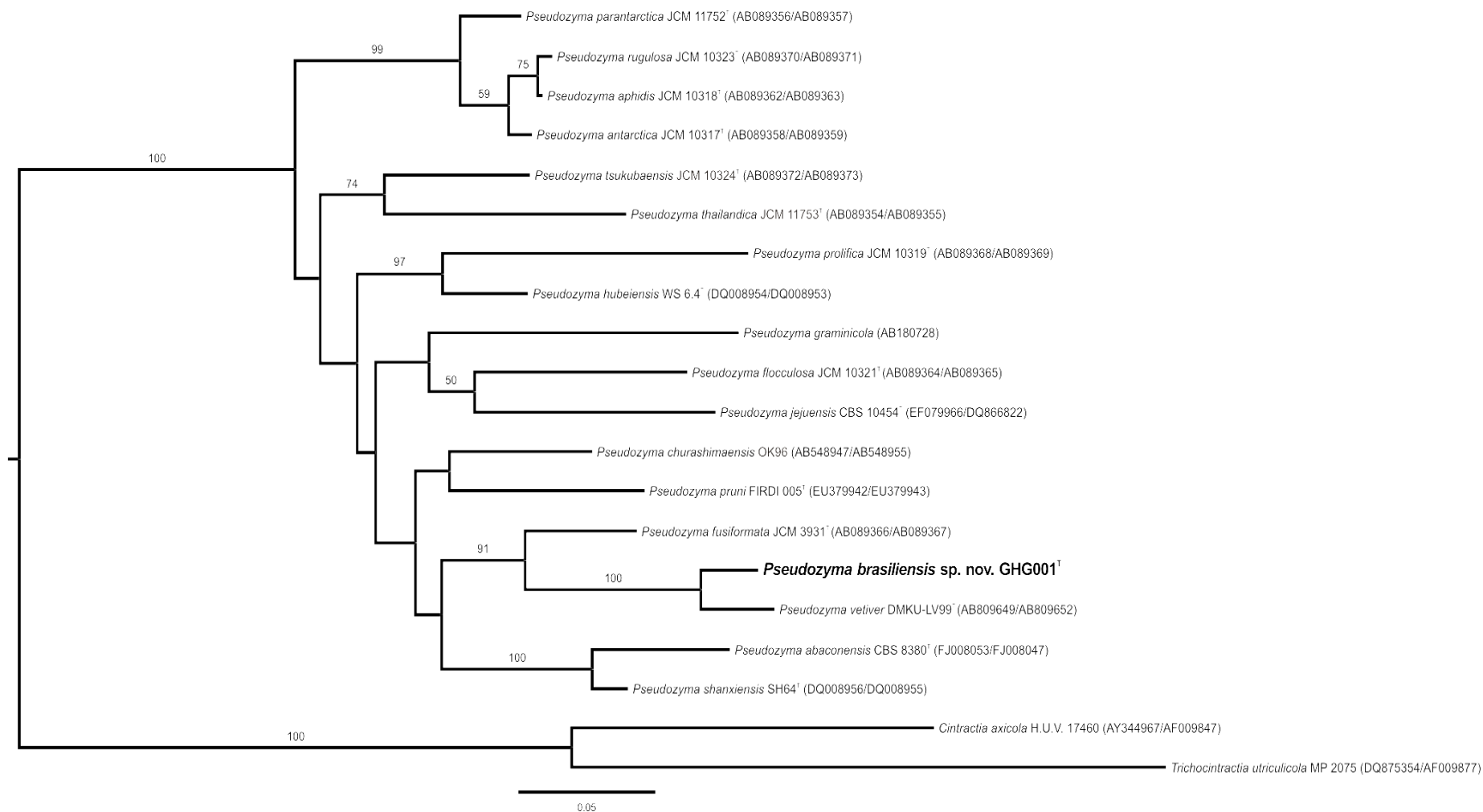


Figura 8 – Árvore filogenética da *P. brasiliensis* sp.nov. GHG001. É possível visualizar que a GHG001 está mais relacionada filogeneticamente com a *P. vetiver* DMKU-LV99.

5.2. Perfil de crescimento das leveduras

Para avaliar o crescimento destas leveduras em xilose, foi utilizada como controle positivo a levedura *S. stipitis* NRRL-Y7124, já que ela é um microorganismo capaz de crescer e fermentar eficientemente pentoses, especialmente a xilose (JEFFRIES, 2006). Já a *S. cerevisiae* Sc9721 foi utilizada como controle negativo, uma vez que esta cepa não utiliza a xilose para o seu crescimento (JEPPSSON et al., 2003; CAI et al., 2012).

Na curva de crescimento contendo 1% de glicose (Figura 9), observa-se o crescimento de todas as leveduras, sendo que a *R. dairenensis* TAB01, a *P. brasiliensis* GHG001 e a *S. stipitis* apresentaram crescimento até mesmo superior ao da *S. cerevisiae*, evidenciando a ótima assimilação da glicose como fonte de carbono por essas leveduras. Em relação ao crescimento em meio com xilose, ele foi realizado em diferentes concentrações deste açúcar (0,5, 1 e 5%), pois se almejava saber se havia alguma influência da concentração no crescimento das mesmas. Como previsto, o crescimento da *S. cerevisiae* é insignificante, visto que ela não é uma boa metabolizadora de xilose. Comparando-se as demais leveduras nas três diferentes concentrações de xilose, observa-se um maior crescimento no meio com 1% e 5% de xilose (Figura 9). Nota-se que a *R. dairenensis* TAB01 cresce mais que a *P. brasiliensis* GHG001, mas que ambas apresentam elevado crescimento nesses meios, ultrapassando em alguns pontos o crescimento da *S. stipitis*. Isso mostra que essas leveduras contêm genes eficientes para a metabolização de xilose, sendo promissores nesse estudo.

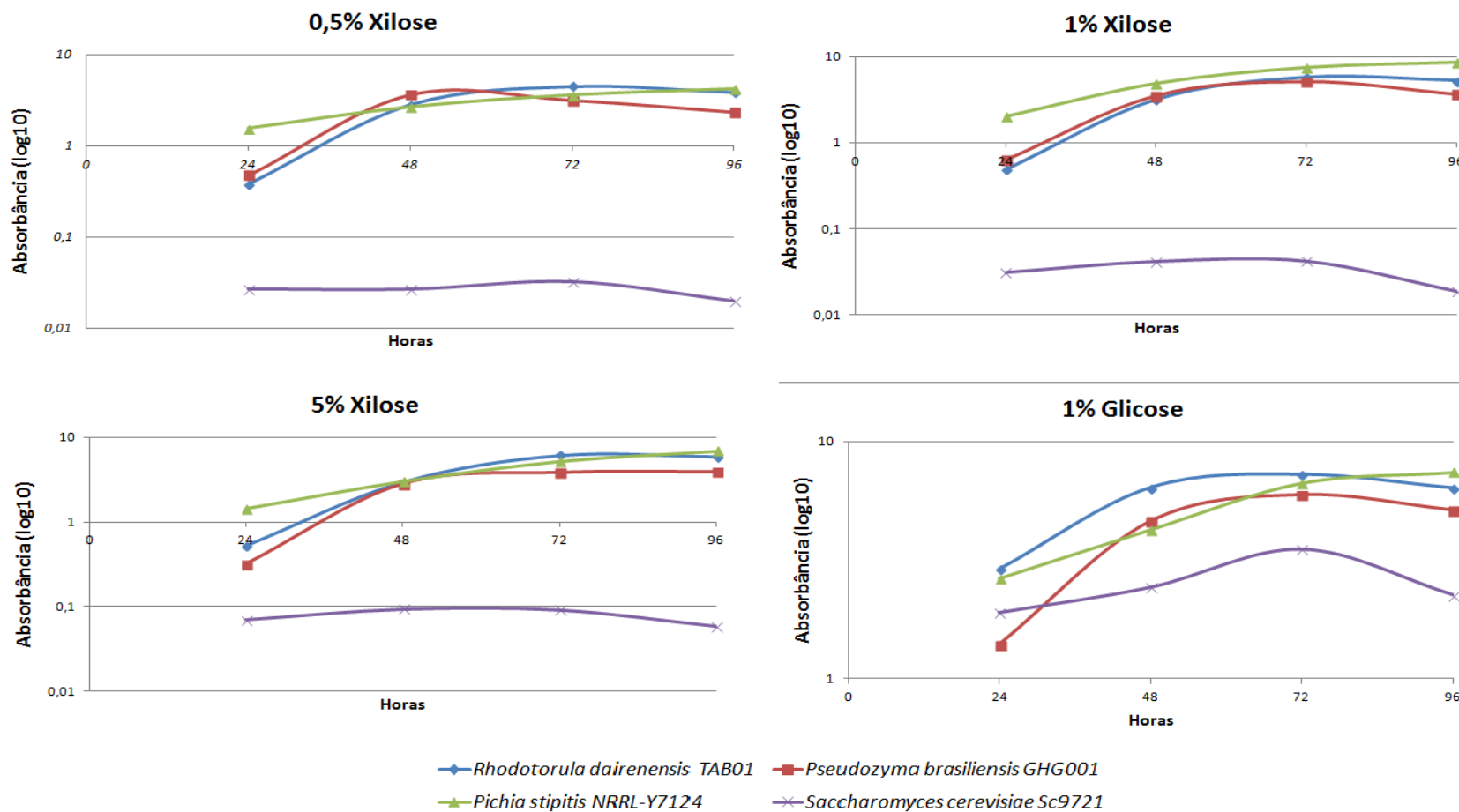


Figura 9 – Curvas de crescimento da *R. dairenensis* TAB01, *P. brasiliensis* GHG001, *S. cerevisiae* Sc9721 e *S. stipitis* NRRL-Y7124 em diferentes fontes e concentrações de carbono, mostrando a capacidade que os isolados deste trabalho tem de crescer utilizando xilose como fonte de carbono.

5.3. A caracterização molecular dos genes *xyl1* e *xyl2* de *R. dairenensis* TAB01

As enzimas XR e XDH são codificadas respectivamente pelos genes *xyl1* e *xyl2*. Como já mencionado, XR e XDH são as principais enzimas envolvidas na via de assimilação de pentoses através da via de redução-oxidação que ocorre em leveduras. Como o objetivo deste trabalho era o de introduzi-los nas cepas de *S. cerevisiae*, foi necessário a obtenção das regiões codificantes desses genes. Como as sequências de *xyl1* e *xyl2* de *R. dairenensis* TAB01 eram desconhecidas foi necessário desenhar oligonucleotídeos degenerados para uma pequena região, baseados em um alinhamento entre a sequência destes genes de leveduras filogeneticamente próximas à *R. dairenensis*. Após a confirmação por sequenciamento do produto obtido pela PCR com esses oligos degenerados, foi necessário utilizar a técnica de RACE para se obter a sequência da região codante (CDS) de *xyl1* e *xyl2* por completo. Os *contigs* de cada gene foram montados e mostraram que a região codificante do gene *xyl1* possui 1011 nucleotídeos (337 aminoácidos) e do *xyl2*, 1185 nucleotídeos (395 aminoácidos) (Figuras 10 e 11).

atggggaccacagtctcgagacctaataacacagaaacatgtcgagcaaatcccctccgtc
 M G T T V S R P K H R N M S Q Q I P S V
 aagctcggaacggcgccgagttccctctcctcggttcggcacctggcagtcggtccg
 K L A N G A E F P L L G F G T W Q S A P
 ggcgaggtcggcaaagccgctcgaggttgctctcaaggccggctaccgccacctcgacctt
 G E V G K A V E V A L K A G Y R H L D L
 gccaaaggtctacggcaaccagaaggagatcgctccggcgattgccaaactcgggagtcgac
 A K V Y G N Q K E I A P A I A N S G V D
 cgcaaggacatcttcatcacctcgaagctctggaacactcagcacaagccggacctcggtt
 R K D I F I T S K L W N T Q H K P D L V
 gaggtcgcgcttgacgacaccctcaaggagctcggcctcgagtacctcgacctctaccts
 E A A L D D T L K E L G L E Y L D L Y X
 atccactggccggttgctttcccgccgagggcgacccccactcgaacctcttcccgaag
 I H W P V A F P A E G D P H S N L F P K
 gagaacgacgtctgcaagatcgacacttcggtctcgatcgctcgacacctggaaggcgatg
 E N D V C K I D T S V S I V D T W K A M
 atcaagctcctcgacactggcaagaccaaggccgctcgggtgtctccaacttctcgccggca
 I K L L D T G K T K A V G V S N F S P A
 atggctcgacgccatcacccgagggcgaccggcgtaagccgcttgtaaccagatcgagcgc
 M V D A I T E A T G V K P L V N Q I E R
 caccgcgcctcctccagaaggacctcctcaagcaccacaaggagaagaacattgtcgtc
 H P R L L Q K D L L K H H K E K N I V V
 actgcctactctggcttcggaacaacagcgctcggcgagccgctcctcctcgagcacccg
 T A Y S G F G N N S V G E P L L L E H P
 accgtcaagaagatcgccgagggccaagggcgccaacccggggccaggtcctcatcgctgg
 T V K K I A E A K G A N P G Q V L I A W
 ggcattgcacggtggccacgccatcatccccaagtcggttaccgcgtcccgtatcgagtcg
 G M H G G H A I I P K S V T P S R I E S
 aacttcaaggctcgtctcccttaccgacgaggaggttcaggagatcaacaagatcgggcgag
 N F K V V S L T D E E V Q E I N K I G E
 gagaggccccgcacgttttaacctgccgatcgactacacgccagcctggccgattaacgctc
 E R P A R F N L P I D Y T P A W P I N V
 tttagacaccccggaagagaaggacgccaagtaccaggtcaagatccagtaa
 F D T P E E K D A K Y Q V K I Q -

Figura 10 – Sequência de nucleotídeos e aminoácidos da CDS da *xyl1* de *R. dairenensis* TAB01.

atggccagcagcaacaccccacttgcgccctccactcgccgaagacacgacggccacctcg
 M A S S N T P L A P P L A E D T T A T S
 cgcaacgtctccttcgtcctgcatggcatcgacgatgtccgcttcgaagagagagcagtc
 R N V S F V L H G I D D V R F E E R A V
 ccagtcgaatgcatgacgatgcagcaatcgtggccccaaggcgaccggaatctgcggc
 P V E C D D D A A I V A P K A T G I C G
 agtgatgtccactacctcaagcacggccggatcgagactttatcgtcaaggagccgatg
 S D V H Y L K H G R I G D F I V K E P M
 gtcctcggacacgagagcgcggcagtggtcgtcaaggtgggaagaatgtcaagaacatc
 V L G H E S A A V V V K V G K N V K N I
 aaggtcggagaccgtgtcgcgctcgaaccgggcaagtcgtgccgatcctgctacgactgc
 K V G D R V A L E P G K S C R S C Y D C
 aaggggtggtcactacgagcgtgccccgacatgatctttgcgggcaccccccttatgac
 K G G H Y E R C P D M I F A A T P P Y D
 ggtaccctcgtggccgctatgtcctcccgccgacctttgctacaaactgccggacaac
 G T L A G R Y V L P A D L C Y K L P D N
 atgtcgatggaggagggcgccctcctcgagcccatgtccgtggcggtgcacgcagttgcc
 M S M E E G A L L E P M S V G V H A V A
 aaggttgccgagctcaagcccggtcgaatgtggttggtttggggccgggcccgtcggt
 K V A E L K P G S N V V V F G A G P V G
 ctccttaccgcgcgcgcgcaaaggtctcggggccgcccgtgtcattgccgtcgacatt
 L L T A A A A K G L G A A R V I A V D I
 caagagagtcgtctgcaattcgcaaaggagaacggcttgatccacgacttctacgtcccg
 Q E S R L Q F A K E N G L I H D F Y V P
 tcaaagccgcaagaaggggaagacaaggtcgacttccagcgacggaatgccaaaggaaatt
 S K P Q E G E D K V D F Q R R N A K E I
 caggagcgcctttggcttcacggagcggggggcgactggcgctcgactatgtctttgagtgt
 Q E R F G F T E R G A T G V D Y V F E C
 tccggcgctgaagtctgcatcggcacctcggtattcctgctcaagcacggcggtaccttg
 S G A E V C I G T S V F L L K H G G T L
 gtccagatcggcatgggtcgtccggacattaccctcgacatgcacaccgtgctcacgcac
 V Q I G M G R P D I T L D M H T V L T H
 gaactgaccctcaagggcagcttccggtacggacggacgtgtaccgactgtcgtcgac
 E L T L K G S F R Y G P D V Y R L S L D
 ctcgtcgcgcgcgggtgccgtcaacctcaagtcctcatcacgcaccgctacgcattcaag
 L V A R G A V N L K S L I T H R Y A F K
 gaggcgccggaagcgttcgaagccaacaccaagggcgctcggaaggacggcaacgctgtc
 E A P E A F E A N T K G V G K D G N A V
 atcaagatcatcattgccggccccgaccgagtcggacactgcgtaa
 I K I I I A G P T E S D T A -

Figura 11 – Sequência de nucleotídeos e aminoácidos da CDS da *xyl2* de *R. dairenensis* TAB01.

A fim de comparar estas sequências contra um banco de dados, foi realizado um *blastp* dessas sequências. A proteína XR apresentou como melhor *hit* “xylose reductase [*Rhodotorula mucilaginosa*]” com 96% de cobertura, *e-value* 0,0 e 95% de identidade (Tabela 17). Já a XDH, apresentou como melhor *hit* “xylitol dehydrogenase [*R. mucilaginosa*]” com 100% de cobertura, *e-value* 0,0 e 95% de identidade (Tabela 18). Esses resultados indicam que as proteínas codificadas pelos genes *xyl1* e *xyl2* codificam realmente as proteínas de interesse do presente estudo.

Tabela 17 – Resultado do *blastp* da sequência correspondente ao gene *xyl1*.

Sequências que produzem alinhamentos significantes	Cobertura	e-value	Identidade
xylose reductase [<i>Rhodotorula mucilaginosa</i>]	96%	0.0	95%
xylose reductase [<i>Rhodospiridium toruloides</i> NP11]	96%	0.0	77%
xylose reductase [<i>Rhodotorula graminis</i>]	92%	1e-165	73%

Tabela 18 – Resultado do *blastp* da sequência correspondente ao gene *xyl2*.

Sequências que produzem alinhamentos significantes	Cobertura	e-value	Identidade
xylitol dehydrogenase [<i>Rhodotorula mucilaginosa</i>]	100%	0.0	95%
xylitol dehydrogenase [<i>Rhodospiridium toruloides</i> NP11]	96%	0.0	79%
xylitol dehydrogenase [<i>Rhodotorula graminis</i>]	86%	1e-173	71%

5.4. A expressão dos genes *xyl1* e *xyl2* da *R. dairenensis* TAB01

A fim de avaliar se os genes *xyl1* e *xyl2* estavam sendo expressos, foi realizado o PCR quantitativo desses genes. Comparando as expressões de *xyl1* e *xyl2* em meio com glicose e xilose observa-se que elas são mais elevadas nos meios contendo xilose (principalmente em meio 5%), o que evidencia que esses genes são induzidos e regulados por este açúcar. Pode-se observar uma baixa

expressão destes genes em glicose, o que sugere que os genes da via metabólica da xilose são reprimidos catabolicamente por glicose (Figura 12). Xu e colaboradores (2011) concluiu que a xilose é um eficiente indutor desses genes e isto é uma tendência encontrada também em *Candida guilliermondii* e *S. stipitis*.

Comparando a expressão dos dois genes, pode-se observar que a expressão de *xyl2* é muito inferior à de *xyl1*, resultado contrário aos da *R. graminis* WP1 e *R. mucilaginosa* PTD3 no trabalho de Xu et al. (2011). Contudo, vale ressaltar que no trabalho citado, para avaliar a expressão dos genes, foi utilizada a técnica de RT-PCR e no presente estudo, nós utilizamos qRT-PCR, que é uma técnica bem mais sensível para medir a expressão gênica.

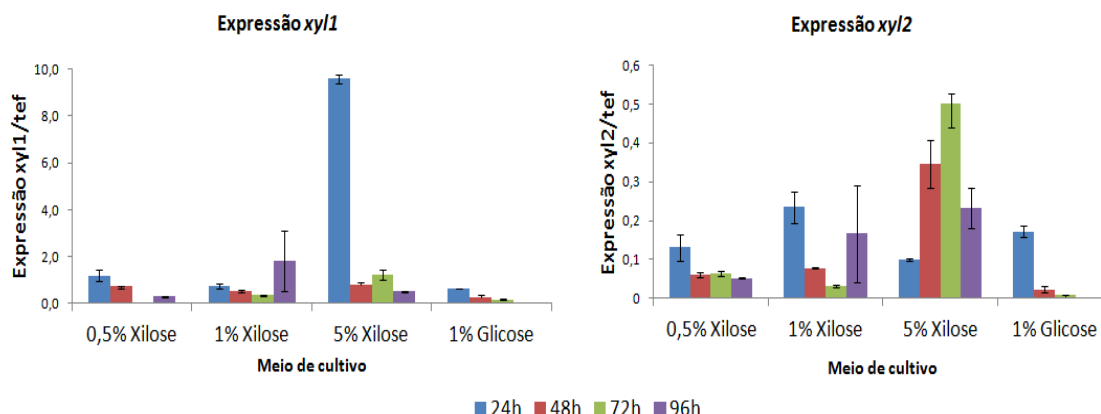


Figura 12 – Expressão de *xyl1* e *xyl2* da *R. dairenensis* TAB01 em diferentes meios ao longo de 96 horas, mostrando que esses genes são induzidos por xilose, principalmente nos meios contendo 5% do açúcar.

Para ambos os genes, na maioria das condições, observa-se uma queda na expressão no decorrer do tempo, inversamente proporcional ao crescimento (Figuras 9 e 12). Isso pode ocorrer porque nas primeiras horas a levedura ativa todos os genes da via metabólica para que ela assimile a fonte de carbono disponível e com isso, obtenha energia necessária para crescer. Com o passar do tempo, a fonte de energia vai se esgotando, a levedura entra em fase estacionária e de declínio e, conseqüentemente a expressão desses genes diminui.

5.5. A atividade enzimática da XR e XDH endógenas da *R. dairenensis* TAB01

Como os genes de assimilação de xilose estavam sendo expressos (Figura 12), foi verificado se as enzimas XR e XDH (codificadas por esses genes) apresentavam atividade, já que essas enzimas são as responsáveis pela conversão de xilose a xilulose. Nesse experimento, *S. stipitis* NRRL-Y7124 foi usada como controle positivo por ser uma boa assimiladora de xilose e por apresentar, segundo a literatura, altas atividades de XR e XDH. Além disso, a xilose redutase de algumas leveduras apresenta afinidade por dois cofatores, NADPH e NADH, sendo chamada de duplo-específica, como é o caso da *S. stipitis*. A XR dessa levedura utiliza o NADH como cofator quase tão bem quanto o NADPH, evitando o desbalanço de cofatores, diferentemente do que ocorre para a xilose redutase estritamente NADPH-dependente da *S. cerevisiae* (HECTOR et al., 2011). O desbalanço de cofatores é um dos principais gargalos do processo de metabolização de xilose visando à produção de etanol, pois em leveduras onde a xilose redutase é estritamente NADPH-dependente, a xilitol desidrogenase utiliza especificamente o NAD^+ como cofator, causando uma depleção de NADPH e um excesso de NADH. Em condições anaeróbicas, o excesso de NADH acumula e leva a uma escassez de NAD^+ , consequentemente a utilização de xilose fica lenta ou em alguns casos cessa (HECTOR et al., 2011). Além disso, este é o principal fator que leva ao acúmulo do subproduto xilitol, indesejável na produção de etanol (MATSUSHIKA et al., 2009; KIM et al., 2013)(KIM et al., 2013)(KIM et al., 2013)(KIM et al., 2013)(KIM et al., 2013)(KIM et al., 2013)(KIM et al., 2013)(KIM et al., 2013)(KIM et al., 2013). Por estes motivos, nas leveduras deste trabalho foi testada a atividade enzimática de XR na presença dos dois cofatores, NADPH e NADH.

Nota-se que ao contrário da *S. stipitis* NRRL-Y7124, a xilose redutase da *R. dairenensis* TAB01 praticamente não tem atividade na presença do cofator NADH (Figura 13). Visto que a atividade de XR da *S. stipitis* NRRL-Y7124 é tida como

superior à atividade de XDH nos trabalhos de Ho et al. (1998) e Wahlbom et al. (2003), os valores obtidos para o controle positivo no nosso estudo estão de acordo com os reportados na literatura. Além disso, a atividade de XR-NADPH dependente da *R. dairenensis* TAB01 (0,19 mg/U) (Figura 13) foi considerada como um resultado satisfatório, pois apesar de serem inferiores a atividade da *S. stipitis* NRRL-Y7124, foram maiores que atividade da XR-NADPH dependente (0,034 U/mg) da levedura *Rhodotorula glutinis* NRRL-Y2502 (Freer et al., 1997). Quanto à atividade de XDH (Figura 13), a *R. dairenensis* TAB01 (0,033 U/mg) apresentou valor inferior à atividade de XDH da *S. stipitis* e também da *Rhodotorula glutinis* (0,108 U/mg). Isso pode indicar um possível acúmulo do sub-produto xilitol no meio, o que não é desejável no nosso trabalho que objetiva o etanol como sendo o produto final.

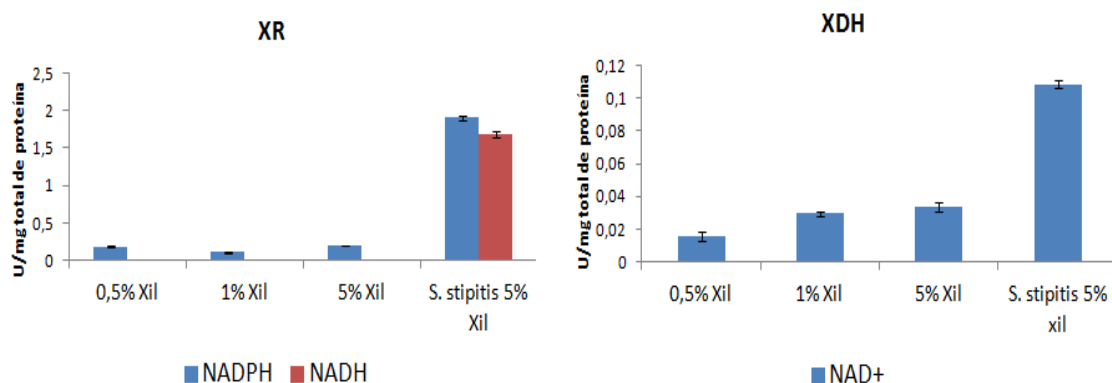


Figura 13 – Atividade enzimática de XR da *R. dairenensis* TAB01 na presença de NADPH e NADH e de XDH na presença do cofator NAD⁺. Nota-se a ausência de atividade de XR na presença do cofator NADH.

Comparando estes dados com os obtidos na expressão gênica, verifica-se que há uma correlação positiva entre eles, pois a mesma diferença vista na quantidade dos mRNAs produzidos é vista na atividade enzimática (XR é cerca de 10 vezes maior que XDH, em ambos os casos). Entretanto, para confirmar que esta correlação se estende realmente para a proteína, seria necessário realizar

algum outro experimento, como por exemplo, a medição da concentração das proteínas através de *immunoblotting*.

5.6. A identificação molecular dos genes *xyl1* e *xyl2* da *Pseudozyma brasiliensis* GHG001

Como feito para a *R. dairenensis* TAB01, foi realizado um alinhamento global desses genes de todas as leveduras depositadas no banco de dados do NCBI e desenhados oligonucleotídeos degenerados para regiões mais conservadas. Entretanto, não obtivemos nenhum resultado positivo com os oligonucleotídeos desenhados e nem como os descritos por Xu e colaboradores (2011), mesmo depois de testar diferentes condições de PCR (diferentes temperaturas de anelamento, diferentes concentrações de Mg, uso de *enhancers*, entre outros). Isso provavelmente se deve ao fato de que a maioria das sequências depositadas são ascomicetos, diferentemente de *P. brasiliensis* GHG001 que é uma levedura basidiomicética (WANG et al., 2006).

Assim, devido ao fato de se tratar de uma nova espécie e dado o potencial biotecnológico desta levedura, decidimos sequenciar seu genoma. Assim, o seu DNA total foi extraído e enviado para USC *Epigenome Center Data Production Facility*, na Califórnia-EUA. Lá foram construídas bibliotecas *pair-end* (100bp) para sequenciamento no *Hi-Seq* (Illumina). Os resultados relativos ao sequenciamento do genoma estão apresentados no Anexo I.

A partir dos dados gerados e analisados por bioinformatas do CTBE, foi identificada uma região do DNA que correspondia à *xyl2* da *P. brasiliensis* GHG001 (Figura 14). A partir dessa sequência, foi realizado o RACE para se obter as extremidades 5' e 3' do gene. Após montar o *contig* foi obtido a sequência da região codante (CDS) do possível candidato do gene da xilitol desidrogenase, que apresenta 1164 nucleotídeos (388 aminoácidos) (Figura 15). Ao realizar uma comparação com o banco de dados do NCBI utilizando o *blastx*, o *contig* da

sequência de aminoácidos apresentou 92% de identidade e e-value 0,0 com “probable xylitol dehydrogenase” de *Sporisorium reilianum* SRZ2.

```
ACCGGTCTCTGCGCCTCGGACTGCCACTACCTCCACCACGGTCGCATCGGCGACTTTGTG
GTGCGCAAGCCCATGGTGTGGGTACGAATCGAGCGGTATCGTCACCGCTGTCGGCGAG
GGTGTACCACGCACAAGGTCGGGGATCGTGTTGCGCTCGAGCCTGGTGTGCCGTGTCGC
TCCTGCCCCGTCTGCTTGAGTGGCAGCTACAACCACTGTGCCAACCTCGAGTTCGCCGCC
ACTCCGCCCTACGATGGCAGCTGTGCACGTACTACAACATCCACTCGACGTTTCGCGCAC
CACGTTCCAGACACCATGTGCTCGAAGAAGCGTCGTTGATGGAGCCGCTCTCGGTGGCG
GTCTACTCGGCGGCCATGCGCGGCCAGGTCAAGGCGATGGAGAACGTGCTTGTGTTGGC
GCCGGTCCCATCGGTCTGCTCAACGCGGCGGTCTGCAAGGCCTACTCTGCCAAGCGTGTC
GTCGTGGTCGACGTGGTCGACTCAAAGCTCGAATTTGCCAGTCCTTCTGCGCCACATCG
ACGTTCAAGCCCAGCCTGCCGCTCCAAGGCGAGGCCAAGATCGACTCCGCTATGCGCAAC
GCCCAGGACCTCATCAGTCAGCTCGGCGACGACGTCGTCGCGCGCAAGGCTTCGACCTC
GTGCTCGAATGCACCGGCGCCGAACCGTGCATCCAGATGGGCATCCAGGCTCTGCGCCCC
AAGGGCCGCTTCGTCCAGGTCGGCATGGGCAGGTCCGAAGTCGAATCCCCATCACCCGG
GTTTGCCTCAAGGAGATCGACGTCACCGGAAGCTTCAGGTACGGCGCGGGGACCTACAAG
AC
```

Figura 14 – Região do DNA predita através do genoma da *xyl2* da *P. brasiliensis* GHG001.

```

atgccagctccagtcgcgcccaacaactcgctctccacccacccggtcgagggcaacctc
M P A P V A P N N S L S T H P V E G N L
tcgttcgtcctgcaagagatcgagaatgtctcgttcgaggagcgaccggtcgtcgcgccc
S F V L Q E I E N V S F E E R P V V A P
aagcccgggcaggtgcaggtcaacatccgtcagaccgggtctctgcgccctcggactgccac
K P G Q V Q V N I R Q T G L C A S D C H
tacctccaccacggtcgcatcgggcgactttgtggtgcgcaagcccatggtgttgggtcac
Y L H H G R I G D F V V R K P M V L G H
gaatcgagcgggtatcgctcaccgctgtcggcgaggggtgttaccacgcacaaggtcggggat
E S S G I V T A V G E G V T T H K V G D
cgtgtttgcgtcgcagcctggtgtgccgtgtcgtcctcctgccccgtctgcttgagtggcage
R V A L E P G V P C R S C P V C L S G S
tacaaccactgtgccaacctcgagttcgccgcccactccgcccctacgatggcacgctgtgc
Y N H C A N L E F A A T P P Y D G T L C
acgtactacaacatccactcgacgttcgcgccaccacgttcagacacccatgctcgctcgaa
T Y Y N I H S T F A H H V P D T M S L E
gaagcgtcggttgatggagccgctctcggtggcggtctactcggcgcccatgcgcgccag
E A S L M E P L S V A V Y S A A M R G Q
gtcaaggcgatggagaacgtgcttgtgtttggcgccgggtcccatcggtctgctcaacgcg
V K A M E N V L V F G A G P I G L L N A
gcggtctgcaaggcctactctgccaagcgtgtcgtcgtgggtcgacgtgggtcgactcaaag
A V C K A Y S A K R V V V V D V V D S K
ctcgaatttgcccagtccttctgcgccacatcgacgttcaagcccagcctgccgctccaa
L E F A Q S F C A T S T F K P S L P L Q
ggcgaggccaagatcgactccgctatgcgcaacgccaggacctcatcagtcagctcggc
G E A K I D S A M R N A Q D L I S Q L G
gacgacgtcgtcgcgcgcgaaggcttcgacctcgtgctcgaatgcaccggcgccgaaccg
D D V V A R E G F D L V L E C T G A E P
tgcattccagatgggcatccaggtctctgcgcccccaaggcgcttcgtccaggtcggcatg
C I Q M G I Q A L R P K G R F V Q V G M
ggcaggtccgaagtccaattccccatcaccggggtttgcgtcaaggagatcgacgtcacc
G R S E V E F P I T R V C V K E I D V T
ggaagcttcaggtacggcgcggggacctaagaccagcatcaacctcgtcagcactgggt
G S F R Y G A G T Y K T S I N L V S T G
ctgatcgacgtcaccaagatggctcactcacggtttctcttcaaggacgcagstcaagg
L I D V T K M V T H R F L F K D A V K X
tttgagaccaccacaaagggtgttgggtgaggacggaaagaccgccmtcaagggtgcagatc
F E T T T K G V G E D G K T A X K V Q I
agccagggcgagggccagcagtag
S Q G E G Q Q -

```

Figura 15 – Sequência de nucleotídeos e aminoácidos da CDS da *xyI2* da *P. brasiliensis* GHG001.

Em relação ao gene *xyI1*, não foi possível encontrar um possível candidato ao gene utilizando a mesma abordagem que foi utilizada para o *xyI2*. Isso é justificado pelo fato das redutases em geral apresentarem um domínio em comum, o que dificultou identificar com segurança a xilose redutase de *P. brasiliensis*

GH001. Assim, foi feito uma busca de dez redutases anotadas de diversos micro-organismos (vide material e métodos) e foi inferida uma árvore filogenética de todas as sequências obtidas. Pela figura 16 foi possível observar, agrupadas em um mesmo clado junto com uma proteína de *P. brasiliensis* (em azul), xiloses redutases de outros micro-organismos as quais foram experimentalmente comprovadas, inclusive a da *R. dairenensis* TAB01 (em vermelho). Esta proteína de *P. brasiliensis* foi considerada como sua provável xilose redutase a qual apresenta 981 nucleotídeos e 327 aminoácidos (Figura 17). A confirmação da funcionalidade dos genes *xyl1* e *xyl2* só foi certificada após análises experimentais detalhadas nos próximos itens.

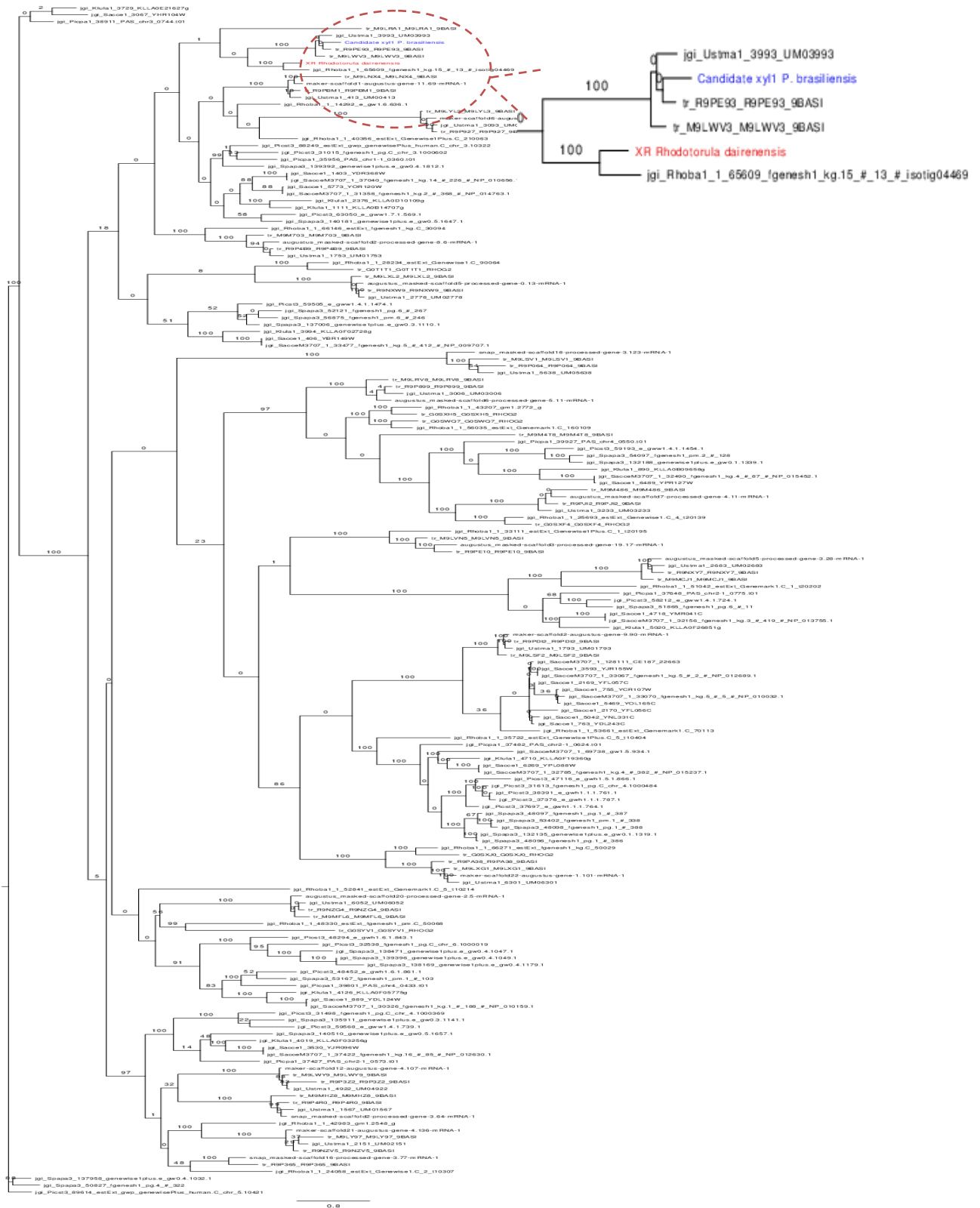


Figura 16 – Árvore filogenética das proteínas que possuem domínio de redutases.

No clado em destaque podemos visualizar o agrupamento das xiloses redutases

experimentalmente comprovadas de outros micro-organismos, e em azul o possível candidato ao gene *xyl1* da *P. brasiliensis* GHG001.

```

atgtcgctcaacaagactctcactctcaacgacggcaacaaaatccccagattggtctc
M S L N K T L T L N D G N K I P Q I G L
ggcacctggctctccaagcccgagaggtcgcaaacgccgtcgaagtcgccatcaaggct
G T W L S K P G E V A N A V E V A I K A
ggctatcgtcacattgactgtgcactcatctacaaaattcaaccgaggttggcgaagcg
G Y R H I D C A L I Y Q N S T E V G E A
tttaagaagacgatcccatcagtgggtcaagcgcggaagatctcttcgtcacgtcgaagctt
F K K T I P S V V K R E D L F V T S K L
tggaacaatgctcacaagcctgagcttgtcgaggcggcctataacgagacattgcaggag
W N N A H K P E L V E A A Y N E T L Q E
ctcggactcgaatacctcgatctgtatctgatccactggcctgtagccttcaaggacggc
L G L E Y L D L Y L I H W P V A F K D G
aaggacctcgtccccaagtcggaggacggcaagcaggccgcactcgaccgccagacctcg
K D L V P K S E D G K Q A A L D R Q T S
atcgctcgacacgtggaaggctctcgtcaagctgcagaaggagggcaagggtcaagtcgac
I V D T W K A L V K L Q K E G K V K S I
ggtgtctcgaatttcaccaccgatgacctggacgccgtcatcaacgcgaccgggtgtcact
G V S N F T T D D L D A V I N A T G V T
ccgcgcgtcaaccagatcgaggctcacccttctgctgcccaggacgaccttgtcgcttac
P A V N Q I E A H P L L P Q D D L V A Y
gctaagtcgaagaacatccacctgaccgcatactccccgctcggcaacaacctcaatggt
A K S K N I H L T A Y S P L G N N L N G
aagactaagattgtggattaccctcaggtcagcgaggtcgccaagaagtacaacgccgac
K T K I V D Y P Q V S E V A K K Y N A D
cctgctcaggttcttctgtcgcttgggggtgtcaagagggggtacagtgtcattcccaagagt
P A Q V L V A W G V K R G Y S V I P K S
gtcaccgagtcgaggatcaagagcaacttcgaccagctcgaactcaaggacgaggactac
V T E S R I K S N F D Q L E L K D E D Y
gagaaggtcaccagcctctacaaggagctcggcaagggtccgcttcaatgtcccaaaaaca
E K V T S L Y K E L G K V R F N V P K T
tatgttccgttctgtggacatcgatatctttggcgaggaggccgaacagggcgctacgtac
Y V P F W D I D I F G E E A E Q G A T Y
aagggtcaaccgtggcaactaa
K V N R G N -

```

Figura 17 – Sequência de nucleotídeos e aminoácidos da CDS do candidato à *xyl1* da *P. brasiliensis* GHG001.

5.7. Análise da expressão gênica de *xyl1* e *xyl2* de *P. brasiliensis* GHG001

A análise da expressão gênica de *xyl1* e *xyl2* de *P. brasiliensis* GHG001 foi feita por qRT-PCR. Não há na literatura, até o momento, trabalhos que avaliam a expressão desses genes em espécies do gênero *Pseudozyma* spp. Portanto, foi feita uma comparação da expressão desses genes com os da *R. dairenensis* TAB01.

A expressão de *xyl1* da *P. brasiliensis* GHG001 apresentou um comportamento diferente da *R. dairenensis* TAB01 no meio contendo glicose. Inesperadamente, esse gene foi expresso em glicose assim como nos meios contendo xilose, principalmente nas primeiras 24 horas de crescimento, o que indica que esses genes estão sendo induzidos também pela glicose (Figura 18). Além disso, de acordo com a figura 18 é possível observar que na condição 5% xilose, a expressão se mantém praticamente idêntica nos diferentes pontos de crescimento, sugerindo que essa quantidade de carbono é suficiente para induzir a expressão dos genes e mantê-los expressos durante as 96 horas de crescimento. Também pode-se visualizar que a expressão de *xyl2* apresenta uma maior expressão no meio contendo 5% xilose após 72 horas de cultivo, além de apresentar expressões similares no ponto 24 horas em todas as concentrações de xilose. Ao contrário de *xyl1*, a expressão de *xyl2* no meio contendo glicose, é baixa. Porém vale ressaltar que no geral a expressão de *xyl2* é menor que a de *xyl1*, com exceção das primeiras 24 horas de crescimento e do meio contendo 5% de xilose o qual apresentaram níveis de expressão relativamente altos e similares aos de *xyl1*. Contudo, em comparação com a *R. dairenensis* TAB01, os níveis de expressão de *xyl2* são bem maiores e satisfatórios.

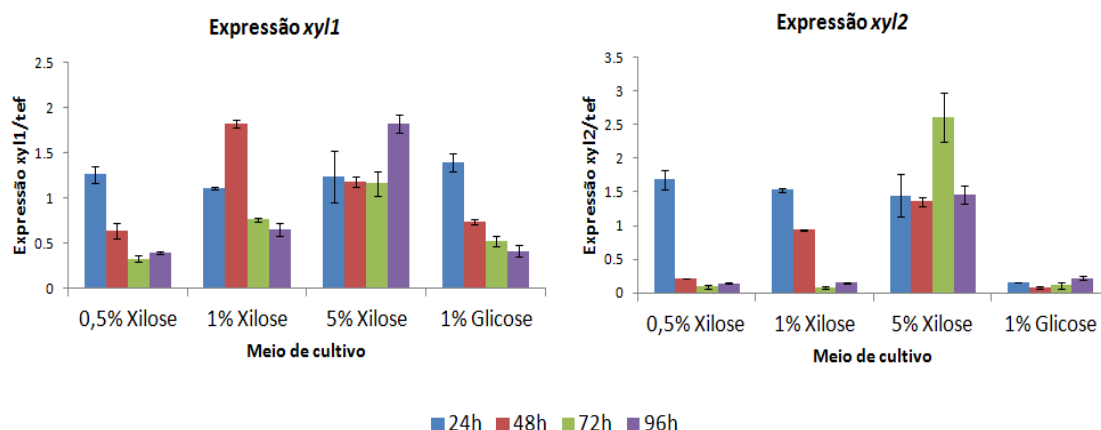
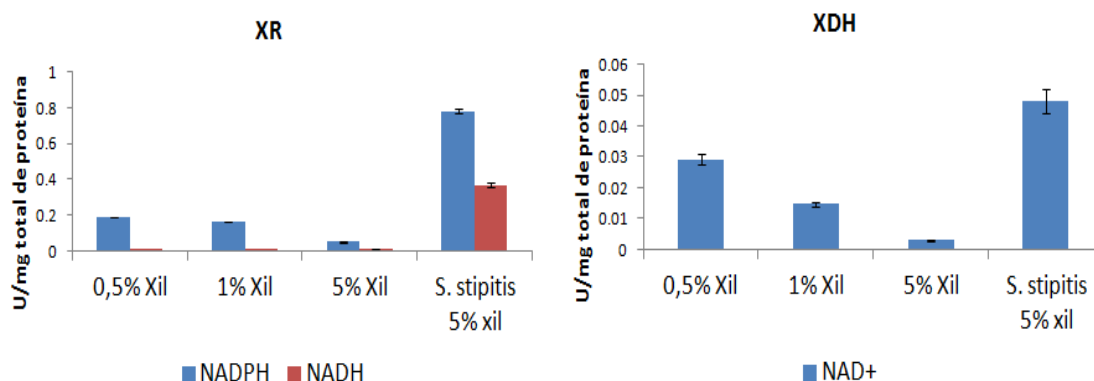


Figura 18 – Expressão de *xyl1* e *xyl2* da *P. brasiliensis* GHG001 em diferentes meios ao longo de 96 horas. Pelos gráficos é possível visualizar que os genes são induzidos pela xilose e, ao contrário do que ocorreu com a *R. dairenensis*, o gene *xyl1* é também induzido por glicose.

5.8. A atividade enzimática da XR e XDH endógenas da *P. brasiliensis* GHG001

Foi observado que a XR da *P. brasiliensis* GHG001 praticamente não teve atividade na presença do cofator NADH (Figura 19), assim como a xilose redutase da *R. dairenensis* TAB01 (Figura 13). Além disso, tanto a atividade de XR como a de XDH de *P. brasiliensis* GHG001 são bastante similares as da *R. dairenensis* TAB01 (Fig. 13) e inferiores às da *S. stipitis* NRRL-Y7124. É importante observar também que a atividade de ambas as enzimas são inversamente proporcionais à concentração do meio, e pode-se sugerir que quando a levedura cresce em pouca fonte de carbono, ela ativa ao máximo suas enzimas e vias de metabolização, para maximizar a assimilação dos açúcares.

Do mesmo modo que a *S. stipitis* CBS 6054 no trabalho de Wahlbom e colaboradores (2003), a XR da *P. brasiliensis* GHG001 apresentou maior atividade que XDH.



Fi

gura 19 – Atividade enzimática de XR e XDH da *P. brasiliensis* GHG001 na presença de NADPH e NADH e de XDH na presença do cofator NAD⁺. É possível visualizar que a XR da *P. brasiliensis* GHG001 não apresenta atividade na presença do cofator NADH.

Comparando esses resultados de atividade com os de expressão, podemos ver que diferentemente da *R. dairenensis* TAB01, não há uma correlação entre eles, visto que os dados de expressão foram semelhantes para ambos os genes, enquanto que a atividade de XR foi bastante superior a de XDH.

5.9. A clonagem dos genes *xyl1* e *xyl2* no vetor pSP-G1 e construção das cepas de *S. cerevisiae* recombinantes

A xilose redutase e a xilitol desidrogenase são as enzimas-chave na fermentação de xilose e, portanto elas têm sido amplamente usadas na construção de leveduras recombinantes fermentadoras de xilose (KHATTAB et al., 2011). Devido às leveduras estudadas neste trabalho serem boas assimiladoras de xilose, foi proposta a construção de cepas recombinantes de *S. cerevisiae* com esses genes, sendo um dos objetivos do projeto a expressão heteróloga das enzimas XR e XDH na cepa industrial *S. cerevisiae* PE-2. Assim, após verificar que as leveduras *R. dairenensis* TAB01 e *P. brasiliensis* GHG001 são capazes de utilizar xilose para crescimento, que seus genes *xyl1* e *xyl2* são expressos na

presença deste açúcar e que as respectivas enzimas XR e XDH possuem atividades enzimáticas, realizou-se a clonagem dos genes de interesse no vetor de expressão de levedura pSP-G1.

A confirmação da clonagem no vetor pSP-G1 foi realizada por PCR das colônias crescidas em meio seletivo utilizando os oligonucleotídeos referentes aos genes de interesse e após essa confirmação, o mesmo foi transformado nas cepas industriais da PE-2 e na PE-2 XKS1, utilizando como controle o vetor pSP-G1 vazio. Em seguida, as transformações das leveduras foram confirmadas extraindo o DNA total de alguns transformantes crescidos em meio YNB sem uridina, e transformando esse DNA extraído em bactéria, com posterior confirmação através de PCR de colônia (Figura 20).

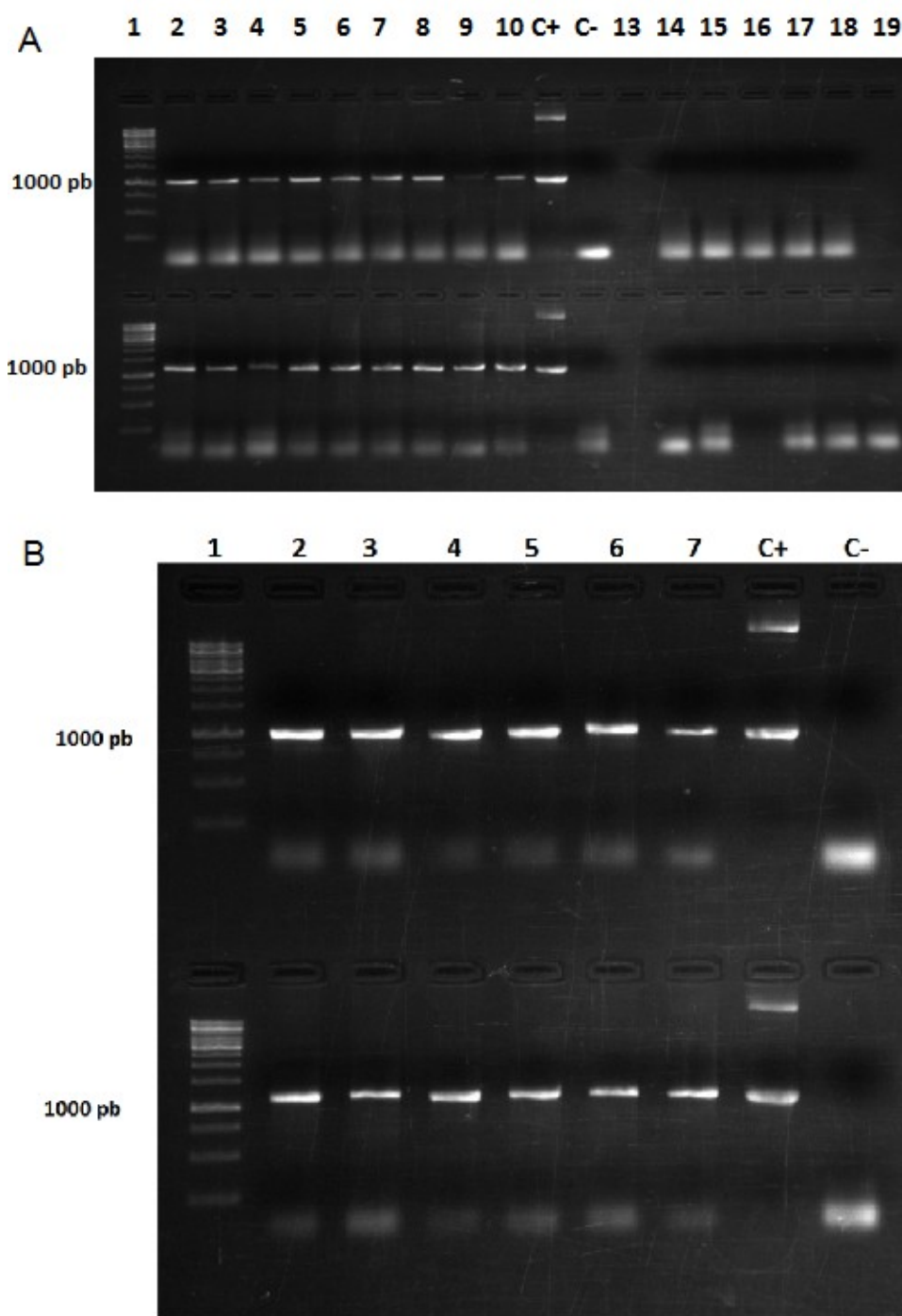


Figura 20 – A) PCR de colônia dos transformantes referentes à transformação da PE-XKS1 Rd. Pente 1: **1.** Marcador de peso molecular 1Kb Plus Fermentas; **2 a 10.** Transformantes da PE-XKS1 Rd testados com os oligos de clonagem

referente ao gene *xyl1*; **11.** Controle positivo; **12.** Controle negativo; **13 a 19.** Transformantes da PE-XKS1 pSP-G1 testados com os oligos de clonagem referente ao gene *xyl1*. Pente 2: **1.** Marcador de peso molecular 1Kb Plus Fermentas; **2 a 10.** Transformantes da PE-XKS1 Rd testados com os oligos de clonagem referente ao gene *xyl2*; **11.** Controle positivo; **12.** Controle negativo; **13 a 19.** Transformantes da PE-XKS1 pSP-G1 testados com os oligos de clonagem referente ao gene *xyl2*. **B)** PCR de colônia dos transformantes referentes à transformação da PE-XKS1 Pb. Pente 1: **1.** Marcador de peso molecular 1Kb Plus Fermentas; **2 a 7.** Transformantes da PE-XKS1 Pb testados com os oligos de clonagem referente ao gene *xyl1*; **8.** Controle positivo; **9.** Controle negativo; Pente 2: **1.** Marcador de peso molecular 1Kb Plus Fermentas; **2 a 7.** Transformantes da PE-XKS1 Pb testados com os oligos de clonagem referente ao gene *xyl2*; **8.** Controle positivo; **9.** Controle negativo.

5.9.1. A atividade enzimática de XR e XDH das cepas PE-2 recombinantes

Para verificar se as construções das cepas de *S. cerevisiae* PE-2 estavam corretas, ou seja, se a superexpressão dos genes estava implicando em um aumento de atividade dessas enzimas, foi quantificada a atividade enzimática de XR e XDH das cepas recombinantes. Assim, foram aferidas as atividades da PE Rd e PE Pb, utilizando como controle positivo as cepas PE 274 (aquela com o plasmídio pRH274 que contém os genes *xyl1* e *xyl2* da *S. stipitis* e o gene *xks1* da *S. cerevisiae* hiperexpresso), PE 209 (com o plasmídio pRH209 que contém os genes *xyl1* e *xyl2* da *S. stipitis*), e a *S. stipitis* NRRL-Y7124 (Fig. 21).

Em relação às atividades de XR, pudemos observar que as cepas recombinantes que contém os genes de *S. stipitis* são as únicas que utilizam o cofator NADH. Em relação às atividades utilizando o NADPH, destacou-se a atividade da PE Rd a qual, apesar de ser inferior a da *S. stipitis* NRRL-Y7124, apresentou-se similar às atividades das construções PE 274 e PE 209. A cepa

recombinante TMB 3001 é uma cepa de laboratório que consegue converter xilose em etanol (JEPPSSON et al., 2003), e apresenta os genes *xyl1* e *xyl2* da *S. stipitis*. A TMB 3001 apresentou atividade usando NADPH como cofator igual a 0,32 U/mg de proteína, assim, podemos considerar que os resultados da PE Rd (0,74 U/mg proteína), PE 274 (0,45 U/mg proteína) e PE 209 (0,59 U/mg proteína) são bastante satisfatórios. Segundo Jeppsson et al. (2003), uma alta atividade de XR pode resultar numa alta conversão de xilose em xilitol, forçando assim, a reação cineticamente desfavorável de XDH, resultando num maior consumo de xilose e menor rendimento de xilitol. Já a cepa PE Pb apresentou atividade abaixo do esperado e semelhante aos controles negativos (PE 416 e PE pSP-G1). Assim, após verificar as construções e certificar que o vetor pSP-G1 contendo os genes de interesse estavam na levedura através de sequenciamento, pode-se deduzir que possivelmente o candidato ao *xyl1* da *P. brasiliensis* GHG001 inferido pela árvore das aldeído redutases pode não estar correto e consequentemente não se trata da xilose redutase de nosso interesse. Isso já havia sido previamente sugerido quando do experimento de qRT-PCR mostrou que o gene *xyl1* da *P. brasiliensis* GHG001 mostrou acumulação de RNAm na presença de glicose em níveis comparáveis aos de xilose.

Em relação às atividades de XDH, observou-se que as atividades das cepas PE 274 (0,072 U/mg proteína), PE 209 (0,051 U/mg proteína) e PE Pb (0,031 U/mg proteína) têm um bom resultado, superando a atividade da XDH endógena de *S. stipitis* (0,027 U/mg), sugerindo que a superexpressão de *xyl2* funcionou. Comparando com a TMB 3001 no trabalho de Eliasson et al. (2000) que apresentou atividade de XDH de 1,78 U/mg proteína, esse resultado não é tão satisfatório quanto de XR.

Além disso, analisando os dois gráficos, vemos que a atividade de XR continua sendo bem maior que a atividade de XDH, resultado contrário às cepas de *S. cerevisiae* construídas no trabalho de Wahlbom et al. (2003), porém similar aos da *S. stipitis* CBS 6054. Esse fato pode estar originando um acúmulo de xilitol nas células. A literatura relata que o desequilíbrio causado pela redução realizada

por XR NADPH-dependente e a oxidação por XDH NAD⁺-dependente, ou seja, o desbalanço de cofatores é responsável pela maioria do xilitol acumulado em cepas de *S. cerevisiae* recombinantes (JEPPSSON et al., 2003).

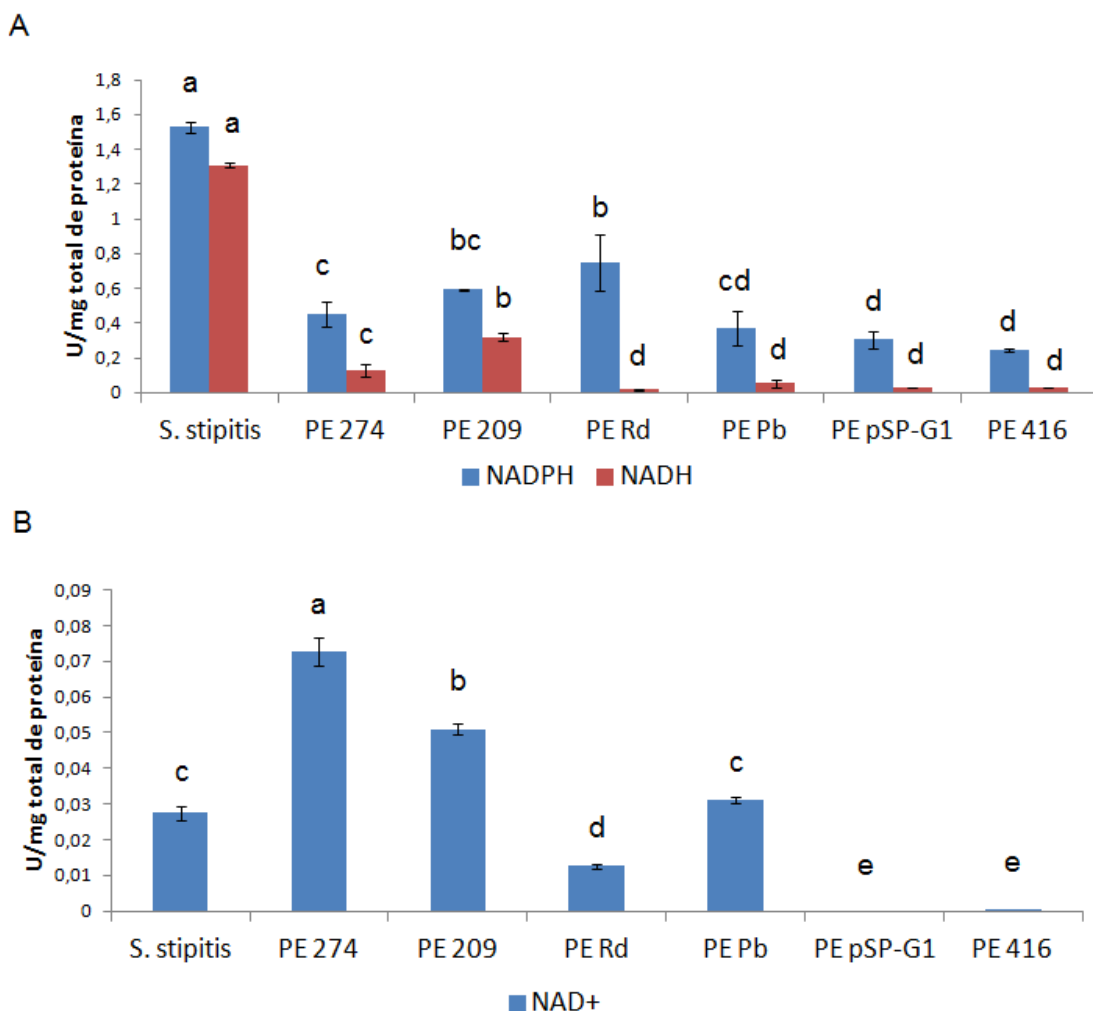


Figura 21 – A) Atividade de XR das cepas PE-2 recombinantes na presença de NADPH e NADH; **B)** Atividade de XDH das cepas PE-2 recombinantes na presença do cofator NAD⁺. As análises estatísticas foram realizadas através do teste de Tukey.

5.9.2. O perfil de crescimento das cepas PE-2 recombinantes

Foi também avaliado o perfil de crescimento das cepas recombinantes PE Rd e PE Pb; como controles, foi utilizada a cepa PE 274 (positivo) e como controle negativo utilizou-se PE pSP-G1 (plasmídio sem genes). Além disso, os perfis de crescimento também foram comparados com a cepa PE 209, que possui os mesmos genes das cepas agora avaliadas, no entanto, provenientes de *S. stiptis* (Figura 22).

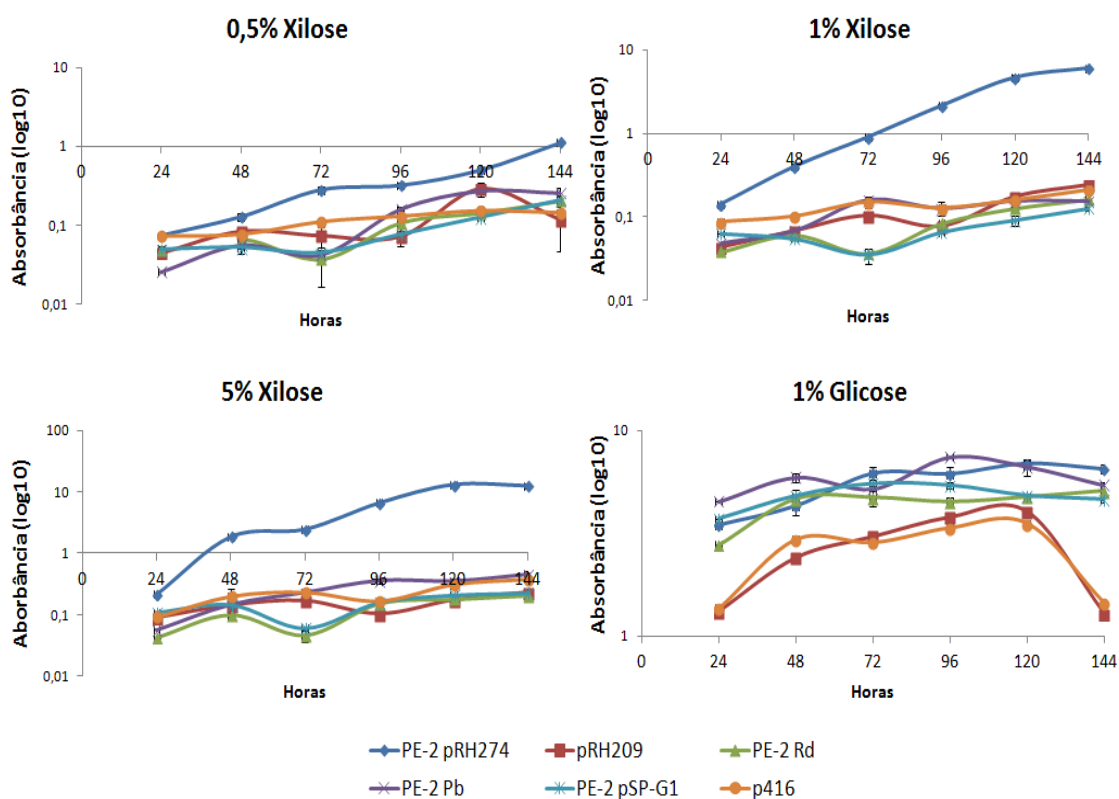


Figura 22 – Perfil de crescimento das cepas PE-2 recombinantes em glicose e em diferentes concentrações de xilose.

Observando as curvas de crescimento dessas cepas em xilose e glicose (Figura 21), nota-se que no meio contendo glicose, todas as cepas cresceram bem, destacando-se inclusive as cepas PE Rd e PE Pb. Entretanto, nos meios

contendo xilose, a PE 274 é a única cepa capaz de utilizar esse açúcar para crescimento. Portanto, vale ressaltar a importância do papel da xilulose quinase na assimilação de xilose, visto que todas as outras cepas, inclusive a PE 209, que possui os mesmos genes de *S. stipitis* *xyl1* e *xyl2* hiperexpressos, apresentaram um crescimento em xilose muito inferior a PE 274, comportamento que corrobora com os resultados de Ho et al.(1998). Nesse trabalho, eles construíram plasmídios com os genes de assimilação de xilose *xyl1* e *xyl2* de *S. stipitis* e o *XKS1* de *S. cerevisiae* e transformaram na cepa de *S. cerevisiae* 1400. Foi visto, então, que a cepa recombinante construída conseguiu crescer em xilose como fonte de carbono e fermentar xilose na presença de glicose (HO et al., 1998). Kahrumaa et al. (2007) também demonstraram que a superexpressão do gene *xks1* é necessária para o consumo de xilose. Ainda, segundo Jin et al. (2002), a superexpressão de *xks1* endógeno da *S. cerevisiae*, com *xyl1* e *xyl2*, aumenta a produção de etanol e diminui a produção de xilitol a partir de xilose. Dessa forma, comprova-se que a inserção dos genes *xyl1+xyl2* combinada com a hiperexpressão de *xks1* seria uma melhor estratégia, por isso foi realizada a transformação destas construções na cepa PE-XKS1.

5.9.3. O perfil de crescimento das cepas PE-XKS1 recombinantes

Segundo Ho et al (1998) a maioria das cepas de *S. cerevisiae* apresentam níveis bastante baixos da enzima XK. Sendo assim, estudos recentes demonstraram que a superexpressão da enzima xilulose quinase (XK), responsável pela fosforilação da xilulose em xilulose 5-fosfato, melhora a utilização de xilose (WAHLBOM et al., 2003). Kahrumaa et al. (2007) relataram que um consumo eficiente de xilose só é observado quando altas atividades de XR e XDH são combinadas com alta atividade de XK, como na cepa recombinante TMB 3062. Sendo assim, foi avaliado o perfil de crescimento das cepas PE-2 recombinantes com superexpressão de *xks1* em um dos alelos, que expressa este gene pelo menos 40% a mais que a PE-2 selvagem.

Foi utilizado como controle a cepa PE-XKS1 274 (positivo) e PE-XKS1 pSP-G1 (negativo). Analisando a figura 23, é possível observar que todas as leveduras crescem em glicose, como esperado. Porém, nos meios contendo xilose, novamente só há crescimento da PE-XKS1 274. Desta forma, existem algumas hipóteses que podem justificar estes resultados: (i) as baixas atividades endógenas de XR e XDH da *R. dairenensis* TAB01 e da *P. brasiliensis* GHG001 e (ii) a permanência do desbalanço redox das enzimas já que não há utilização do cofator NADH pela enzima XR.

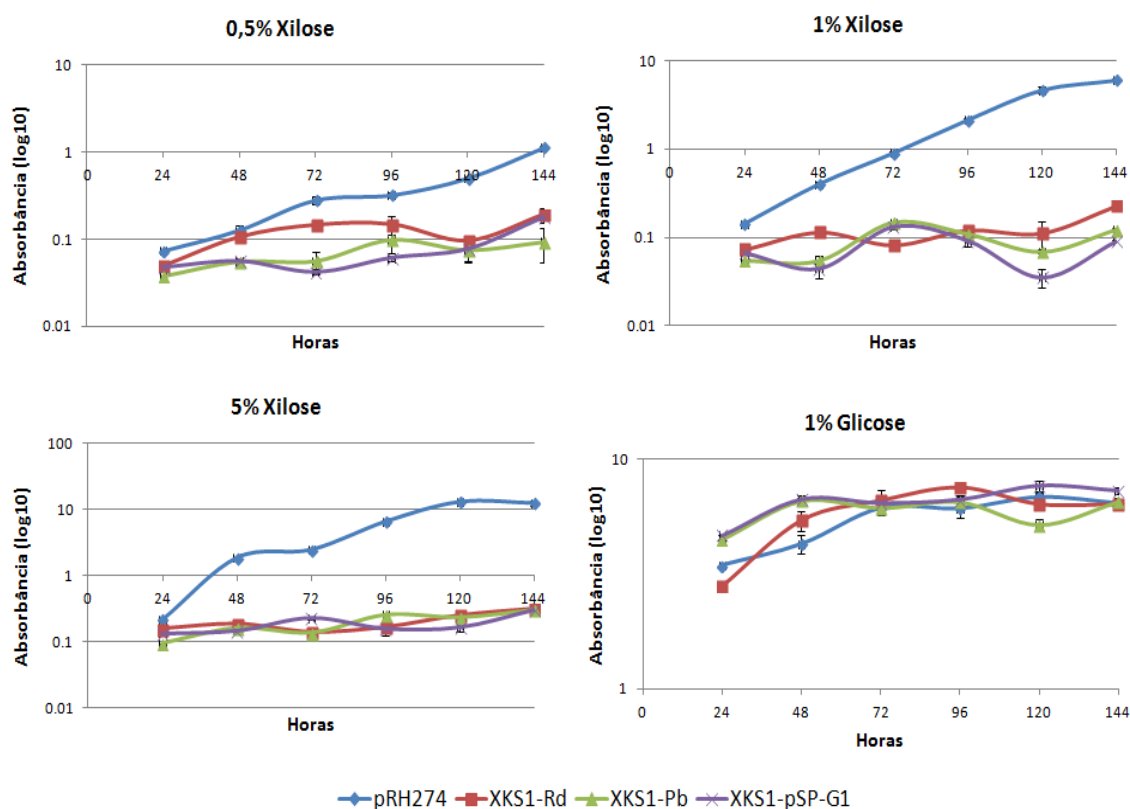


Figura 23 – Perfil de crescimento das cepas PE-XKS1 recombinantes em glicose e em diferentes concentrações de xilose.

Como citado por Wahlbom et al. (2003), a introdução de genes que codificam XR, XDH e XK é necessária devido às baixas atividades dessas enzimas na *S. cerevisiae* selvagem. Contudo, utilizar apenas a engenharia

metabólica não tem se mostrado suficiente para uma utilização de xilose eficiente. Eles propõem então, uma combinação da engenharia metabólica, que é a integração desses genes codificadores de XR, XDH e XK, com a mutagênese aleatória, provando que assim pode-se obter uma *S. cerevisiae* capaz de bons crescimentos em xilose e produção de etanol. Nesse trabalho ele utiliza mutagênese química com etil metanosulfonato, e o melhor mutante gerado (TMB 3400) provou que a combinação das técnicas gerou uma cepa recombinante superior quanto à utilização de xilose. Portanto, essa abordagem poderia ser utilizada nos próximos estudos com as leveduras deste trabalho.

6. Artigo: *Pseudozyma brasiliensis* sp. nov., a novel xylanolytic Ustilaginomycetous yeast species isolated from an insect pest of sugarcane roots

Uma das leveduras isoladas neste trabalho foi identificada como uma nova espécie, sendo denominada *Pseudozyma brasiliensis* sp. nov. GHG001. Essa espécie da ordem Ustilaginales foi isolada do trato intestinal de uma parasita da cana-de-açúcar pertencente à família *Chrysomelidae* e identificada com base em análises filogenéticas e moleculares das regiões D1/D2 do LSU e ITS, sendo seu parente mais próximo filogeneticamente, a *P. vetiver*. Maiores detalhes e informações podem ser visualizadas no artigo a seguir (artigo aceito para publicação).

***Pseudozyma brasiliensis* sp. nov., a novel xylanolytic Ustilaginomycetous yeast species isolated from an insect pest of sugarcane roots**

Juliana Velasco de Castro Oliveira¹, Thuanny A. Borges¹, Renato Augusto Corrêa dos Santos¹, Larissa F. D. Freitas², Carlos Augusto Rosa², Gustavo Henrique Goldman^{1,3#}, and Diego Mauricio Riaño-Pachón^{1#}

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

² Departamento de Microbiologia, Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas, Belo Horizonte, Minas Gerais - Brazil

³ Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil

Corresponding authors: ggoldman@usp.br and diego.riano@bioetanol.org.br

Contents category: New Taxa – Eukaryotic micro-organism

Running title: *Pseudozyma brasiliensis* sp. nov.

The GenBank accession number for SSU ITS LSU sequence of *Pseudozyma brasiliensis* is KF737866

32 A novel Ustilaginomycetous yeast isolated from the intestinal tract of an insect pest of
 33 sugarcane roots in Ribeirão Preto, São Paulo State, Brazil, represents a novel *Pseudozyma*
 34 species based on molecular analyses of the D1/D2 LSU and the ITS (ITS1 + ITS2) regions.
 35 The name *Pseudozyma brasiliensis* sp. nov. is proposed for this species with GHG001^T (=
 36 CBS 13268^T = UFMG-CM-Y307^T) as type strain. *P. brasiliensis* sp. nov. is the sister
 37 species of *P. vetiver*, originally isolated from leaves of vetiver grass and sugarcane in
 38 Thailand. *P. brasiliensis* sp. nov. is able to grow well in xylan as sole carbon source and
 39 produces high levels of an endo-1,4-xylanase, which has a higher specific activity in
 40 comparison to other eukaryotic xylanases. This enzyme has a variety of industrial
 41 applications, indicating the great biotechnological potential of *P. brasiliensis*.

42
 43 The anamorphic yeast-like genus *Pseudozyma* belongs to the order Ustilaginales
 44 (Ustilaginomycetes, Ustilaginomycotina), with close relationship to species of the
 45 teleomorphic genera *Ustilago* and *Sporisorium*, which are responsible for serious plant
 46 diseases. *Pseudozyma* encompasses a small group of basidiomycetous yeasts that have
 47 20 species described: *P. abaconensis*, *P. alboarmeniaca*, *P. antarctica*, *P. aphidis*, *P.*
 48 *churashimaensis*, *P. flocculosa*, *P. fusiformata*, *P. graminicola*, *P. crassa*, *P. hubeiensis*, *P.*
 49 *jejuensis*, *P. parantarctica*, *P. prolifica*, *P. pruni*, *P. rugulosa*, *P. shanxiensis*, *P. siamensis*
 50 *P. thailandica*, *P. tsukubaensis* and *P. vetiver* (Wang *et al.*, 2006; Golubev, 2007; Seo *et al.*,
 51 2007; Liou *et al.*, 2009; Statzell-Tallman *et al.*, 2010; Boekhout, 2011; Konishi *et al.*, 2011;
 52 Chamnanpa *et al.*, 2013; Mekha *et al.*, 2014). Many of these are isolated most frequently
 53 from plants, although they are not plant pathogens. *P. parantarctica*, *P. thailandica*, *P.*
 54 *alboarmeniaca*, *P. crassa* and *P. siamensis* (Sugita *et al.*, 2003; Mekha *et al.*, 2014) were
 55 isolated from blood samples and *P. abaconensis* was isolated from coral reef waters
 56 (Statzell-Tallman *et al.*, 2010).

57 Noteworthy, members of this genus are able to produce a range of biotechnologically
 58 relevant biomolecules such as, the enzymes cutinases (Seo *et al.*, 2007), lipases
 59 (Bussamara *et al.*, 2010; Bussamara *et al.*, 2012), and xylanases (Adsul *et al.*, 2009;
 60 Borges *et al.*, 2013); and biosurfactants as mannopyranosyl-meso-erythritol (Morita *et al.*,
 61 2009; Morita *et al.*, 2010; Konishi *et al.*, 2011). Furthermore, some species have been used
 62 as biocontrol agents, particularly targeting plant pathogens (Avis & Belanger, 2002;
 63 Marchand *et al.*, 2009).

During the study of the yeast-like organisms associated to the intestinal tract of insects found in the roots of sugarcane (*Saccharum* spp.) plantations located in Ribeirão Preto, São Paulo State, Brazil, a novel species of the genus *Pseudozyma* was identified. Particularly, we were looking for yeasts that metabolize either xylose or their polysaccharides, e.g., xylan. A single isolate (GHG001) showed ability to grow very well in these sugars as single carbon source and produced high levels of an endoxylanase (PbXynA) (Borges *et al.*, 2013). Although a diversity of filamentous fungi and bacteria is able to produce xylanase, this feature is rare in yeasts. PhXylA is an endo-1,4-xylanase from the GH family 11 that shows higher specific activity than other eukaryotic xylanases previously studied (Borges *et al.*, 2013). Most of the xylanases reported so far produce xylooligosaccharides along with xylose as end-products (Adsul *et al.*, 2009). Particularly, PhXylA produces small (degree of polymerization around 3 to 7) xylooligosaccharides (XOS) with negligible production of monomers as one of the end-products of xylan hydrolysis. It is worth noting that this property of this xylanase can enhance XOS production since little raw material will be lost as xylose. XOS can be used as ingredients of functional food, selectively stimulating the growth of beneficial bacteria in the digestive system and also facilitating nutrient absorption. In addition, XOS have antioxidant and hypolipemic activity (against cholesterol and triglycerides), can inhibit melanin and melanoma cell proliferation, it has anti-inflammatory activity, and function as anti-allergy agents (Moure *et al.*, 2006). Besides XOS production, xylanases are responsible for the degradation of the heteroxylans constituting the lignocellulosic plant cell wall, and can be used for the enzymatic saccharification of the plant biomass and second-generation bioethanol production. In addition, xylanases have a wide range of applications in industry, for example in the manufacture of paper, food, animal feed, pharmaceuticals and textiles (Polizeli *et al.*, 2005). Thus, this novel species of the genus *Pseudozyma* has a high biotechnological potential, and on the basis of phenotypic and phylogenetic analysis we proposed the new species as *Pseudozyma brasiliensis* sp. nov.

91

92 *Yeast Isolation and identification*

The yeast strain GHG001 was isolated from the intestinal tract of a Chrysomelidae larvae associated to sugarcane roots. Briefly, Chrysomelidae at larval stages were collected

95 from *Saccharum* spp. roots in sugarcane plantations located in Ribeirão Preto, São Paulo,
 96 Brazil. The larvae were anesthetized at 4 °C for 10 min, then externally sterilized with 70%
 97 ethanol for 1 minute, and rinsed twice with sterilized water. After larvae dissection, the
 98 intestinal juices were transferred to a sterile flask containing 1 mL of cold sterile water. A
 99 small volume of each gut sample (50 µL) was then inoculated into 250 mL Erlenmeyer flask
 100 containing 50 mL of minimal medium (MM) composed of 50 mL solution salt L⁻¹ (6 g NaNO₃
 101 L⁻¹, 1.5 g KH₂PO₄ L⁻¹, 0.5 g KCl L⁻¹ and 0.5 g MgSO₄ L⁻¹) and 1 mL trace elements L⁻¹ (10 g
 102 EDTA L⁻¹, 4.4 g ZnSO₄·H₂O L⁻¹, 1.0 g MnCl₂·4H₂O L⁻¹, 0.32 g CoCl₂·6H₂O L⁻¹, 0.315 g
 103 CuSO₄·5H₂O L⁻¹, 0.22 g (NH₄)₆Mo₇O₂₄·4H₂O L⁻¹, 1.47 g CaCl₂·2H₂O L⁻¹ and 1 g
 104 FeSO₄·H₂O L⁻¹) pH 6.5 with 1% xylose, supplemented with a cocktail of antibiotics (100 µg
 105 ampicillin µL⁻¹, 34 µg kanamycin µL⁻¹, 20 µg tetracycline µL⁻¹ and 35 µg chloramphenicol
 106 µL⁻¹). The inoculum was incubated on a rotary shaker at 30 °C for enrichment of
 107 microorganisms that could use xylose more efficiently. Two weekly transfers were
 108 performed by changing the inoculum to fresh medium and after the second week, 100 µL of
 109 a tenfold dilution were spread on solid MM with 1% xylose (w/v), and incubated at 30 °C
 110 until colonies appeared. Yeast colonies were repeatedly streaked out until pure cultures
 111 were obtained. The isolated strains were growth in YPD (1% yeast extract (w/v), 2%
 112 peptone (w/v), 2% glucose (w/v)), and stored at -80 °C (vials of cell suspensions in 1 mL
 113 YPD/10% glycerol (v/v)). Among the isolates, strain GHG001^T was selected due its high
 114 capacity to grow on xylose (Borges *et al.*, 2013).

115 Morphological description and physiological tests were carried out by standard
 116 methods (Kurtzman *et al.*, 2011). Presence of budding cells, hyphae and pseudohyphae
 117 was investigated by cultivation on potato dextrose agar (PDA), diluted V8 and YM agar at
 118 20 and 25 °C for up to 21 days and MM containing 0.5% of ammonium sulphate at 30 °C for
 119 up 5 days. Basidiospore formation was investigated on 5% malt extract agar (w/v), corn
 120 meal, PDA and diluted V8 agars at 20 and 25 °C for up to 21 days.

122 *DNA sequence and phylogenetic analysis*

123 Due to its biotechnological potential we have sequenced the genome of this
 124 strain. Briefly, the extraction of genomic DNA followed the protocol proposed by Sambrook
 125 & Russell (Sambrook & Russell, 2001), and then whole genome shotgun sequence was

126 carried out on an Illumina HiSeq2000 at the University of Southern California Epigenome
 127 Center Data Production Facility, Los Angeles, California, USA. The genome sequence was
 128 assembled into 45 scaffolds using Velvet and SSPACE BASIC (Accession number:
 129 AWXO000000000). The full ITS region (ITS1 + ITS2, including 5.8S rDNA) as well as the
 130 LSU D1/D2 region were extracted from the assembled genome sequence and deposited at
 131 GenBank under accession number KF737866 (Oliveira *et al.*, 2013). The sequence was
 132 compared using a MEGABLAST similarity search (Altschul *et al.*, 1990) against the NCBI
 133 nucleotide database to identify the genus of the yeast. The genus assignment was
 134 confirmed using the ITS2 sequence as a query in a search in the ITS2 database
 135 (Koetschan *et al.*, 2012). Further sequence comparisons were carried out with the
 136 *Pseudozyma* spp. sequences reported by Chamnanpa *et al.* (2013), Liou *et al.* (2009),
 137 Statzell-tallman *et al.* (2010), Seo *et al.* (2007) and Morita *et al.* (2011). Sequences of each
 138 individual region, i.e., ITS1, 5.8S, ITS2, D1/D2 were structurally aligned using MXSCARNA
 139 (Tabei *et al.*, 2008). Regions were concatenated into a supermatrix using FASConCAT
 140 (Kuck & Meusemann, 2010). Phylogenetic inference was carried out with the Maximum
 141 Likelihood method implemented in the software RAxML (Stamatakis, 2006). Each region
 142 was kept as an independent partition for the analysis. We used different models for stem
 143 and loop regions. For the loop regions the evolutionary model GTR was used. The stem
 144 regions were analyzed under the 7A evolutionary model, thus taking into account
 145 secondary structure topology, i.e., compensatory mutations (Higgs, 2000; Jow *et al.*, 2002;
 146 Hudelot *et al.*, 2003). Substitution rate heterogeneity was taken into account using the
 147 gamma-model of Yang (Yang, 1996). Bootstrap values were computed over 1000
 148 replicates.

149 Genome-wide comparisons were carried out against the genome sequences of
 150 *Pseudozyma antarctica* T-34 (ACC: GCA_000334475.1), *Pseudozyma hubeiensis* SY62
 151 (ACC: GCA_000403515.1), *Pseudozyma flocculosa* PF-1 (ACC: GCA_000417875.1),
 152 *Pseudozyma aphidis* DSM 70725 (ACC: GCA_000517465.1), *Ustilago maydis* 521 (ACC:
 153 GCF_000328475.1) and *Malassezia globosa* [CBS 7966](#) (ACC: GCF_000181695.1).
 154 Clusters of orthologous genes, i.e., putative gene families, were identified using OrthoMCL
 155 with an inflation value of 2.5 (Li *et al.*, 2003).

156 *Pseudozyma* spp. are not known to be plant pathogens, as such we decided to look
 157 for typical pathogenicity genes found in the closely related species *U. maydis*. In order to

158 enter into the pathogenic state, yeast-like cells must mate, i.e., fuse cells of different
 159 compatibility to form a dikaryotic mycelium, thus triggering a sexual developmental phase.
 160 The genes responsible for the different compatibility are located, in *U. maydis* and other
 161 pathogenic species, in what is now called the mating-type locus, and different alleles of
 162 these genes are responsible of different mating types (Bakkeren & Kronstad, 1994; Urban
 163 *et al.*, 1996; Casselton & Olesnick, 1998; Casselton, 2002; Smith *et al.*, 2004; Lefebvre *et al.*,
 164 2013). We used as baits the protein sequences of the genes in *U. maydis*, and
 165 performed `tblastn` and `blastp` similarity searches to identify the homologues. Pairwise
 166 alignments were then carried out with `water`, a program from the EMBOSS suite, that
 167 implements the pairwise local alignment algorithm originally described by Smith and
 168 Waterman (Smith & Waterman, 1981; Rice *et al.*, 2000).

169

170 We have isolated a strain named GHG001^T from the intestinal tract of Chrysomelidae
 171 that can grow on xylose and xylan as sole carbon sources and produces high levels of
 172 xylanase when grown in xylan. Initial identification based on ITS2 region revealed that this
 173 strain belongs to the genus *Pseudozyma*, and a more detailed study based on
 174 morphological and molecular typing (D1/D2 LSU + ITS) showed that this strain was a novel
 175 yeast species in the genus *Pseudozyma*.

176 A sequence similarity search using MEGABLAST against the GenBank database
 177 and the ITS2 region as query revealed that GHG001^T belongs to the genus *Pseudozyma*.
 178 Accurate identification at the species level was carried out with the complete ITS regions,
 179 5.8S rDNA and the D1/D2 LSU region, which were extracted from the genome sequence.
 180 The Maximum Likelihood phylogenetic tree showed that strain GHG001^T forms a clade with
 181 *P. vetiver* DMKU-LV99^T and *P. fusiformata* JCM 3931^T (Fig. 1). The main differences
 182 between these species are found on the ITS region (ITS1 + ITS2). GHG001^T has 19
 183 substitutions (and 6 indels) compared to *P. vetiver*, and 40 substitutions (32 indels)
 184 compared to *P. fusiformata*. On the D1/D2 LSU region there are 3 base substitutions,
 185 additionally one indel, between strain GHG001^T and *P. vetiver*, and 14 substitutions, plus 2
 186 indels, when compared with *P. fusiformata* (Phylogenetic tree and multiple sequence
 187 alignment are available in TreeBase with accession number TB2:S14822). The number of
 188 substitutions found is suggestive of a new species in the genus; as strains differing by two

189 or more nucleotides on the D1/D2 LSU region could represent different taxa (Fell *et al.*,
190 2000).

191 Physiological characteristics of *P. brasiliensis* sp. nov. and all other *Pseudozyma*
192 species are presented in Table 1. In terms of physiological characteristics, strain GHG001^T
193 differed from *P. vetiver*, its closest relative, on its ability to assimilate inulin, D-galactose, L-
194 arabinose, D-ribose, ethanol, glycerol, erythritol, ribitol and D-gluconic acid. In addition, *P.*
195 *brasiliensis* sp. nov. was unable to assimilate soluble starch and lactic and citric acids. No
196 mating reaction or sexual structure was observed when the single isolate of *P. brasiliensis*
197 sp. nov. was grown on 5% malt extract agar, corn meal, PDA and diluted V8 agars at 20
198 and 25 °C for up to 21 days. Notably, although pseudohyphae formation is a characteristic
199 of the genus *Pseudozyma* grown on 5% malt extract agar or Dalmau plate (Boekhout,
200 2011), we only observed *P. brasiliensis* forming pseudohyphae in minimal medium
201 containing 0.5% ammonium sulphate (Fig. 2). We also looked for the mating type locus and
202 related genes in the genome of *P. brasiliensis* sp. nov., and other *Pseudozyma* species
203 whose genome sequence is available using as reference the genes found in *U. maydis*
204 (Table 2, Levebvre *et al.*, 2013). Both *a* and *b* loci, responsible for mating-type are present
205 in *P. brasiliensis* sp. nov., as well as in all the other *Pseudozyma* spp., indicating that *P.*
206 *brasiliensis* sp. nov. is a homothallic yeast as well as all the other *Pseudozyma* spp.
207 included in this analysis (Table 2). The pheromone receptor genes, *mf*, *prf* and *rba1*, are
208 found in the *a* locus, which in *P. brasiliensis* sp. nov. is located in scaffold
209 PSEUBRA_SCAF4. Particularly the *mfa* genes is not annotated in the current release of the
210 genome sequence of *P. brasiliensis* sp. nov., but the *tblastn* search clearly revealed that
211 it is encoded by the genome, the same is true for *P. hubeiensis* and *P. antarctica*, and it is
212 most likely related to the small size of this gene, which will be overlooked by most *ab initio*
213 gene prediction software. On the other hand, the genes *bW* and *bE*, locus *b*, which are
214 homeodomain transcription factors, are located in scaffold PSEUBRA_SCAF3. There is as
215 well a *smu1* homologue in *P. brasiliensis* sp. nov., located in scaffold PSEUBRA_SCAF14.
216 Smu1 is a protein kinase required for pathogenicity and mating (Smith *et al.*, 2004). The
217 product of the *prf* gene regulates the transcription of the gene encoded by loci *a* and *b*
218 (Hartmann *et al.*, 1999), and in *P. brasiliensis* sp. nov. is located in scaffold
219 PSEUBRA_SCAF5. Thus, *P. brasiliensis* sp. nov. appears to have a full complement of the
220 most important genes related to mating, even more, these genes are present in all

221 *Pseudozyma* species sequenced up to now. It still remains to be proved whether these
222 genomic loci are functional, i.e., transcribed and translated into the active protein. Recent
223 results showed that there was not mating interaction between *P. flocculosa* and *U. maydis*
224 (Lefebvre *et al.*, 2013).

225 Just a few species in the genus *Pseudozyma* have had their genome fully
226 sequenced so far. These are: *P. antarctica* T-34 (ACC: GCA_000334475.1; Morita *et al.*,
227 2013), *P. aphidis* DSM 70725 (ACC: GCA_000517465.1; (Lorenz *et al.*, 2014), *P.*
228 *hubeiensis* SY62 (ACC: GCA_000403515.1; Konishi *et al.*, 2013), *P. flocculosa* PF-1 (ACC:
229 GCA_000417875.1; Lefebvre *et al.*, 2013) and *P. brasiliensis* GHG001^T (ACC:
230 GCA_000497045.1; Oliveira *et al.*, 2013). All *Pseudozyma* species have a genome of
231 similar size, around 18Mbp, as estimated by the total scaffold length, *P. flocculosa* has the
232 largest genome, i.e., 23.31Mbp and *P. brasiliensis* sp. nov., currently has, the smallest
233 17.33Mbp. They all have similar numbers of protein-coding genes. Although it is important
234 to note that all these genome assemblies are still in a draft stage (Table 3). There have
235 been no report of species of *Pseudozyma* that are pathogens, which is in contrast to other
236 species in the Ustilaginales, such as *Ustilago maydis*, or in other Ustilamycotina, as
237 *Malessezia globosa*. The gene content in these species do not differ to a large extent, but
238 there are clear differences in the GC content. For instance, *Pseudozyma* spp. tend to have
239 a higher GC content than *U. maydis* or *M. globosa* genome-wide. This trend is even
240 stronger when looking at the coding regions in the genome. Particularly, regarding the gene
241 content, we have clustered all the predicted proteins of these *Pseudozyma* species,
242 including also *U. maydis* (ACC: GCF_000328475.1; (Kamper *et al.*, 2006) and *M. globosa*
243 [CBS 7966](#) (ACC: GCF_000181695.1; (Xu *et al.*, 2007), uncovering 6608 putative protein-
244 coding gene families. Fig. 3 shows the number of protein-coding gene families among the
245 different species. The species inside the genus *Pseudozyma* share 62.6% of the families
246 (Fig. 3(a)), this number drops to 45.1% when including the pathogenic species, *U. maydis*
247 and *M. globosa* (Fig. 3(b)). It is also important to note that there are lineage-specific gene
248 families, and these might be directly involved with the maintenance of the particular
249 lifestyles of these groups or species.

250 The molecular and physiological differences found strongly support that strain
251 GHG001 is a new species in the genus *Pseudozyma*. The new species was isolated from
252 the intestinal tract of a Chrysomelidae larvae associated to sugarcane roots. However, only

one isolate *P. brasiliensis* sp. nov. was isolated in the present study. Probably, *P. brasiliensis* sp. nov. is a minor component of the yeast communities associated with the insects living in sugarcane roots. We describe this new species based on only a single strain because of their ability to produce high levels of an endoxylanase, a biotechnological trait of industrial interest for production functional foods, and biofuels through degradation of the heteroxylans of plant hydrolysates.

Based on the phenotypic and phylogenetic analyses, strain GHG001^T (= CBS 13268^T = UFMGCM-Y307^T) represents a novel species in the genus *Pseudozyma*, for which we propose the name of *Pseudozyma brasiliensis* sp. nov. Finally, we present and updated key for the identification of species in the genus *Pseudozyma* as supplementary material 1.

Description *Pseudozyma brasiliensis* sp. nov. Oliveira, Borges, Santos, Freitas, Rosa, Goldman & Riaño-Pachón.

The epithet *brasiliensis* (bra.si.li.en'sis) N.L. fem. adj. *brasiliensis*, of or pertaining to the country (Brazil) where this yeast was found.

Growth on YM agar after 3 days at 25 °C, cells are ellipsoidal to fusoid and variable in size (1.5-4 x 2-6 µm), and contain what appear to be oil droplets (Fig. 4a and 4b). Budding is polar on a short denticle. On YM agar after 3 days at 25 °C, colonies are smooth to wrinkled, butyrous, glistening, cream-colored to light salmon with an entire margin. Pseudohyphae are formed when grown on minimal medium containing 0.5% ammonium sulphate. After 3 weeks in Dalmau plate culture on cornmeal agar, pseudomycelium and true mycelium are not formed. Sexual reproduction was not observed. Ballistoconidia are not produced. Fermentation is absent. Assimilation of carbon sources: D-glucose, sucrose, raffinose, galactose, trehalose, maltose, melezitose, cellobiose, salicin, L-sorbose, D-xylose, L-arabinose, D-ribose, ethanol, glycerol, erythritol, ribitol, D-mannitol, D-glucitol, myo-inositol (latent), succinic acid, D-gluconate, D-glucosamine (latent), N-Acetyl-D-glucosamine and xylitol (latent) are positive. No growth was detected in inulin, melibiose, lactose, L-rhamnose, D-arabinose, methanol, galactitol, DL-lactate, citrate and hexadecane. Assimilation of nitrogen compounds: positive for nitrate, nitrate, lysine and cadaverine (latent). Growth in amino-acid-free medium is positive. Growth at 37 °C is negative. Growth on YM agar with 10% sodium chloride is negative. Growth in 50% glucose is negative.

283 Starch-like compounds are not produced. In 100 µg cycloheximide mL⁻¹ growth is negative.
 284 Diazonium Blue B reaction is positive. The type strain accession number of *P. brasiliensis*
 285 sp. nov. is GHG001^T. It was isolated from intestinal tract of a Chrysomelidae larvae
 286 associated with sugarcane roots in sugarcane plantations located in Ribeirão Preto, São
 287 Paulo, Brazil. It has been deposited in the collection of the Yeast Division of the
 288 Centraalbureau voor Schimmelcultures, Utrecht, the Netherlands, as strain CBS 13268, in
 289 the Culture Collection of Microorganisms and Cells of the Federal University of Minas
 290 Gerais (UFMG), Belo Horizonte, Minas Gerais, Brazil as strain UFMG-CM-Y307, and in the
 291 Brazilian Collection of Environmental and Industrial Microorganisms (Coleção Brasileira de
 292 Micro-organismos de Ambiente e Indústria, CBMAI), Campinas, São Paulo, Brazil, as strain
 293 CBMAI 1606. The Mycobank number is MB 805996.

294

295 Acknowledgements

296

297 TAB and RACS hold FAPESP scholarships (2012/00080-1 and 2011/22690-3). We thank
 298 Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP 10/513224-2),
 299 Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação do
 300 Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG) and Vale S.A., from Brazil, for
 301 financial support.

302

303 References

- 304 Adsul, M. G., Bastawde, K. B. & Gokhale, D. V. (2009). Biochemical characterization of two xylanases from
 305 yeast *Pseudozyma hubeiensis* producing only xylooligosaccharides. *Bioresource technology* **100**,
 306 6488-6495.
- 307 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool.
 308 *Journal of molecular biology* **215**, 403-410.
- 309 Avis, T. J. & Belanger, R. R. (2002). Mechanisms and means of detection of biocontrol activity of
 310 *Pseudozyma* yeasts against plant-pathogenic fungi. *FEMS yeast research* **2**, 5-8.
- 311 Bakkeren, G. & Kronstad, J. W. (1994). Linkage of mating-type loci distinguishes bipolar from tetrapolar
 312 mating in basidiomycetous smut fungi. *Proc Natl Acad Sci U S A* **91**, 7085-7089.
- 313 Boekhout, T. (2011). *Pseudozyma* Bandoni emend. Boekhout (1985) and a comparison with the yeast state
 314 of *Ustilago maydis* (De Candolle) Corda (1842). In *The yeasts, a taxonomic study*, 5th edn, pp. 1139-
 315 1494. Edited by C. P. Kurtzman, J. W. Fell & T. Boekhout. Amsterdam: Elsevier.
- 316 Borges, T. A., Souza, A. T., Squina, F. M., Riaño-Pachón, D. M., Santos, R. A. C., Machado, E., Oliveira, J. V.
 317 C., Damásio, R. L. & Goldman, G. H. (2013). Biochemical characterization of an endoxylanase from

318 *Pseudozyma brasiliensis* sp. nov. strain GHG001 isolated from the intestinal tract of Chrysomelidae
319 larvae associated to sugarcane roots. *Process Biochemistry* **49**, 77-83.

320 Bussamara, R., Dall'agnol, L., Schrank, A., Fernandes, K. F. & Vainstein, M. H. (2012). Optimal Conditions for
321 Continuous Immobilization of *Pseudozyma hubeiensis* (Strain HB85A) Lipase by Adsorption in a
322 Packed-Bed Reactor by Response Surface Methodology. *Enzyme research* **2012**, 329178.

323 Bussamara, R., Fuentefria, A. M., de Oliveira, E. S., Broetto, L., Simcikova, M., Valente, P., Schrank, A. &
324 Vainstein, M. H. (2010). Isolation of a lipase-secreting yeast for enzyme production in a pilot-plant
325 scale batch fermentation. *Bioresource technology* **101**, 268-275.

326 Casselton, L. A. (2002). Mate recognition in fungi. *Heredity* **88**, 142-147.

327 Casselton, L. A. & Olesnick, N. S. (1998). Molecular genetics of mating recognition in basidiomycete fungi.
328 *Microbiology and molecular biology reviews* : MMBR **62**, 55-70.

329 Chamnanpa, T., Limtong, P., Srisuk, N. & Limtong, S. (2013). *Pseudozyma vetiver* sp. nov., a novel
330 anamorphic ustilaginomycetous yeast species isolated from the phylloplane in Thailand. *Antonie Van*
331 *Leeuwenhoek* **104**, 637-644.

332 Fell, J. W., Boekhout, T., Fonseca, A., Scorzetti, G. & Statzell-Tallman, A. (2000). Biodiversity and
333 systematics of basidiomycetous yeasts as determined by large-subunit rDNA D1/D2 domain
334 sequence analysis. *International journal of systematic and evolutionary microbiology* **50 Pt 3**, 1351-
335 1371.

336 Golubev, V. I. (2007). Mycotinogeny in smut yeast-like fungi of the genus *Pseudozyma*. *Mikrobiologiya* **76**,
337 813-816.

338 Hartmann, H. A., Kruger, J., Lottspeich, F. & Kahmann, R. (1999). Environmental signals controlling sexual
339 development of the corn Smut fungus *Ustilago maydis* through the transcriptional regulator Prf1.
340 *Plant Cell* **11**, 1293-1306.

341 Higgs, P. G. (2000). RNA secondary structure: physical and computational aspects. *Quarterly reviews of*
342 *biophysics* **33**, 199-253.

343 Hudelot, C., Gowri-Shankar, V., Jow, H., Rattray, M. & Higgs, P. G. (2003). RNA-based phylogenetic
344 methods: application to mammalian mitochondrial RNA sequences. *Molecular phylogenetics and*
345 *evolution* **28**, 241-252.

346 Jow, H., Hudelot, C., Rattray, M. & Higgs, P. G. (2002). Bayesian phylogenetics using an RNA substitution
347 model applied to early mammalian evolution. *Molecular biology and evolution* **19**, 1591-1601.

348 Kamper, J., Kahmann, R., Bolker, M., Ma, L. J., Brefort, T., Saville, B. J., Banuett, F., Kronstad, J. W., Gold, S.
349 E. & other authors (2006). Insights from the genome of the biotrophic fungal plant pathogen
350 *Ustilago maydis*. *Nature* **444**, 97-101.

351 Koetschan, C., Hackl, T., Muller, T., Wolf, M., Forster, F. & Schultz, J. (2012). ITS2 database IV: interactive
352 taxon sampling for internal transcribed spacer 2 based phylogenies. *Molecular phylogenetics and*
353 *evolution* **63**, 585-588.

354 Konishi, M., Nagahama, T., Fukuoka, T., Morita, T., Imura, T., Kitamoto, D. & Hatada, Y. (2011). Yeast
355 extract stimulates production of glycolipid biosurfactants, mannosylerythritol lipids, by *Pseudozyma*
356 *hubeiensis* SY62. *Journal of bioscience and bioengineering* **111**, 702-705.

357 Kuck, P. & Meusemann, K. (2010). FASconCAT: Convenient handling of data matrices. *Molecular*
358 *phylogenetics and evolution* **56**, 1115-1118.

359 Kurtzman, C. P., Fell, J. W., Boekhout, T. & Robert, V. (2011). Methods for isolation, phenotypic
360 characterization and maintenance of yeasts. In *The yeasts, a taxonomic study*, 5th edn. Edited by K.
361 C.P., J. W. Fell & T. Boekhout. Amsterdam: Elsevier.

362 Lefebvre, F., Joly, D. L., Labbe, C., Teichmann, B., Linning, R., Belzile, F., Bakkeren, G. & Belanger, R. R.
363 (2013). The transition from a phytopathogenic smut ancestor to an anamorphic biocontrol agent
364 deciphered by comparative whole-genome analysis. *Plant Cell* **25**, 1946-1959.

365 Li, L., Stoeckert, C. J., Jr. & Roos, D. S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic
366 genomes. *Genome Res* **13**, 2178-2189.

367 Liou, G. Y., Wei, Y. H., Lin, S. J., Wen, C. Y. & Lee, F. L. (2009). *Pseudozyma pruni* sp. nov., a novel
368 ustilaginomycetous anamorphic fungus from flowers in Taiwan. *International journal of systematic
369 and evolutionary microbiology* **59**, 1813-1817.

370 Lorenz, S., Guenther, M., Grumaz, C., Rupp, S., Zibek, S. & Sohn, K. (2014). Genome Sequence of the
371 Basidiomycetous Fungus *Pseudozyma aphidis* DSM70725, an Efficient Producer of Biosurfactant
372 Mannosylerythritol Lipids. *Genome Announc* **2**, e00053-00014.

373 Marchand, G., Remus-Borel, W., Chain, F., Hammami, W., Belzile, F. & Belanger, R. R. (2009). Identification
374 of genes potentially involved in the biocontrol activity of *Pseudozyma flocculosa*. *Phytopathology* **99**,
375 1142-1149.

376 Mekha, N., Takashima, M., Boon-Long, J., Cho, O. & Sugita, T. (2014). Three new basidiomycetous yeasts,
377 *Pseudozyma albaarmeniaca*, sp. nov., *P. crassa* sp. nov., *P. slamensis* sp. nov. Isolated from Thai
378 patients. *Microbiology and immunology* **58**, 9-14.

379 Morita, T., Takashima, M., Fukuoka, T., Konishi, M., Imura, T. & Kitamoto, D. (2010). Isolation of
380 basidiomycetous yeast *Pseudozyma tsukubaensis* and production of glycolipid biosurfactant, a
381 diastereomer type of mannosylerythritol lipid-B. *Applied microbiology and biotechnology* **88**, 679-
382 688.

383 Morita, T., Fukuoka, T., Konishi, M., Imura, T., Yamamoto, S., Kitagawa, M., Sogabe, A. & Kitamoto, D.
384 (2009). Production of a novel glycolipid biosurfactant, mannosylmannitol lipid, by *Pseudozyma*
385 *parantarctica* and its interfacial properties. *Applied microbiology and biotechnology* **83**, 1017-1025.

386 Moure, A., Gullón, P., Domínguez, H. & Parajo, J. C. (2006). Advances in the manufacture, purification and
387 applications of xylo-oligosaccharides as food additives and nutraceuticals. *Process Biochemistry* **41**,
388 1913-1923.

389 Oliveira, J. V., Dos Santos, R. A., Borges, T. A., Riano-Pachon, D. M. & Goldman, G. H. (2013). Draft Genome
390 Sequence of *Pseudozyma brasiliensis* sp. nov. Strain GHG001, a High Producer of Endo-1,4-Xylanase
391 Isolated from an Insect Pest of Sugarcane. *Genome Announc* **1**, e00920-00913.

392 Polizeli, M. L., Rizzatti, A. C., Monti, R., Terenzi, H. F., Jorge, J. A. & Amorim, D. S. (2005). Xylanases from
393 fungi: properties and industrial applications. *Applied microbiology and biotechnology* **67**, 577-591.

394 Rice, P., Longden, I. & Bleasby, A. (2000). EMBL: the European Molecular Biology Open Software Suite.
395 *Trends in genetics : TIG* **16**, 276-277.

396 Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*, 3rd edition edn. New York:
397 Cold Spring Harbor Laboratory Press.

398 Seo, H. S., Um, H. J., Min, J., Rhee, S. K., Cho, T. J., Kim, Y. H. & Lee, J. (2007). *Pseudozyma jejuensis* sp. nov.,
399 a novel cutinolytic ustilaginomycetous yeast species that is able to degrade plastic waste. *FEMS*
400 *yeast research* **7**, 1035-1045.

401 Smith, D. G., Garcia-Pedrajas, M. D., Hong, W., Yu, Z., Gold, S. E. & Perlin, M. H. (2004). An ste20
402 homologue in *Ustilago maydis* plays a role in mating and pathogenicity. *Eukaryotic cell* **3**, 180-189.

403 Smith, T. F. & Waterman, M. S. (1981). Identification of common molecular subsequences. *Journal of*
404 *molecular biology* **147**, 195-197.

405 Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of
 406 taxa and mixed models. *Bioinformatics* 22, 2688-2690.

407 Statzell-Tallman, A., Scorzetti, G. & Fell, J. W. (2010). *Candida spencermartinsiae* sp. nov., *Candida taylori*
 408 sp. nov. and *Pseudozyma abaconensis* sp. nov., novel yeasts from mangrove and coral reef
 409 ecosystems. *International journal of systematic and evolutionary microbiology* 60, 1978-1984.

410 Sugita, T., Takashima, M., Poonwan, N., Mekha, N., Malaithao, K., Thungmuthasawat, B., Prasarn, S.,
 411 Luangsook, P. & Kudo, T. (2003). The first isolation of ustilaginomycetous anamorphic yeasts,
 412 *Pseudozyma* species, from patients' blood and a description of two new species: *P. parantarctica*
 413 and *P. thailandica*. *Microbiology and immunology* 47, 183-190.

414 Tabei, Y., Kiryu, H., Kin, T. & Asai, K. (2008). A fast structural multiple alignment method for long RNA
 415 sequences. *BMC Bioinformatics* 9, 33.

416 Urban, M., Kahmann, R. & Bolker, M. (1996). The biallelic a mating type locus of *Ustilago maydis*: remnants
 417 of an additional pheromone gene indicate evolution from a multiallelic ancestor. *Molecular &*
 418 *general genetics : MGG* 250, 414-420.

419 Wang, Q. M., Jia, J. H. & Bai, F. Y. (2006). *Pseudozyma hubeiensis* sp. nov. and *Pseudozyma shanxiensis* sp.
 420 nov., novel ustilaginomycetous anamorphic yeast species from plant leaves. *International journal of*
 421 *systematic and evolutionary microbiology* 56, 289-293.

422 Xu, J., Saunders, C. W., Hu, P., Grant, R. A., Boekhout, T., Kuramae, E. E., Kronstad, J. W., Deangelis, Y. M.,
 423 Reeder, N. L. & other authors (2007). Dandruff-associated *Malassezia* genomes reveal convergent
 424 and divergent virulence traits shared with plant and human fungal pathogens. *Proc Natl Acad Sci U S*
 425 *A* 104, 18730-18735.

426 Yang, Z. (1996). Among-site rate variation and its impact on phylogenetic analyses. *Trends in ecology &*
 427 *evolution* 11, 367-372.

428

429 **Figure Legends:**

430 Fig. 1. Maximum-likelihood phylogenetic tree of *Pseudozyma brasiliensis* sp. nov. and related species, based
 431 on the analysis of the ITS1 and ITS2 regions, the 5.8S rDNA and the D1/D2 region of the LSU. Bootstrap
 432 values are based on 1000 replicates, only values above 50% are shown. GenBank accession numbers
 433 appear in parentheses. *Cintractia axicola* and *Trichocintractia utriculicola* were used as outgroups. ^T:
 434 Sequences from type strains.

435 Fig. 2. Rudimentary pseudohyphae produced in minimal medium containing 0.5% ammonium sulphate after .3
 436 days at 30 °C. Bar 10µm.

437 Fig. 3. Venn diagrams showing number of shared groups of orthologues, i.e., putative gene families, between
 438 *P. brasiliensis* sp. nov. and related species. (a) Showing only the sequenced *Pseudozyma* species. (b). *P.*
 439 *brasiliensis* and most closely related *Pseudozyma* species (from Fig. 1), in addition to *U. maydis* and *M.*
 440 *globosa*. Groups of Orthologous genes were identified by OrthoMCL using an inflation value of 2.5.

441 Fig. 4. *Pseudozyma brasiliensis* sp. nov. GHG001^T: cells showing possible oil droplets inside, after 7 days at
 442 25 °C grown on potato dextrose agar (a); cells showing polar budding on short denticles, grown on corn meal
 443 agar after 21 days at 25 °C. Bar 10 µm.

444

445

446 **Supplementary material:**

447 Supp. Mat. 1. Taxonomy key for the identification of species in the genus *Pseudozyma*.

448

449 Table 1 - Comparison of phenotypic properties of *Pseudozyma brasiliensis* sp. nov. with
450 other recognized species of the genus *Pseudozyma*.

Characteristic	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Assimilation of:																					
Inulin	+	-	-	-	-	-	-	n	d/w	-	-	-	-	-	-	w	-	-	d/w	-	-
Melibiose	-	-	v	s	-	+	+	+	d	s	+	+	-	+	-	+	s	+	+	+	-
D-Galactose	+	-	-	+	+	+	+	n	+	s	s	+	s	+	-	+	+	+	d/w	+	s
Lactose	-	-	-	-	+	+	+	+	-	-	s	+	+	+	+	+	-	+	+	-	+
Trehalose	+	+	+	+	+	+	+	n	-	+	s	+	+	+	+	+	+	+	+	+	+
Maltose	+	+	+	w	+	+	+	n	+	+	+	+	+	+	+	+	+	+	+	+	+
Melezitose	+	+	+	s	+	+	+	n	+	+	+	+	+	+	+	+	+	+	+	+	+
Soluble starch	-	+	-	-	+	+	+	+	+	s	s	d/w	-	+	+	+	+	d/w	+	+	+
Cellulose	+	w	s	-	+	+/d	s	n	d	s	w	+	d	+	+	+	s	-	d	+	+
Salicin	+	w	v	-	+	+/w	d	n	+	+	w	-	d	+	-	+	d	-	+	+	-
D-Sorbitose	s	s	d	w	+	+	d	n	d/w	d	w	-	d	w	d	w	+	-	+	+	d
D-Rhamnose	-	-	-	-	v	-	+	n	-	-	-	-	-	+	+	w	+	-	-	-	-
D-Xylose	+	+	+	s	+	+	+	n	+	+	+	+	+	+	+	+	+	+	+	+	+
D-Arabinose	+	-	+	s	+	+	+	n	+	+	+	+	+	+	+	+	+	+	+	+	+
D-Ribose	-	-	+	s	+	+/d	+	n	+	+	-	+	-	+	+	s	+	v	d	-	s
Methanol	-	-	-	-	-	-	-	n	+	-	-	-	-	-	-	-	w	-	-	-	-
Ethanol	+	-	+	w	+	+	+	d	-	+	-	-	-	+	+	+	+	-	+	+	+
Glycerol	+	-	s	-	+	+	s	n	+	-	s	+	-	+	+	+	+	+	+	+	+
Erythritol	+	-	s	s	+	+	+	-	+	+	w	+	d	d	+	+	s	+	+	+	+
Ribitol	+	-	s	s	s	+	s	n	d	s	w	+	-	+	s	v	s	+	d	d	-
D-Mannitol	+	s	+	w	+	+	+	n	+	+	s	+	+	+	+	+	+	+	+	+	-
D-Glucitol	+	+	+	w	+	+	+	n	+	+	s	+	+	+	+	+	+	+	+	+	-
myo-Inositol	s	d	+	+	+	+	+	n	+	+	w	-	-	+	+	+	+	-	+	+	+
DL-Lactic acid	-	+	s	s	+	+	+	n	w	s	s	-	-	+	s	w	+	-	+	+	+
Succinic acid	+	+	+	+	+	+	+	-	+	+	s	+	-	+	+	+	+	+	+	+	+
Citric acid	-	+	s	+	+	+	+	n	+	+	s	-	+	+	-	w	+	-	+	+	+
D-Gluconic acid	+	-	+	+	+	+	s	n	d	s	s	n	d	+	+	w	+	n	+	w	-
Xylitol	d/w	+	n	n	n	n	n	n	n	n	-	n	-	d	n	s	n	n	-	n	n
L-Lysine	+	+	+	n	+	n	+	n	+	+	+	+	+	+	+	+	+	+	n	+	-
Growth:																					
Vitamin-free medium	+	+	+	+	-	-	-	n	+	+	+	+	+	-	+	+	w	+	+	-	+
50% glucose	-	+	-	n	-	-	-	n	-	-	-	-	-	-	-	-	-	-	-	-	-
10% sodium chloride	-	w	n	-	n	n	n	n	n	n	w	n	n	n	n	n	n	n	n	n	n
0.01% cycloheximide	-	-	n	n	n	n	n	n	n	n	n	n	n	n	n	+	n	n	n	n	n
37 °C	-	-	-	-	v	+	+	-	+	-	n	+	+	d/w	-	+	+	+	+	+	-

451 1, *Pseudozyma brasiliensis* sp. nov.; 2, *P. vetiver*; 3, *P. fusiformata*; 4, *P. abaconensis*; 5, *P. antarctica*; 6, *P.*
452 *alboarmeniaca*; 7, *P. aphidis*; 8, *P. churashimaensis*; 9, *P. crassa*; 10, *P. flocculosa*; 11, *P. graminicola*; 12, *P. hubeiensis*;
453 13, *P. jejuensis*; 14, *P. parantarctica*; 15, *P. prolifica*; 16, *P. pruni*; 17, *P. rugulosa*; 18, *P. shanxiensis*; 19, *P. siamensis*;
454 20, *P. thailandica*; 21, *P. tsukubaensis*. Abbreviations: +, positive; -, negative; w, weak; v, variable; d, delayed growth; s,
455 slow growth; n, no data.

456

Table 2 – Identification of Mating-type related genes in *Pseudozyma* spp. a. P: Percent identity. C. Coverage query protein (%). Species abbreviations: Bra: *P. brasiliensis* sp. nov.; Hub: *P. hubeiensis*; Aphi: *P. aphidis*; Flo: *P. flocculosa*; Ant: *P. antarctica*. *: Protein in *U. maydis*. †: Comparison against the corresponding protein in *U. maydis*, resulting from a pairwise local alignment. M: Mis- or non-annotated gene, the coverage column in these cases is expressed in relation to the *U. maydis* protein.

Ref Gene Name	Ref protein (length)	Species	Genomic Position	Protein accession number	p [†]	C
Pheromone receptor (pra)	XM_753437 (357 aa)	Bra	PSEUBRA_SCAF4 (KJ545884.1): 730180-731367	EST05906.1	52	100
		Hub	DF238821.1: 335177- 335870	GAC98648.1	37	58
		Aphi	AWNIO1000010.1: 416482 -417484	ETS62568.1	62	98
		Flo	scaffold00004(KE361626.1): 662653- 662730	EPQ30811.1	45	90
		Ant	scaffold00012(DF196778.1): 428476 - 429478	GAC74678.1	43	97
Mating pheromone (mla)	CCF49902 (42 aa)	Bra	PSEUBRA_SCAF4(KJ545884.1): 733620-733727	M	62	86
		Hub	DF238821.1: 333635-333748	M	55	93
		Aphi	AWNIO1000010.1: 413919-414035	ETS62567.1	54	26
		Flo	scaffold00004(KE361626.1): 653355-653726	EPQ30812.1	35	93
		Ant	scaffold00012 (DF196778.1): 425891- 426007	M	55	93
rba1	XP_758531 (119 aa)	Bra	PSEUBRA_SCAF4(KJ545884.1): 729004-729273	EST05905.1	79	97
		Hub	DF238821.1: 339486-339928	M	80	98
		Aphi	AWNIO1000010.1: 403403-403672	ETS62563.1	82	97
		Flo	scaffold00004(KE361626.1): 653355-653726	EPQ30806.1	63	94
		Ant	scaffold00012(DF196778.1): 415664-415933	GAC74673.1	82	97
Pheromone response factor (prf)	XP_758860 (840 aa)	Bra	PSEUBRA_SCAF5(KJ545891.1): 277324 - 279591	EST05476.1	51	100
		Hub	DF238776.1(275713-278307)	GAC93481.1	62	100
		Aphi	AWNIO1000004.1: 286205-288613	ETS64722.1	42	98
		Flo	scaffold00012(KE361634.1): 380670 - 381247	EPQ28567.1	27	75
		Ant	scaffold00005(DF196771.1): 282068-284503	GAC71737.1	42	98
Ste-20 like protein kinase (smu1)	AAM97788 (746 aa)	Bra	PSEUBRA_SCAF14(KJ545856.1): 220258-222378	EST08823.1	84	100
		Hub	DF238786.1: 153912-156029	GAC94823.1	84	95
		Aphi	AWNIO1000038.1: 245809- 247920	ETS60212.1	77	91

		Flo	scaffold00023(KE361645.1): 258154- 260889	EPQ26394.1	64	94
		Ant	scaffold00027(DF196793.1): 245527-247686	GAC77637.1	81	100
bW	XP_756725 (626 aa)	Bra	PSEUBRA_SCAF3(KI545873.1): 1480206 - 1482317	EST06506.1	40	99
		Hub	DF238831.1: 125908-127963	GAC99319.1	47	92
		Aphi	AWN101000008.1: 1628892-1630348	ETS63816.1	35	96
		Flo	scaffold00010(KE361632.1): 86085-86246	EPQ28975.1	28	93
		Ant	scaffold00009(DF196775.1): 1634549- 1636121	GAC73812.1	36	92
bE	CAJ59728 (462 aa)	Bra	PSEUBRA_SCAF3(KI545873.1): 1482868 - 1484199	M	38	95
		Hub	DF238831.1: 124148-125587	GAC99318.1	47	71
		Aphi	AWN101000008.1: 1630687-1632102	ETS63817.1	50	100
		Flo	scaffold00010(KE361632.1): 84397-85371	GAC73813.1	25	94
		Ant	scaffold00009(DF196775.1): 1636427 - 1637572	EPQ28974.1	49	100

464

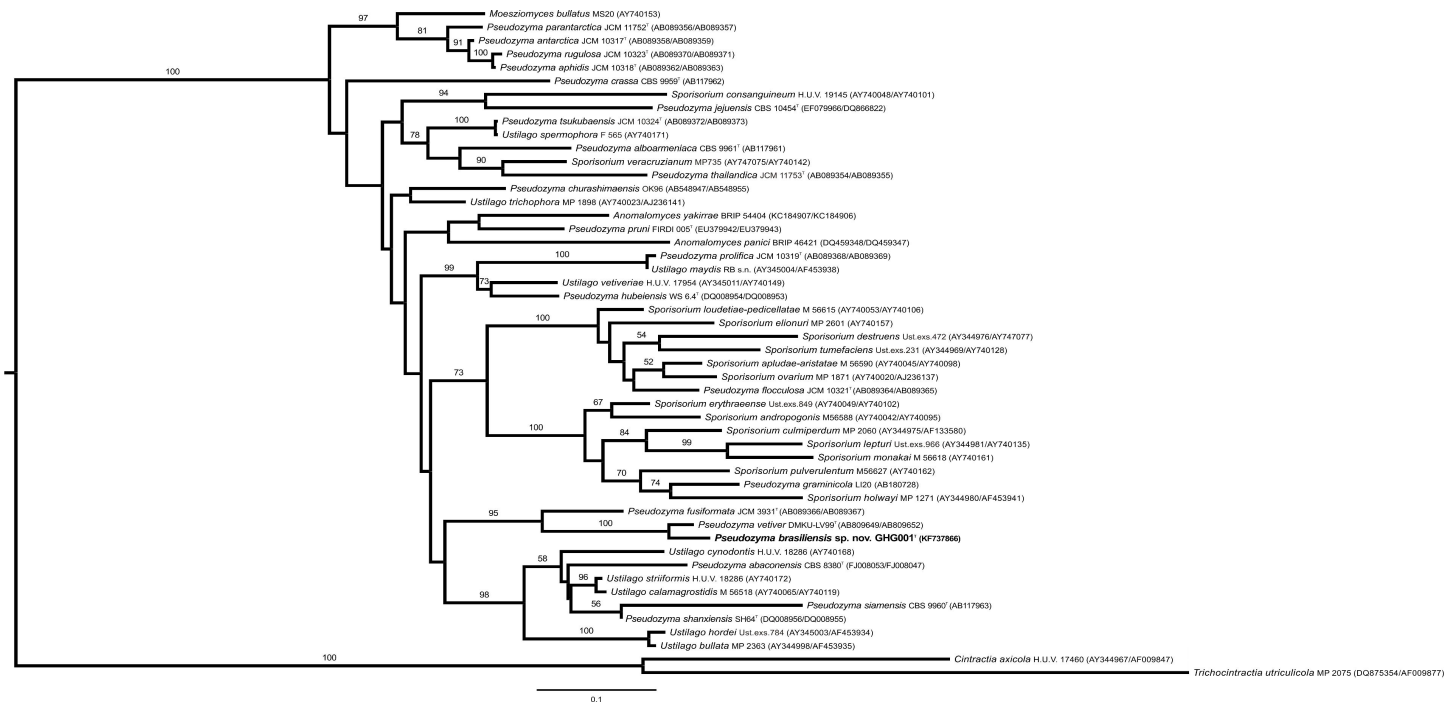
465 Table 3 - Comparison of genome properties of *Pseudozyma brasiliensis* sp. nov. with other
 466 sequenced related species.

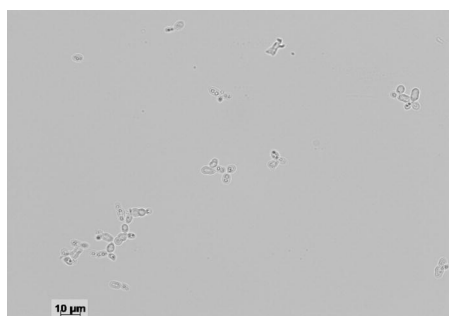
467

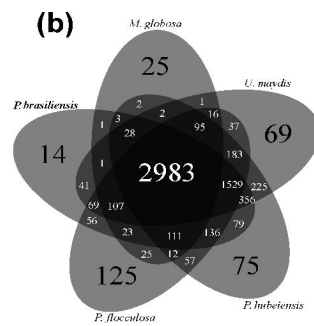
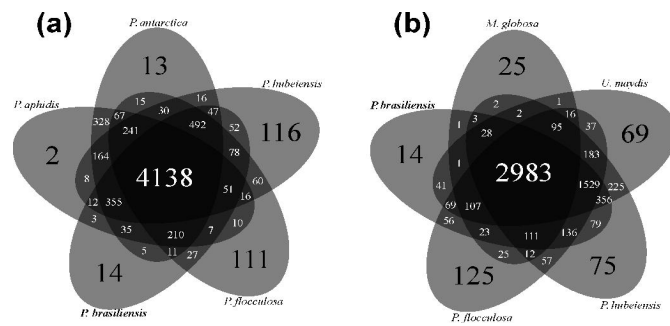
Characteristic	<i>P. brasiliensis</i>	<i>P. antarctica</i>	<i>P. aphidis</i>	<i>P. flocculosa</i>	<i>P. hubeiensis</i>	<i>U. maydis</i>	<i>M. globosa</i>
Assembly statistics							
Total scaffold length (Mbp)	17.33	18.07	17.92	23.31	18.44	19.74	8.96
N50 scaffolds (Mbp)	2.21	2.38	2.38	1.98	1.40	0.51	1.41
N90 scaffolds (Mbp)	1.72	2.07	2.07	1.48	1.09	0.48	1.01
GC content (%)							
Assembled	33.85	42.03	37.51	36.12	39.33	34.13	37.00
Protein coding	58.10	60.28	60.50	64.30	56.49	53.95	52.06
Number of protein coding genes	5768	6640	6011	6877	7472	6522	4286

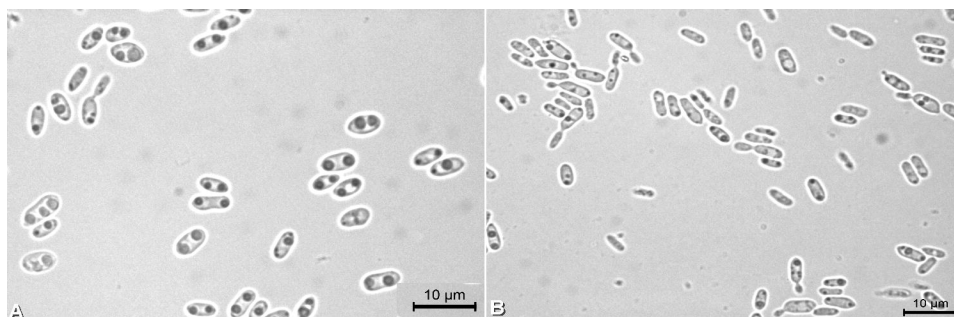
468

469









***Pseudozyma brasiliensis* sp. nov., a novel xylanolytic Ustilaginomycetous yeast species isolated from an insect pest of sugarcane roots**

Juliana Velasco de Castro Oliveira¹, Thuanny A. Borges¹, Renato Augusto Corrêa dos Santos¹, Larissa F. D. Freitas², Carlos Augusto Rosa², Gustavo Henrique Goldman^{1,3#}, and Diego Mauricio Riaño-Pachón^{1#}

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

² Departamento de Microbiologia, Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas, Belo Horizonte, Minas Gerais - Brazil

³ Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil

Corresponding authors: ggoldman@usp.br and diego.riano@bioetanol.org.br

**Taxonomic key for the species in the genus
Pseudozyma.**

1	a.	Galactitol is assimilated	<i>U. maydis</i>
	b.	Galactitol is not assimilated	2
2(1)	a.	Succinic acid is assimilated	3
	b.	Succinic acid is not assimilated	21
3(2)	a.	L-Arabinose is assimilated	4
	b.	L-Arabinose is not assimilated	<i>P. vetiver</i>
4(3)	a.	D-Ribose is assimilated	5
	b.	D-Ribose is not assimilated	<i>P. fusiformata</i>
5(4)	a.	Trehalose is assimilated	6

	b.	Trehalose is not assimilated	<i>P. crassa</i>
6(5)	a.	D-Galactose is assimilated	7
	b.	D-Galactose is not assimilated	<i>P. prolifica</i>
7(6)	a.	Melibiose is assimilated	8
	b.	Melibiose is not assimilated	19
8(7)	a.	Soluble starch is assimilated	9
	b.	Soluble starch is not assimilated	<i>P. abaconensis</i>
9(8)	a.	Cellobiose is assimilated	10
	b.	Cellobiose is not assimilated	<i>P. shanxiensis</i>
10(9)	a.	Salicin is assimilated	11
	b.	Salicin is not assimilated	<i>P. hubeiensis</i>
11(10)	a.	Glycerol is assimilated	12
	b.	Glycerol is not assimilated	<i>P. flocculosa</i>
12(11)	a.	Ethanol is assimilated	13
	b.	Ethanol is not assimilated	<i>P. aphidis</i>
13(12)	a.	Methanol is assimilated	<i>P. rugulosa</i>
	b.	Methanol is not assimilated	14
14(13)	a.	Lactose is assimilated	15
	b.	Lactose is not assimilated	<i>P. thailandica</i>
15(14)	a.	Inulin is assimilated	16
	b.	Inulin is not assimilated	17
16(15)	a.	L-Rhamnose is assimilated	<i>P. pruni</i>
	b.	L-Rhamnose is not assimilated	<i>P. siamensis</i>
17(15)	a.	D-Arabinose is assimilated	18
	b.	D-Arabinose is not assimilated	<i>P. gramminicola</i>
18(17)	a.	L-Rhamnose is assimilated	<i>P. parantarctica</i>
	b.	L-Rhamnose is not assimilated	<i>P. alboarmeniaca</i>

19(7)	a.	Inulin is assimilated	<i>P. brasiliensis</i>
	b.	Inulin is not assimilated	20
20(19)	a.	Salicin is assimilated	<i>P. antarctica</i>
	b.	Salicin is not assimilated	<i>P. tsukubaensis</i>
21(2)	a.	Melibiose is assimilated	<i>P. churashimaensis</i>
	b.	Melibiose is not assimilated	<i>P. jejuensis</i>

7. Artigo: Draft Genome Sequence of *Pseudozyma brasiliensis* sp. nov. Strain GHG001, a High Producer of Endo-1,4-Xylanase Isolated from an Insect Pest of Sugarcane

Nesse item pode ser visualizado o artigo referente ao sequenciamento do genoma da *P. brasiliensis* sp. nov. GHG001. O genoma foi sequenciado devido ao alto potencial biotecnológico da GHG001 e por se tratar de uma levedura basidiomicética, com características morfológicas e gênicas peculiares, diferentes das demais leveduras. O genoma foi sequenciado pela plataforma Illumina HiSeq2000, gerando 73.703.370 *paired-end reads* de 100pb, com um tamanho de 22,09 Mpb.

Draft Genome Sequence of *Pseudozyma brasiliensis* sp. nov. Strain GHG001, a High Producer of Endo-1,4-Xylanase Isolated from an Insect Pest of Sugarcane

Juliana Velasco de Castro Oliveira,^a Renato Augusto Corrêa dos Santos,^a Thuanny A. Borges,^a Diego Mauricio Riaño-Pachón,^a Gustavo Henrique Goldman^{a,b}

Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Campinas, São Paulo, Brazil^a; Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil^b

J.V.D.C.O. and R.A.C.D.S. contributed equally to this work.

Here, we present the nuclear and mitochondrial genome sequences of *Pseudozyma brasiliensis* sp. nov. strain GHG001. *P. brasiliensis* sp. nov. is the closest relative of *Pseudozyma vetiver*. *P. brasiliensis* sp. nov. is capable of growing on xylose or xylan as a sole carbon source and has great biotechnological potential.

Received 2 October 2013 **Accepted** 18 November 2013 **Published** 19 December 2013

Citation Oliveira JVDC, dos Santos RAC, Borges TA, Riaño-Pachón DM, Goldman GH. 2013. Draft genome sequence of *Pseudozyma brasiliensis* sp. nov. strain GHG001, a high producer of endo-1,4-xylanase isolated from an insect pest of sugarcane. *Genome Announc.* 1(6):e00920-13. doi:10.1128/genomeA.00920-13.

Copyright © 2013 Oliveira et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Diego Mauricio Riaño-Pachón, diego.riano@bioetanol.org.br, or Gustavo Henrique Goldman, ggoldman@usp.br.

Pseudozyma brasiliensis sp. nov. strain GHG001 is a yeast-like species that belongs to the order *Ustilaginales*. This strain was isolated from the intestinal tract of a *Chrysomelidae* larva associated with sugarcane roots in plantations in Ribeirão Preto, São Paulo, Brazil, following an enrichment protocol for microorganisms that use xylose as a sole carbon source. Based on the phylogenetic analysis of the ribosomal operon, we suggest that GHG001 represents a novel species that we named *P. brasiliensis* sp. nov.; its closest relative is *Pseudozyma vetiver* (1). GHG001 can grow well in xylose or xylan as its sole carbon source, where it produces high levels of endo-1,4-xylanase from the glycoside hydrolase (GH) family GH11 (2), the members of which show higher specific activity than other eukaryotic xylanases. Xylanases are essential for breaking down hemicellulose of plant cell walls, and they are routinely added to enzyme cocktails for the saccharification of pretreated biomass and second-generation ethanol production. Xylanases have further commercial applications, such as in bread making, the manufacture of food, drinks, and textiles, bleaching of cellulose pulp, and xylitol production (3).

Here, we present the genome sequence of *P. brasiliensis* sp. nov. strain GHG001. This genome was sequenced on the Illumina HiSeq2000 system, generating 73,703,379 paired-end reads of 100 bp (insert size, 250 bp). The reads were preprocessed with the Fastx-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). The genome size was estimated to be 22.09 Mbp based on *k*-mer count statistics (4), with an estimated coverage of 585×. The reads were randomly subsampled to a genome coverage of approximately 100×, and this subset was assembled using VelvetOptimiser and Velvet (5,6). The remaining reads were used to extend the contigs and perform scaffolding using SSPACE Basic (7). The resulting assembly has 45 scaffolds, with a total length of 17,323,620 bp and an *N*₅₀ of 720,612 bp. The average G+C content of the genome is 56.3%, which is similar to those of *Pseudozyma hubeiensis* SY62 (8)

and *Pseudozyma antarctica* T-34 (9). We evaluated the completeness of the gene space using CEGMA (10), which revealed that the current assembly is 97.98% complete. The scaffolds were masked for repeats using RepeatMasker, and gene prediction was carried out with GeneMark (11), Augustus (12), and STAP (<http://korflab.ucdavis.edu/software.html>), using MAKER (13). Gene finders were trained with the CEGMA-produced gene models. A total of 5,768 protein-encoding genes were identified, which is similar to the gene content of other *Pseudozyma* spp. A search against the NCBI nr database revealed 2,361 protein-encoding genes with strong sequence similarity hits to proteins in that database, providing a preliminary landscape of the genomic and metabolic capabilities of *P. brasiliensis*. Ribosomal genes were identified with RNAmmer (14), and the rRNA operon repeats (small subunit [SSU], internal transcribed spacer 1 [ITS1], 5.8 S, ITS2, and long subunit [LSU]) were collapsed into a single scaffold (PSEUBRA_SCAF27). One hundred nineteen tRNA genes were identified with tRNAscan-SE version 1.3.1 (15). The scaffold PSEUBRA_SCAF26 contains the mitochondrial genome.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. **AWXO00000000**. The version described in this paper is version AWXO01000000.

ACKNOWLEDGMENTS

R.A.C.D.S. and T.A.B. hold FAPESP scholarships (no. 2011/22690-3 and 2012/00080-1). This work was financially supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo (no. FAPESP 10/513224-2), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Vale SA, Brazil.

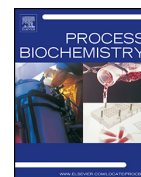
REFERENCES

1. Chamnanpa T, Limtong P, Srisuk N, Limtong S. 2013. *Pseudozyma vetiver* sp. nov., a novel anamorphic ustilaginomycetous yeast species iso-

- lated from the phylloplane in Thailand. *Antonie Van Leeuwenhoek* 104: 637–644.
2. Paës G, Berrin JG, Beaugrand J. 2012. GH11 xylanases: structure/function/properties relationships and applications. *Biotechnol. Adv.* 30: 564–592.
 3. Polizeli ML, Rizzatti AC, Monti R, Terenzi HF, Jorge JA, Amorim DS. 2005. Xylanases from fungi: properties and industrial applications. *Appl. Microbiol. Biotechnol.* 67:577–591.
 4. Marçais G, Kingsford C. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of *k*-mers. *Bioinformatics* 27:764–770.
 5. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 18:821–829.
 6. Zerbino DR. 2010. Using the Velvet *de novo* assembler for short-read sequencing technologies. *Curr. Protoc. Bioinformatics* 31:11.5.1–11.5.12. doi:10.1002/0471250953.bi1105s31.
 7. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27: 578–579.
 8. Konishi M, Hatada Y, Horiuchi J. 2013. Draft genome sequence of the basidiomycetous yeast-like fungus *Pseudozyma hubeiensis* SY62, which produces an abundant amount of the biosurfactant mannosylerythritol lipids. *Genome Announc.* 1(4):e00409-13. doi:10.1128/genomeA.00409-13.
 9. Morita T, Koike H, Koyama Y, Hagiwara H, Ito E, Fukuoka T, Imura T, Machida M, Kitamoto D. 2013. Genome sequence of the basidiomycetous yeast *Pseudozyma antarctica* T-34, a producer of the glycolipid biosurfactants mannosylerythritol lipids. *Genome Announc.* 1(2):e00064-13. doi:10.1128/genomeA.00064-13.
 10. Parra G, Bradnam K, Korf I. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* 23:1061–1067.
 11. Borodovsky M, Lomsadze A, Ivanov N, Mills R. 2003. Eukaryotic gene prediction using GeneMark.hmm. *Curr. Protoc. Bioinformatics* 35: 4.6.1–4.6.10. doi:10.1002/0471250953.bi0406s35.
 12. Stanke M, Schöffmann O, Morgenstern B, Waack S. 2006. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics* 7:62. doi:10.1186/1471-2105-7-62.
 13. Cantarel BL, Korf I, Robb SM, Parra G, Ross E, Moore B, Holt C, Sánchez Alvarado A, Yandell M. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 18:188–196.
 14. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108.
 15. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25: 955–964.

8. Artigo: Biochemical characterization of an endoxylanase from *Pseudozyma brasiliensis* sp. nov. strain GHG001 isolated from the intestinal tract of *Chrysomelidae* larvae associated to sugarcane roots

A *P. brasiliensis* sp.nov. GHG001 possui a importante propriedade de produzir xilanase, propriedade muito comum em vários fungos filamentosos, porém não vista frequentemente em leveduras (ADSUL; BASTAWDE; GOKHALE, 2009). As xilanases são glicosil-hidrolases que degradam xilano, o polissacarídeo mais abundante na hemicelulose proveniente da cana-de-açúcar, e possuem muitos usos comerciais, tais como na fabricação de papel, alimentação animal, panificação, indústrias de suco e vinho, produção de xilitol, além de ser utilizada na hidrólise do material lignocelulósico para produção do etanol de segunda geração (POLIZELI *et al.*, 2005). Nesse artigo foi caracterizada bioquimicamente uma xilanase (PbXynA) produzida pela *P. brasiliensis* sp. nov. GHG001 a qual se mostrou com alto potencial biotecnológico e industrial. A xilanase produzida pela *P. brasiliensis* GHG001 foi purificada por Gel Filtração apresentando 24 kDa, pertencente a família GH11 e constituída de β -folhas. A caracterização da xilanase resultou em pH e temperatura ótimos, 4 e 55°C, respectivamente. A enzima possui influência de alguns íons como Ca^{2+} . A cinética indicou um comportamento sigmoidal característico de enzimas alostéricas e de acordo com a eletroforese capilar, a xilanase produz xilooligossacarídeos que podem ser usados industrialmente como prébióticos. O artigo pode ser visualizado abaixo.



Biochemical characterization of an endoxylanase from *Pseudozyma brasiliensis* sp. nov. strain GHG001 isolated from the intestinal tract of *Chrysomelidae* larvae associated to sugarcane roots

Thuanny A. Borges^a, Aline Tieppo de Souza^a, Fabio M. Squina^a,
Diego Mauricio Riaño-Pachón^a, Renato Augusto Corrêa dos Santos^a, Ednildo Machado^b,
Juliana Velasco de Castro Oliveira^a, André R.L. Damásio^a, Gustavo H. Goldman^{a,c,*}

^a Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais, CNPEM, Brazil

^b Instituto de Biofísica Carlos Chagas, Universidade Federal do Rio de Janeiro, Brazil

^c Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Brazil

ARTICLE INFO

Article history:

Received 3 September 2013

Received in revised form 8 October 2013

Accepted 9 October 2013

Available online 18 October 2013

Keywords:

Pseudozyma brasiliensis

Endoxylanase

Biochemical characterization

GH11

ABSTRACT

Endo-xylanases play a key role in the hydrolysis of xylan and recently they have attracted much attention due to their potential applications on the biofuel and paper industries. We isolated a *Pseudozyma brasiliensis* sp. nov. strain from the intestinal tract of *Chrysomelidae* larvae that parasitize sugarcane roots. This basidiomycetous yeast produces a xylanase designated PbXynA which was purified and characterized. The molecular weight of PbXynA is 24 kDa, it belongs to the GH11 family and its optimum pH and optimum temperature are 4.0 and 55 °C, respectively. PbXynA has as secondary structure predominantly β -sheets and sigmoidal kinetic behavior with elevated speed conversion from substrate-to-products ($V_{\max} = 2792.0 \mu\text{mol product/min/mg protein}$). It is highly activated by bivalent cations such as Ca^{2+} , however in the presence of Cu^{2+} xylanase activity was inhibited. It has a high specific activity and produces xylooligosaccharides that have a variety of industrial applications, indicating PbXynA has a great biotechnological potential.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Lignocellulosic materials account for approximately 50% of the biomass in the world [1]. Degradation and conversion of lignocellulosic biomass are attracting attention because of its potential for the development of a sustainable and environmentally friendly bioenergy, biorefining and biomaterials industry [2]. Hemicellulose is the second most abundant renewable biomaterial available after cellulose representing 20–35% in lignocellulosic polymers [1]. Hemicelluloses consist of heteropolymers such as xylan, glucomannan, galactoglucomannan and arabinogalactan [3]. Xylan is a linear polymer of β -D-xylopyranosyl units linked by (1–4) glycosidic bonds and is the major hemicellulose. The complete hydrolysis of this heteropolysaccharide requires a set of enzymes, which the endo-1,4- β -xylanase is the most abundant [4].

Xylanases, the enzymes responsible for xylan backbone breakdown, have been classified into different GH (glycoside hydrolase) families: 5, 8, 10, 11, 16, 26, 30, 43, and 62 [1], but the major group can be classified into families 10 and 11 based on their amino-acid sequence similarities [5]. GH11 family is one of the best characterized GH families and it displays several interesting properties like high substrate selectivity and high catalytic efficiency, small size (around 20 kDa), variety of optimum pH and optimum temperature, making them suitable for various conditions in many applications [1]. Xylanases have played an important role in many industrial processes and have been applied as additives to enhance the quality of baked goods, animal feeds, and bleachers to kraft pulp [6]. Recently, they have received much attention owing to their use in degradation of lignocellulosic biomass for biofuel production [5]. The xylanolytic systems are extensively studied in filamentous fungi (*Aspergillus* sp., *Trichoderma* sp., *Penicillium* sp., etc.) and bacteria (*Bacillus* sp., *Streptomyces* sp.) but there are very few yeast examples [7].

Pseudozyma brasiliensis sp. nov. strain GHG001 is a basidiomycetous yeast, closely related to *Pseudozyma fusiformata* and *Pseudozyma vetiver*. *Pseudozyma* spp. are able to produce xylanases [7] and biosurfactants such as mannosylerythritol lipids [8–13]. We are currently looking for yeast strains that colonize the intestinal

* Corresponding author at: Laboratório Nacional de Ciência e Tecnologia do Bioetanol – CTBE/CNPEM, Giuseppe Máximo Scolaro Street, 10.000 Polo II de Alta Tecnologia, Campinas, São Paulo, Brazil. Tel.: +55 19 3512 1010; fax: +55 19 3518 3104.

E-mail addresses: andre.damasio@gmail.com (A.R.L. Damásio), ggoldman@usp.br (G.H. Goldman).

tract of insects that colonize sugarcane roots and evaluating them for the more efficient use of either single pentoses or their polysaccharides, such as xylose or xylan. We have isolated from the intestinal tract of a *Chrysomelidae* a strain that can grow very well in xylose as sole carbon source and produces high levels of xylanase when grown in xylan. Thus, in our studies we purified and characterized a xylanase (PbXynA) produced by this *P. brasiliensis* sp. nov. strain. Although there are hundreds of studies describing the characterization of eukaryotic xylanases, in this work we reported a G11 endo-1,4-xylanase showing a higher specific activity than the eukaryotic xylanases previously studied. Moreover, the PbXynA showed important biochemical characteristics such as a sigmoidal kinetic behavior, high activation by bivalent ions and application in xylooligosaccharides production, showing a strong biotechnological potential and commercial value.

2. Materials and methods

2.1. Isolation of *P. brasiliensis* sp. nov. strain GHG001

Chrysomelidae at last larval stages were collected from *Saccharum officinarum* roots in sugarcane plantations located in Ribeirão Preto, São Paulo State, Brazil. The larvae were anesthetized at 4 °C for 10 min, then externally sterilized with 70% ethanol for about 1 min, and rinsed twice with sterilized water. The guts of different larvae were dissected and the intestinal juices were transferred to a sterile flask containing 1 ml of cold sterile water. A small aliquot of each gut sample (50 µl) was then inoculated into 250 ml Erlenmeyer flask with 50 ml of minimal medium (MM) composed of 50 ml/l salt solution (6 g/l NaNO₃, 1.5 g/l KH₂PO₄, 0.5 g/l KCl and 0.5 g/l MgSO₄) and 1 ml/l trace elements (10 g/l EDTA, 4.4 g/l ZnSO₄·H₂O, 1.0 g/l MnCl₂·4H₂O, 0.32 g/l CoCl₂·6H₂O, 0.315 g/l CuSO₄·5H₂O, 0.22 g/l (NH₄)₆Mo₇O₂₄·4H₂O, 1.47 g/l CaCl₂·2H₂O and 1 g/l FeSO₄·H₂O) pH 6.5 with 1% xylose, supplemented with a cocktail of anti-bacterial antibiotics (100 µg/µl ampicillin, 34 µg/µl kanamycin, 20 µg/µl tetracycline and 35 µg/µl chloramphenicol). The inoculum was incubated at 30 °C under agitation at 130 rpm for enrichment of microorganisms that could use xylose more efficiently. Two weekly transfers were performed by changing the medium and after the second transfer, 100 µl of tenfold dilution were spread on solid MM with 1% xylose, and single colonies were repeatedly streaked out until pure cultures were obtained. These isolates were inoculated in 125 ml Erlenmeyer flasks with 25 ml of liquid YPD (20 g/l peptone, 10 g/l yeast extract and 20 g/l dextrose) and incubated under shaking at 30 °C for 2 days. Among the isolates strain GHG001 was selected and the extraction of genomic DNA followed the protocol proposed by Sambrook and Russell [14], and genome sequencing was carried out on an Illumina HiSeq2000 instrument (unpublished data).

2.2. Xylan/xylose/glucose induction and determination of enzymatic activities

P. brasiliensis were inoculated (1 × 10⁶ cells/ml) in 0.67% YNB medium and 2% xylan, 2% xylose or 1% glucose as carbon sources, at 30 °C for 12, 24, 48, 72, 96 and 120 h. An aliquot was used for spectrophotometric determination of optical density at 590 nm, the supernatant was harvested by centrifugation and was kept at –20 °C for enzymatic analysis. The xylanase enzymatic assay was performed using Azo-Xylan (Birchwood; from Megazyme International, Bray, Ireland) as substrate, according to the manufacturer's protocols. Briefly, supernatant containing enzymes from xylan, xylose or glucose induced *P. brasiliensis* sp. nov. was mixed with 100 mM sodium acetate buffer (pH 4.5) in an appropriate volume. Reaction mixtures consisted of 0.5 ml of buffered enzyme preparation and 0.5 ml of substrate solution (1% w/v Azo-Xylan (Birchwood)). The samples were incubated at 40 °C for 10 min, and the reactions were interrupted by adding 2.5 ml of ethanol (95%, v/v) with vigorous stirring. Non-hydrolyzed ethanol-precipitated substrate was removed by centrifugation at 1000 × g for 10 min, and the absorbance of the supernatant was measured at 590 nm. Enzymatic activity was determined using Mega-CalTM software (Megazyme International). One unit of enzymatic activity was defined as the amount of enzyme required to release 1 mM D-xylose-reducing sugar equivalent per minute from arabinoxylan (pH 4.5) at 40 °C.

2.3. Supernatant activity assay

Supernatant activity was assayed using xylan from beechwood, rye arabinoxylan, β-mannan, β-glucan, laminarin, pectin, arabinogalactan, arabinan from sugar beet, xyloglucan from tamarind and carboxymethylcellulose (CMC), all at a 5% final concentration. The released reducing sugars were quantified by a dinitrosalicylic method [15]. One unit (U) of enzyme activity was defined as the amount of enzyme required to release 1 µmol of reducing sugar per minute. Specific activity was expressed as U/mg protein.

2.4. Xylanase production and purification

Xylanase production was performed in 1 l Erlenmeyer flasks with 200 ml of the production medium. The flasks were inoculated with 10⁶ yeasts/ml and incubated at 30 °C on a rotary shaker (250 rpm). The cell growth was harvested after 96 h by centrifugation (4000 rpm, 20 min) and the supernatant was used as a crude enzyme preparation. The supernatant was concentrated by precipitation with 65% ammonium sulfate saturation. The concentrated enzyme sample was loaded on Superdex 75 column pre-equilibrated with 50 mM sodium acetate buffer, pH 4.0, and the fractions (1.5 ml/tube) were collected. The fractions were analyzed for xylanase activity and by 12% SDS-PAGE followed by staining with 0.1% Coomassie Brilliant Blue R-250 (w/v) in methanol/acetic acid/water (v/v/v) (4:1:5). One distinct xylanase peak was identified which was designated as PbXynA.

2.5. Mass spectrometry

The PbXynA corresponding band was separated by SDS-PAGE, and after extensive washing was submitted to *in situ* trypsin gel digestion with 0.5 µg of modified trypsin (Promega Corporation, Madison, WI, USA). An aliquot of the resulting peptide mixture was separated by C18 (75 µm × 100 mm) RP-nanoUPLC (nanoAcquity, Waters) coupled with a Q-ToF Ultima mass spectrometer (Waters) with nano-electrospray source at a flow rate of 0.6 µl/min. The gradient was 2–90% (v/v) acetonitrile in 0.1% (v/v) formic acid over 45 min. The instrument was operated in the 'top three' mode, in which one MS spectrum is acquired followed by MS/MS of the top three most-intense peaks detected. The spectra were acquired using software MassLynx v.4.1 (Waters Corporation, MA, USA) and the raw data files were converted to a peak list format (mgf) by the software Mascot Distiller v.2.3.2.0, 2009 (Matrix Science Ltd.). The MS/MS profile was searched against predicted protein sequences using engine Mascot v.2.3 (Matrix Science Ltd.) with carbamidomethylation as fixed modification, oxidation of methionine as variable modification, one chymotrypsin missed cleavage and a tolerance of 0.1 Da for both precursor and fragment ions.

2.6. Enzyme characterization

Enzymatic activity was measured colorimetrically using xylan from beechwood as substrate and the reducing sugars determined according to Miller procedure [15]. The reaction mixture, consisting of 50 µl substrate (5%, w/v) in 50 mM sodium acetate buffer, pH 4.0 and 10 µl enzyme solution at 1724.53 U/ml, that was incubated at 55 °C in a thermocycler for 10 min. The reaction was stopped by adding 100 µl of DNS and immediately boiled for 5 min. The reducing sugars released as a result of enzyme activity were measured at 540 nm. One unit of enzymatic activity was defined as the amount of enzyme that produced 1 µmol of reducing sugars per minute. To determine the optimum pH and temperature profiles, the enzymatic reaction was carried out at different pH in McIlvaine's buffer system (pH 2.0–10.0) and various temperatures (25–80 °C).

The PbXynA kinetic constants were determined using xylan from beechwood (1.2–60 mg/ml) as substrate at pH 4.0 and 55 °C and the kinetic data were fitted using the SigraW software [16]. The protein content was measured at 280 nm by Nanodrop and calculated by molar extinction coefficient. The effect of metal ions (Zn²⁺, Mn²⁺, Fe²⁺, Na⁺, Mg²⁺, K⁺, Cu²⁺, Co²⁺ and Ca²⁺) on xylanase activity was analyzed adding 10 mM ions at final concentration to the assay and a xylanase-bivalent ions free solution as control (Chelex 100, from Bio-Rad).

2.7. CD spectroscopy and thermal denaturation

Far-UVCD spectra of PbXynA were measured between 195 and 250 nm in 50 mM sodium acetate buffer, pH 4.0 at 25 °C with a Jasco (Hachioji City, Tokyo, Japan) J-810 spectropolarimeter using 1 mm-path-length cuvette and a protein concentration of 0.16 mg/ml. For each measurement, a total of eight spectra were collected, averaged and corrected by subtraction of a buffer blank and ellipticity was reported as the mean residue molar ellipticity (θ; deg cm² dmol^{–1}). Thermal denaturation was carried out using a 1 mm path-length cuvette at the same protein concentration in the same buffer, and measuring the CD signal change at 224.8 nm over the temperature range 20–100 °C at 2 degree intervals with a settling time of 60 s. The transition temperatures were estimated from the inflection points of the first derivative of the CD signal.

2.8. Capillary zone electrophoresis of oligosaccharides

The oligosaccharides released by the enzyme action on the xylan from beechwood was derivatized with 8-aminopyrene-1,3,6-trisulfonic acid (APTS) by reductive amination as described [17]. Capillary zone electrophoresis (CZE) of oligosaccharides was performed on a P/ACE MQD instrument (Beckman Coulter) equipped with laser-induced fluorescence detection. A fused-silica capillary (TSP050375, Polymicro Technologies) of internal diameter of 50 µm and total length of 31 cm was used as separation column for oligosaccharides. Electrophoresis conditions were 15 kV/70–100 µA at a controlled temperature of 20 °C. Oligomers labeled with APTS were excited at 488 nm and emission was collected through a 520 nm band pass filter.

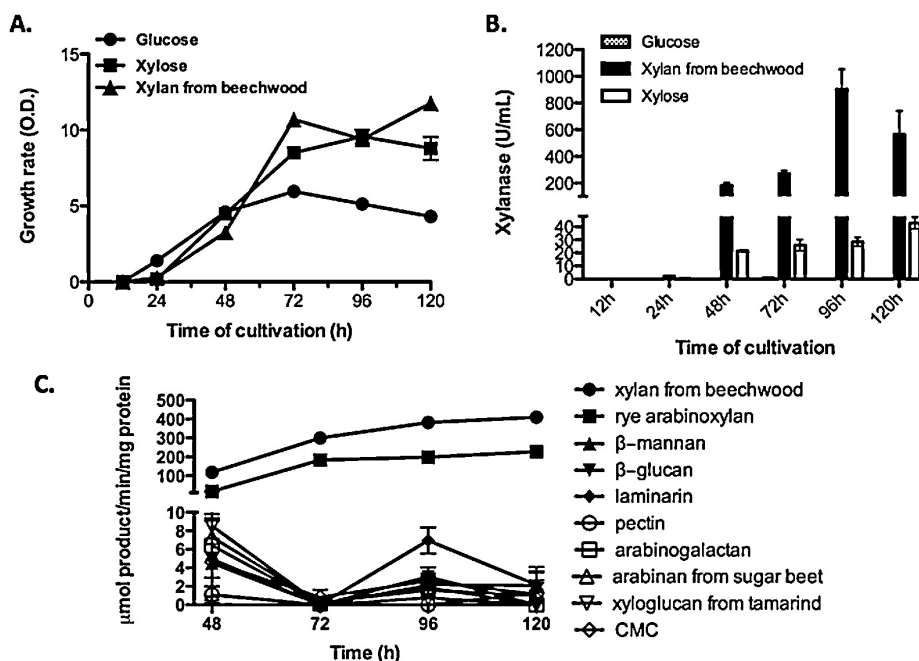


Fig. 1. *P. brasiliensis* sp. nov. has high growth rates in xylan and xylose. (A) Growth of *P. brasiliensis* in MM with either glucose (1%), xylose (2%) or xylan from beechwood (2%) at 30 °C. (B) Endoxylanase activities of the supernatant from cultures grown for 120 h at 30 °C in glucose (1%), xylose (2%) and xylan from beechwood (2%). (C) Endoxylanase activities of the supernatant from culture grown for 96 h at 30 °C in xylan from beechwood (2%). The results are the average of three independent assays, and the error bars represent standard deviations.

3. Results and discussion

3.1. *P. brasiliensis* is a major xylanase producer

We have isolated a yeast strain from the intestinal tract of *Chrysomelidae* larvae by enrichment of xylose as single carbon source. Sequence analysis of the 5.8S and Large ribosomal subunit as well as the internal transcribed spacer regions (ITS) revealed that this strain (GHG001) is a new species, closely related to *P. vetiver*,

which was recently isolated in Thailand associated to vetiver and sugarcane leaves [18]. Thus we propose this as a new species named *P. brasiliensis* sp. nov. (unpublished results). This strain was able to grow in xylose and xylan to higher levels than in glucose (Fig. 1A). Subsequently, we verified the induction of hemicellulases, by evaluating a time course of endo-1,4- β -xylanase activity when *P. brasiliensis* sp. nov. was grown in the presence of xylose, xylan and glucose as carbon sources. Xylanase activity was higher in xylan than in xylose, and it was absent in glucose as carbon

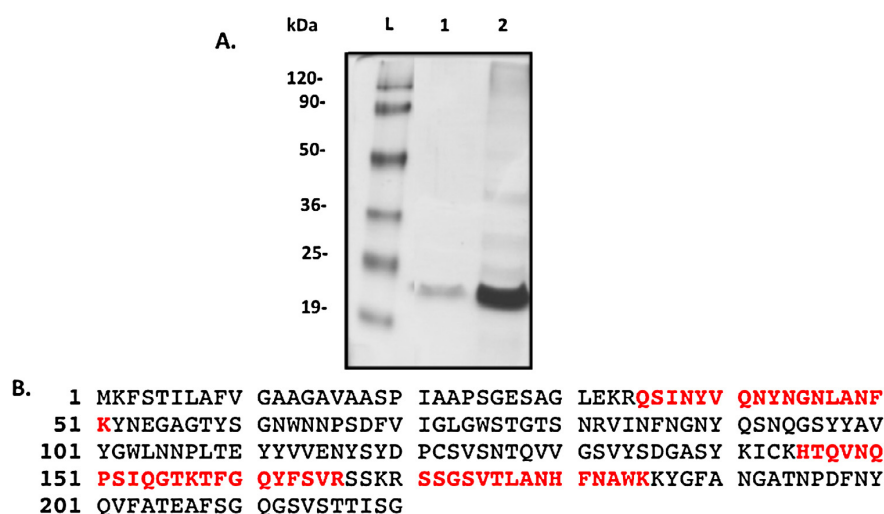


Fig. 2. Purification of PbXynA. (A) Supernatant and PbXynA after gel filtration chromatography were run on a 12% SDS-PAGE; (L) ladder, (1) PbXynA purified, (2) supernatant culture after 2% xylan from beechwood induction for 96 h at 30 °C. The PbXynA was identified by MALDI-TOF-TOF yielding 24% sequence coverage (peptides underlined). Based on BLAST analysis the PbXynA sequence showed a high similarity with GH11 enzymes.

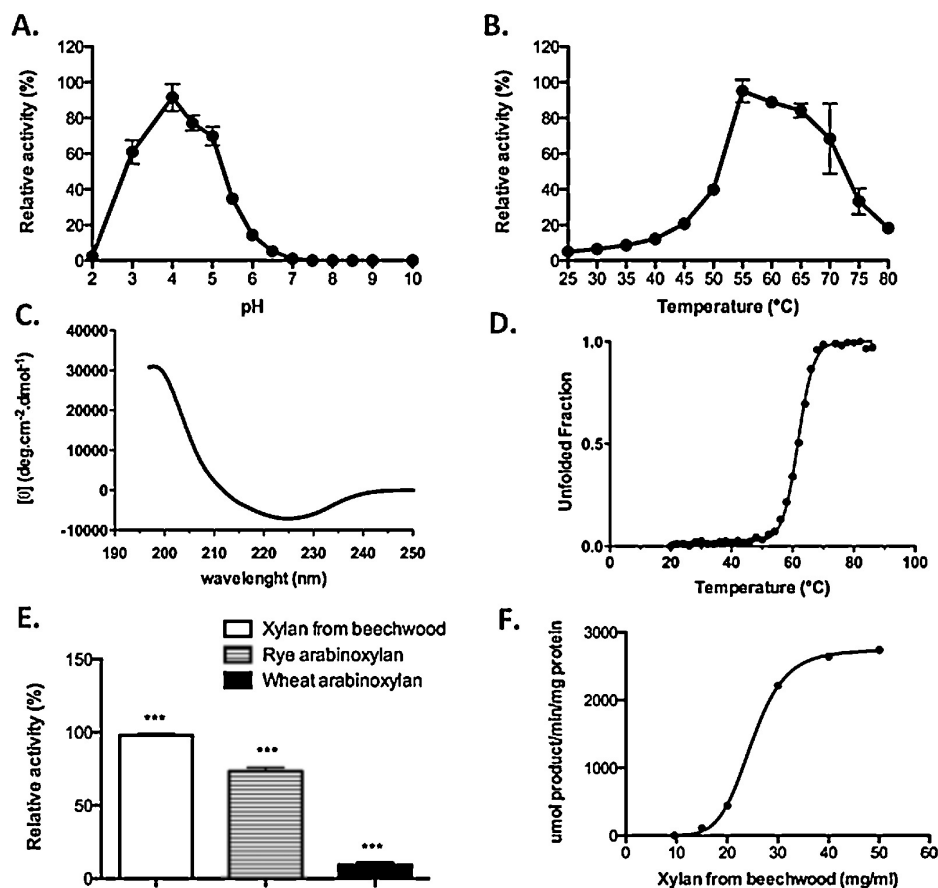


Fig. 3. Physico-chemical properties of PbXynA. (A) The effect of pH on the PbXynA enzymatic activity. (B) The effect of temperature on PbXynA enzymatic activity. (C) Far-UV CD spectra of PbXynA at pH 4.0 and 20 °C. The circular dichroism spectrum represents an average of eight scans. (D) Thermal denaturation curve at pH 4.0. The thermal denaturation curve was obtained by monitoring at 224.8 nm and at a temperature range of 20–100 °C. (E) The influence of branching in PbXynA activity using different substrates: xylan from beechwood, rye arabinoxylan and wheat arabinoxylan. (F) PbXynA has a sigmoidal kinetic behavior.

source, suggesting this xylanase is subjected to catabolite repression (Fig. 1B). Comparison of the *P. brasiliensis* sp. nov. xylanase activities at 48 h with other xylanase activities from filamentous fungi (for example, compare with *Aspergillus niger* under the same conditions [19]) showed that *P. brasiliensis* sp. nov. has about 20 times higher activity. We also tested the activity of *P. brasiliensis* supernatant after 96 h against a set of plant cell wall polysaccharides (Fig. 1C). This supernatant showed a high activity on xylan from beechwood and rye arabinoxylan (arabinose to xylose ratio, 38 to 62), indicating xylanases as major induced hemicellulases. The absence of hydrolysis on carboxymethylcellulose (CMC) indicated the absence of cellulase in the crude extract.

3.2. *P. brasiliensis* endoxylanase (PbXynA) production and purification

PbXynA was purified with a gel filtration step using a Superdex 75 column. The PbXynA homogeneity was evaluated by SDS-PAGE and the apparent molecular weight was approximately 24 kDa (Fig. 2). PbXynA was the major protein in the *P. brasiliensis* supernatant. This band was excised from the gel and digested with trypsin, and the analysis by MS/MS and “Basic Local alignment Tool” (BLAST) clearly showed that PbXynA is very similar to a hypothetical protein PANT.22d00001 from *P. antarctica* T-34 which belongs to GH11 family. The PbXynA was identified by MALDI-TOF-TOF

providing 24% sequence coverage. Blastp analyses of these peptides against the genome sequence of this strain revealed 100% identity with a *P. brasiliensis* sp. nov. open reading frame (data not shown). Even more, Blastp analyses revealed that PbXynA has high identity with other xylanases from basidiomycetous yeasts, such as *P. antarctica* [20] (value: 2e–127; identity: 82%; similarity: 87%), *P. hubeiensis* strain SY62 [21] (6e–115; 77%; 86%), *Ustilago hordei* [22] (3e–105; 71%; 79%), and *U. maydis* (7e–101; 74%; 83%).

Adsul et al. [7] have identified two different xylanases that have molecular masses of 33.3 kDa (PhX33) and 20.1 kDa (PhX20) from a strain of *P. hubeiensis* and purified them to homogeneity. Both xylanases produced only xylooligosaccharides with degree of polymerization 3–7 without formation of xylose and xylobiose. Interestingly, five partial peptides of these xylanases strongly indicate they are novel since there is no significant identity with known fungal xylanase sequences in the current databases. PbXynA and PhX20 have similar molecular weights and other comparable biochemical properties. However, we have compared by using Blastp these five peptides to PbXynA, and observed identity only for one of them (VVNELFQEDGSGFRHTQVNQPSLQGTKNKSGSVTLNHFNAWK) at the aminoacid positions 129 to 185 (value: 2e–14, identity: 49%, similarity: 63%). These results strongly suggest that PbXynA appears to be a novel xylanase in *P. brasiliensis* sp. nov. different from *P. hubeiensis* PhX20.

Table 1
Kinetic parameters of PbXynA.

Kinetic parameters	PbXynA
$V_{\max}$ (μ moles product/min/mg protein)	2792.15 ± 26.23
K (mg/ml)	25.04 ± 1.2
n_H (Hill coefficient)	6.5 ± 0.2877
$V_{\max}/K$	111.49

3.3. PbXynA has a sigmoidal kinetic profile

PbXynA exhibited activity over the pH range 3.0–5.0 with an optimum at pH 4.0 (Fig. 3A), suggesting that it is an acid xylanase. It also exhibited maximum xylanase activity at 55 °C (Fig. 3B). These data are according to the literature where the peak xylanase activity is between pH 4.0 and 6.5 and temperatures are between 40 °C and 80 °C [4].

The PbXynA far-UV CD spectrum showed a minimum negative band at 224.8 nm (Fig. 3C), typical of proteins mainly containing β -sheet structures [23]. This profile was expected because of the family GH11 enzyme consists of a single domain predominantly β -sheet [24]. The CD technique is also a very convenient approach for monitoring protein denaturation and is based on the evaluation of changes in optical activity as a function of temperature or denaturing agent [23]. We followed the PbXynA thermal denaturation by measuring the change in θ at 224.8 nm, from 20 to 100 °C (Fig. 3D). The analysis by CD outlined that the midpoint temperature (T_m) for PbXynA was at 61.18 °C. The optimum temperature was too close to the T_m , then by working with temperatures slightly lower than optimal ensures the integrity of the protein.

We also analyzed the influence of xylan branching by PbXynA catalytic efficiency (Fig. 3E). The activity on rye arabinoxylan was about 75% relative to xylan from beechwood (Fig. 3E). Thus the PbXynA had the ability to bind and hydrolyze decorated xylan. The ratio xylose to arabinose is the same for rye and wheat arabinoxylan (Megazyme) but the wheat arabinoxylan is more insoluble. Thus, apart from branching rate, the solubility was an important substrate parameter for PbXynA catalytic efficiency.

As shown in Fig. 3F, the PbXynA had a sigmoidal profile for kinetic studies by varying the concentrations of xylan from beechwood (Table 1). Enzymes involved in the hydrolysis of plant cell wall cellulose and hemicellulose are typically modular, consisting of conserved catalytic domains and carbohydrate-binding modules (CBMs), joined by flexible linker sequences. Binding modules facilitate catalysis by associating the catalytic domain with its substrates [25]. Despite the clear advantage of having CBMs, some glycoside hydrolases consist of a catalytic domain only. The majority of enzymes from GH11 family lack a CBM, raising the question of how these glycoside hydrolases recognize and efficiently catalyze the hydrolysis of hemicelluloses [25]. Studies revealed the presence of other substrate binding regions situated on the surface of the structural unit that contains the catalytic site, rather than on an auxiliary domain. These substrate binding sites are located at a certain distance from the active site and are called secondary binding sites (SBS) [26]. The SBSs are present in different regions of the GH11 xylanases and indicates that the incorporation of an SBS is a strategy of these single-module xylanases to compensate for the lack of CBMs [27]. In the study of Cuyvers et al. [26], the presence of a SBS gives rise to higher activities on all tested polymeric substrates. For small substrates, the binding of substrate to the SBS and the active site appear to be independent of each other. For larger substrates, on the contrary, cooperative binding occurs and leads to enhanced hydrolysis of these substrates [25].

The kinetic parameters analysis showed that the enzyme-substrate apparent dissociation constant (K) for PbXynA was 25.04 mg/ml, using xylan from beechwood. The $V_{\max}$ value of

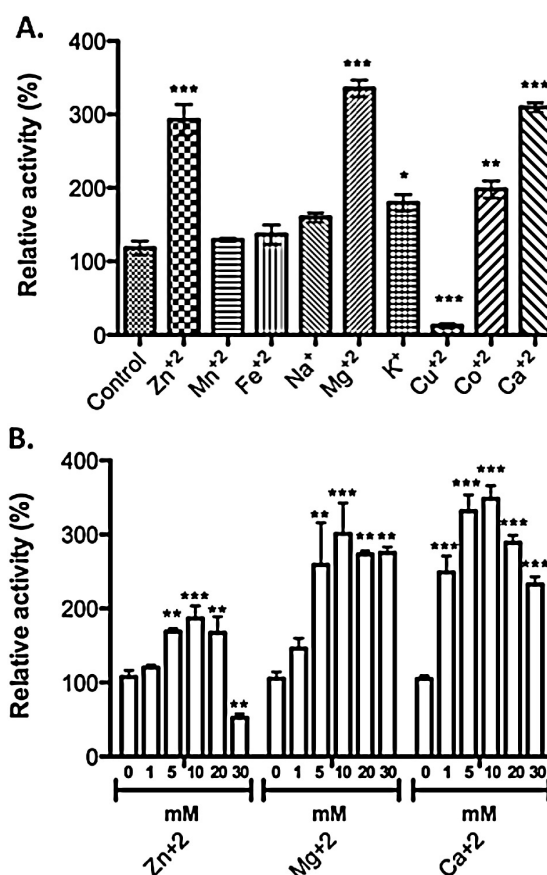


Fig. 4. PbXynA activity is influenced by bivalent ions. (A) Effects of cations on the activity of purified PbXynA. The final cations concentration was 10 mM. The activity of the experiment after treatment with Chelex 100 resin was defined as 100% (control). (B) The determination of optimal ions concentration on the activity of PbXynA.

PbXynA was 2792.15 μ mol product/min/mg protein and the Hill coefficient (n_H) was 6.5 (Table 1). Comparing the $V_{\max}$ of the enzyme in this study with others industrially important xylanases, the high potential of substrate–product conversion of PbXynA can be highlighted (Table 2).

3.4. PbXynA is highly activated by Ca²⁺ and Mg²⁺

The effects of metal ions on the enzyme activity were also investigated (Fig. 4). The results revealed that the enzyme activity decreased 62.43% after bivalent ion removal by Chelex 100 resin (data not shown) and was significantly stimulated by K⁺, Co²⁺, Zn²⁺, Mg²⁺ and Ca²⁺, and completely inactivated by Cu²⁺ (Fig. 4A). The effects of other metal ions on the enzyme activity were not significant (Fig. 4A). The PbXynA activity was completely inhibited by Cu²⁺ and highly stimulated by Ca²⁺. These results were similar to those obtained for *Thermotoga thermarum* xylanase [6]. The xylanase activity was greatly stimulated by Zn²⁺, Mg²⁺ and Ca²⁺, and was enhanced by nearly 200%, 300% and 400%, respectively, in the solution of 10 mM of these ions (Fig. 4B). Our results suggest that a few metal ions, especially Ca²⁺, are probably required for maintaining the structure stability. The activation by bivalent ions has a great biotechnological appeal since the controlled addition of ions can highly increase the hydrolysis efficiency during a process.

Table 2
Properties of G11 eukaryotic endo-1,4-xylanases.

Xylanase	Species	Mass (kDa)	$V_{\max}$ (μ moles product/min/mg protein)	Reference
PbXynA	<i>Pseudozyma brasiliensis</i> sp. nov.	24	2792.0	This study
–	<i>Aspergillus ficuum</i> AF-98	35.0	18.38	[35]
XynG2	<i>Aspergillus oryzae</i> RIB40	21	123.0	[36]
EX2	<i>Trichoderma</i> sp. K9301	20.1	526.0	[32]
XylG	<i>Paecilomyces thermophila</i>	45.0	7.70	[37]

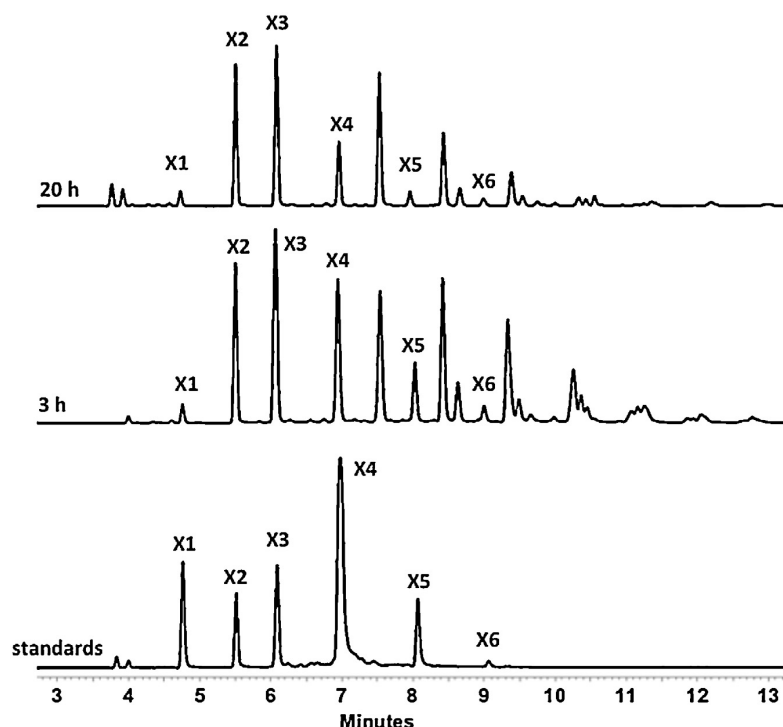


Fig. 5. PbXynA produces longer xylo-oligosaccharides (XOS) with high degree of polymerization. Capillary electrophoresis of APTS-labeled XOS produced after xylan from beechwood hydrolysis by PbXynA.

3.5. PbXynA yielded XOS with high degree of polymerization

We have evaluated which xylose oligomers are produced by PbXynA (Fig. 5). The PbXynA hydrolyzed xylan from beechwood yielded a complete set of xylo-oligosaccharides (XOS), such as xylobiose, xylotriose, xylo-tetraose, xylopentaose and xylohexaose. The production of XOS with a high degree of polymerization (DP) by PbXynA is not common in other eukaryotic xylanases [28,29]. Furthermore, it is worth noting that during the hydrolysis the high DP XOS were not hydrolyzed into shorter XOS. Although there is formation of xylose after 20 h, xylose is considered a byproduct because the enzyme itself did not form xylose as a main product but depending on the points of hydrolysis, there are xylose traces, and this is very common in xylanases. The different *P. hubeiensis* xylanases PhX33 and PhX20 also produced only XOS with DP 3–7 without formation of xylose and xylobiose [7].

XOS are useful for a variety of purposes [30]. XOS are naturally present in fruits, vegetables, bamboo, honey and milk, and can be produced at industrial scale by enzymatic hydrolysis from xylan. They are non-digestible carbohydrates and have been suggested to exert prebiotic activity [31]. XOS have been shown to reduce cholesterol, maintain gastrointestinal health, improve the biological availability of calcium [32] and reduce the risk of colon cancer

in rats [33]. Furthermore, XOS exert beneficial effects on gastrointestinal health by increasing mainly the bifidobacteria population [33], in other words, XOS exert preferential stimulatory effects on a number of the health-promoting microorganisms, inhibiting the population of potential pathogens at relatively low level [34]. In spite of their benefits to human health and biochemical importance, XOs are relatively expensive. The routine methods for XOS production are costly and require pure monosaccharides. Thus, the xylanases producing only XOS will probably be the potential biocatalysts to produce XOS with cost-effective technologies [7].

4. Conclusion

The bioprospection of yeast strains that colonize the intestinal tract of insects that are associated with sugarcane roots provide us a good strategy to find new enzymes with potential biotechnological applications. We have isolated a new basidiomycetous yeast species *P. brasiliensis* sp. nov. strain from the intestinal tract of *Chrysomelidae* larvae by growth enrichment on xylose. This strain produces high levels of a xylanase, PbXynA, when grown in xylan or xylose. This endoxylanase belongs to GH11 family showing a higher specific activity than other G11 eukaryotic xylanases previously studied. Furthermore, PbXynA exhibits sigmoidal

kinetics behavior with a high conversion speed substrate–product, activation by bivalent ions such as Ca^{2+} and application in XOS production with degree of polymerization 3–6 as end products from xylan hydrolysis. Taken together, PbXynA properties show a great biotechnological potential and commercial value.

Acknowledgements

We thank Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Vale S.A., from Brazil, for financial support, and we gratefully acknowledge the provision of time on the MAS facilities at the National Laboratory of Biosciences (LNBio) from the National Center of Energy and Materials (CNPEM), Ministry of Science and Technology, Brazil. ARLD is FAPESP postdoctoral fellow (2011/02169-4).

References

- Paes G, Berrin JG, Beaugrand J. GH11 xylanases: structure/function/properties relationships and applications. *Biotechnology Advances* 2012;30:564–92.
- Damasio AR, Ribeiro LF, Furtado GP, Segato F, Almeida FB, Crivellari AC, Buckenridge MS, Souza TA, Murakami MT, Ward RJ, Prade RA, Polizeli ML. Functional characterization and oligomerization of a recombinant xyloglucan-specific endo-beta-1,4-glucanase (GH12) from *Aspergillus niger*. *Biochimica et Biophysica Acta* 2012;1824:461–7.
- Song Y, Lee YG, Choi IS, Lee KH, Cho EJ, Bae HJ. Heterologous expression of endo-1,4-beta-xylanase A from *Schizophyllum commune* in *Pichia pastoris* and functional characterization of the recombinant enzyme. *Enzyme and Microbial Technology* 2013;52:170–6.
- Polizeli MLTM, Rizzatti ACS, Monti R, Terenzi HF, Jorge JA, Amorim DS. Xylanases from fungi: properties and industrial applications. *Applied Microbiology and Biotechnology* 2005;67:577–91.
- Santos CR, Meza AN, Hoffmam ZB, Silva JC, Alvarez TM, Ruller R, Giesel GM, Verli H, Squina FM, Prade RA, Murakami MT. Thermal-induced conformational changes in the product release area drive the enzymatic activity of xylanases 10B: crystal structure, conformational stability and functional characterization of the xylanase 10B from *Thermotoga petrophila* RKU-1. *Biochemical and Biophysical Research Communications* 2010;403:214–9.
- Shi H, Zhang Y, Li X, Huang Y, Wang L, Wang Y, Ding H, Wang F. A novel highly thermostable xylanase stimulated by Ca^{2+} from *Thermotoga thermarum*: cloning, expression and characterization. *Biotechnology for Biofuels* 2013;6:26.
- Adsul MG, Bastawde KB, Gokhale DV. Biochemical characterization of two xylanases from yeast *Pseudozyma hubeiensis* producing only xylooligosaccharides. *Bioresource Technology* 2009;100:6488–95.
- Konishi M, Morita T, Fukuoka T, Imura T, Kakugawa K, Kitamoto D. Production of different types of mannosylerythritol lipids as biosurfactants by the newly isolated yeast strains belonging to the genus *Pseudozyma*. *Applied Microbiology and Biotechnology* 2007;75:521–31.
- Konishi M, Morita T, Fukuoka T, Imura T, Kakugawa K, Kitamoto D. Efficient production of mannosylerythritol lipids with high hydrophilicity by *Pseudozyma hubeiensis* KM-59. *Applied Microbiology and Biotechnology* 2008;78:37–46.
- Konishi M, Fukuoka T, Nagahama T, Morita T, Imura T, Kitamoto D, Hatada Y. Biosurfactant-producing yeast isolated from *Calyptogena soyoe* (deep-sea cold-seep clam) in the deep sea. *Journal of Bioscience and Bioengineering* 2010;110:169–75.
- Konishi M, Nagahama T, Fukuoka T, Morita T, Imura T, Kitamoto D, Hatada Y. Yeast extract stimulates production of glycolipid biosurfactants, mannosylerythritol lipids, by *Pseudozyma hubeiensis* SY62. *Journal of Bioscience and Bioengineering* 2011;111:702–5.
- Takahashi M, Morita T, Fukuoka T, Imura T, Kitamoto D. Glycolipid biosurfactants, mannosylerythritol lipids, show antioxidant and protective effects against H(2)O(2)-induced oxidative stress in cultured human skin fibroblasts. *Journal of Oleo Science* 2012;61:457–64.
- Morita T, Ishibashi Y, Fukuoka T, Imura T, Sakai H, Abe M, Kitamoto D. Production of glycolipid biosurfactants, cellobiose lipids by *Cryptococcus humicola* JCM 1461 and their interfacial properties. *Bioscience, Biotechnology, and Biochemistry* 2011;75:1597–9.
- Sambrook J, Russell DW. Molecular cloning: a laboratory manual. New York: Cold Spring Harbor Laboratory Press; 2001.
- Miller GL. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Analytical Chemistry* 1959;31:426–8.
- Leone FA, Baranauskas JA, Furriel RP, Borin IA. SigrafW: an easy-to-use program for fitting enzyme kinetic data. *Biochemistry and Molecular Biology Education* 2005;33:399–403.
- Naran R, Pierce ML, Mort AJ. Detection and identification of rhamnogalacturonan lyase activity in intercellular spaces of expanding cotton cotyledons. *Plant Journal: For Cell and Molecular Biology* 2007;50:95–107.
- Chamnanpa T, Limtong P, Srisuk N, Limtong S. *Pseudozyma vetiver* sp. nov., a novel anamorphic ustilaginomycetous yeast species isolated from the phylloplane in Thailand. *Antonie Van Leeuwenhoek* 2013;104:637–44.
- de Souza WR, de Gouvea PF, Savoldi M, Malavazi I, de Souza Bernardes LA, Goldman MH, de Vries RP, de Castro Oliveira JV, Goldman GH. Transcriptome analysis of *Aspergillus niger* grown on sugarcane bagasse. *Biotechnology for Biofuels* 2011;4:40.
- Morita T, Koike H, Koyama Y, Hagiwara H, Ito E, Fukuoka T, Imura T, Machida M, Kitamoto D. Genome sequence of the basidiomycetous yeast *Pseudozyma antarctica* T-34, a producer of the glycolipid biosurfactants mannosylerythritol lipids. *Genome Announcements* 2013;1:e0006413.
- Konishi M, Hatada Y, Horiuchi J. Draft genome sequence of the basidiomycetous yeast-like fungus *Pseudozyma hubeiensis* SY62, which produces an abundant amount of the biosurfactant mannosylerythritol lipids. *Genome Announcements* 2013;1:e00409-13.
- Lefebvre F, Joly DL, Labbe C, Teichmann B, Linning R, Belzile F, Bakkeren G, Belanger RR. The transition from a phytopathogenic smut ancestor to an anamorphic biocontrol agent deciphered by comparative whole-genome analysis. *Plant Cell* 2013;25:1946–59.
- Corrêa DHA, Ramos CHI. The use of circular dichroism spectroscopy to study protein folding, form and function. *African Journal of Biochemistry Research* 2009;3:164–73.
- Collins T, Gerday C, Feller G. Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiology Reviews* 2005;29:3–23.
- Ludwiczek ML, Heller M, Kantner T, McIntosh LP. A secondary xylan-binding site enhances the catalytic activity of a single-domain family 11 glycoside hydrolase. *Journal of Molecular Biology* 2007;373:337–54.
- Cuyvers S, Dornez E, Rezaei MN, Pollet A, Delcour JA, Courtin CM. Secondary substrate binding strongly affects activity and binding affinity of *Bacillus subtilis* and *Aspergillus niger* GH11 xylanases. *FEBS Journal* 2011;278:1098–111.
- Cuyvers S, Dornez E, Delcour JA, Courtin CM. Occurrence and functional significance of secondary carbohydrate binding sites in glycoside hydrolases. *Critical Reviews in Biotechnology* 2012;32:93–107.
- Squina FM, Mort AJ, Decker SR, Prade RA. Xylan decomposition by *Aspergillus clavatus* endo-xylanase. *Protein Expression and Purification* 2009;68:65–71.
- Damásio ARL, Silva TM, Almeida FBR, Squina FM, Ribeiro DA, Leme AFP, Segato F, Prade RA, Jorge JA, Terenzi HF, Polizeli MLTM. Heterologous expression of an *Aspergillus niger* xylanase GH11 in *Aspergillus nidulans* and its characterization and application. *Process Biochemistry* 2011;46:1236–42.
- Vazquez MJ, Alonso JL, Dominguez H, Parajo JC. Xylooligosaccharides: manufacture and applications. *Trends in Food Science & Technology* 2000;11:387–93.
- Makelainen H, Juntunen M, Hasselwander O. Prebiotic potential of xylooligosaccharides. In: Charalampopoulos D, Rastall RA, editors. *Prebiotics and Probiotics Science and Technology*. New York: Springer; 2009.
- Chen LL, Zhang M, Zhang DH, Chen XL, Sun CY, Zhou BC, Zhang YZ. Purification and enzymatic characterization of two beta-endoxylanases from *Trichoderma* sp. K9301 and their actions in xylooligosaccharide production. *Bioresource Technology* 2009;100:5230–6.
- Hsu CK, Liao JW, Chung YC, Hsieh CP, Chan YC. Xylooligosaccharides and fructooligosaccharides affect the intestinal microbiota and precancerous colonic lesion development in rats. *Journal of Nutrition* 2004;134:1523–8.
- Rhew BK, Lee JW, Lee CS, Hyun SY, Park YJ, Ahn JB, Yang CK, Yoon S. Effect of xylooligosaccharides on the growth of intestinal microflora. *Korean Journal of Microbiology and Biotechnology* 2002;30:380–7.
- Lu F, Lu M, Lu Z, Bie X, Zhao H, Wang Y. Purification and characterization of xylanase from *Aspergillus ficuum* AF-98. *Bioresource Technology* 2008;99:5938–41.
- Kimura T, Suzuki H, Furuhashi H, Aburatani T, Morimoto K, Karita S, Sakka K, Ohmiya K. Molecular cloning, overexpression, and purification of a major xylanase from *Aspergillus oryzae*. *Bioscience, Biotechnology, and Biochemistry* 2000;64:2734–8.
- Zhang M, Jiang Z, Yang S, Hua C, Li L. Cloning and expression of a *Paecilomyces thermophila* xylanase gene in *E. coli* and characterization of the recombinant xylanase. *Bioresource Technology* 2010;101:688–95.

9. Considerações finais

Este trabalho propiciou o isolamento de duas leveduras do trato intestinal de insetos parasitas da cana-de-açúcar: *Rhodotorula dairenensis* TAB01 e *Pseudozyma brasiliensis* GHG001. Essas leveduras mostraram-se ótimas assimiladoras de xilose e promissoras para este estudo. As atividades das enzimas XR e XDH (de ambas as leveduras) foram avaliadas e mostraram-se similares, mas inferiores as da *S. stipitis* NRRL-Y7124. Além disso, as enzimas XR destas leveduras não utilizam NADH como cofator, sugerindo que possa haver um desbalanço redox nas células, e possivelmente um acúmulo do subproduto xilitol nas células, o que não é desejável para produção de etanol. Ao superexpressar os genes de assimilação de xilose nas cepas de *S. cerevisiae* industrial PE-2 e PE-XKS1 (com expressão aumentada da *xks1* em um dos alelos), foi verificado que elas não conseguem utilizar a xilose para o seu crescimento. Diante disso, é possível que uma série de fatores contribuíram negativamente para esse resultado, tais como o desbalanço redox na célula causado pela não utilização do NADH como cofator e a baixa atividade endógena das enzimas XR e XDH dessas leveduras. Entretanto, é considerado como um ponto positivo o isolamento destas duas leveduras no trato intestinal de insetos, mostrando o potencial de se bioprospectar micro-organismos em nichos que potencialmente são enriquecidos ou favoráveis ao seu crescimento. Os genes de ambas as leveduras podem ser melhorados utilizando abordagens de engenharia metabólica e evolução dirigida visando seu uso em cepas industriais para a produção de etanol a partir de xilose. Além disso, o trabalho foi de grande importância, pois através dele foi isolada uma nova espécie de levedura e caracterizada uma enzima xilanolítica (PbXynA) com uma alta atividade não encontrada em outras xilanases de eucariotos. Isso implica em uma redução de quantidade de enzima necessária em processos industriais, como por exemplo, nos coquetéis enzimáticos para a quebra da parede celular vegetal no processo de produção de etanol, e consequentemente na redução de custos desses processos.

10. Referências

ABASCAL, F., ZARDOYA, R., POSADA, D. ProtTest: selection of best-fit models of protein evolution. **Bioinformatics** v. 21 n. 9, p. 2104–2105., 2005.

ADSUL, M. G.; BASTAWDE, K. B.; GOKHALE, D. V. Biochemical characterization of two xylanases from yeast *Pseudozyma hubeiensis* producing only xylooligosaccharides. **Bioresource technology**, v. 100, n. 24, p. 6488–95, 2009. Elsevier Ltd. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19692229>>. Acesso em: 26/1/2014.

AMORIM, H. V.; LOPES, M. L.; CASTRO OLIVEIRA, J. V. DE; BUCKERIDGE, M. S.; GOLDMAN, G. H. Scientific challenges of bioethanol production in Brazil. **Applied microbiology and biotechnology**, v. 91, n. 5, p. 1267–1275, 2011. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21735264>>. Acesso em: 7/2/2014.

ARGUESO, J. L.; CARAZZOLLE, M. F.; MIECZKOWSKI, P. A; et al. Genome structure of a *Saccharomyces cerevisiae* strain widely used in bioethanol production. **Genome research**, v. 19, n. 12, p. 2258–70, 2009. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2792172&tool=pmcentrez&rendertype=abstract>>. Acesso em: 7/2/2014.

BABRZADEH, F.; JALILI, R.; WANG, C.; et al. Whole-genome sequencing of the efficient industrial fuel-ethanol fermentative *Saccharomyces cerevisiae* strain CAT-1. **Molecular genetics and genomics : MGG**, v. 287, n. 6, p. 485–94, 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/22562254>>. Acesso em: 7/2/2014.

BASSO, L. C.; AMORIM, H. V DE; OLIVEIRA, A. J. DE; LOPES, M. L. Yeast selection for fuel ethanol production in Brazil. **FEMS yeast research**, v. 8, n. 7, p. 1155–63, 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18752628>>. Acesso em: 24/1/2014.

BASSO, T. P.; BASSO, T. O.; GALLO, C. R.; BASSO, L. C. Towards the Production of Second Generation Ethanol from Sugarcane Bagasse in Brazil. ,2013.

BELLEMAIN, E.; CARLSEN, T.; BROCHMANN, C.; et al. ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. **BMC microbiology**, v. 10, p. 189, 2010. Disponível em:

<<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2909996&tool=pmcentrez&rendertype=abstract>>. .

BORGES, T. A.; SOUZA, A. T. DE; SQUINA, F. M.; et al. Biochemical characterization of an endoxylanase from *Pseudozyma brasiliensis* sp. nov. strain GHG001 isolated from the intestinal tract of Chrysomelidae larvae associated to sugarcane roots. **Process Biochemistry**, v. 49, n. 1, p. 77–83, 2014. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1359511313005655>>. Acesso em: 8/2/2014.

BURA, R.; VAJZOVIC, A.; DOTY, S. L. Novel endophytic yeast *Rhodotorula mucilaginosa* strain PTD3 I: production of xylitol and ethanol. **Journal of industrial microbiology & biotechnology**, v. 39, n. 7, p. 1003–11, 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/22399239>>. Acesso em: 7/2/2014.

CAI, Z.; ZHANG, B.; LI, Y. Engineering *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: reflections and perspectives. **Biotechnology journal**, v. 7, n. 1, p. 34–46, 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/22147620>>. Acesso em: 2/2/2014.

COLLA, E.; PEREIRA, A. B.; HERNALSTEENS, S.; MAUGERI, F.; RODRIGUES, M. I. Optimization of Trehalose Production by *Rhodotorula dairenensis* Following a Sequential Strategy of Experimental Design. **Food and Bioprocess Technology**, v. 3, n. 2, p. 265–275, 2008. Disponível em: <<http://link.springer.com/10.1007/s11947-008-0081-y>>. Acesso em: 20/3/2014.

DIAS, M. O. S.; CUNHA, M. P.; JESUS, C. D. F.; et al. Second generation ethanol in Brazil: can it compete with electricity production? **Bioresource technology**, v. 102, n. 19, p. 8964–71, 2011. Elsevier Ltd. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21795041>>. Acesso em: 5/2/2014.

EDDY, S. R. BIOINFORMATICS REVIEW Profile hidden Markov models. , p. 755–763, 1998.

FELL, J. W.; BOEKHOUT, T.; FONSECA, A.; SCORZETTI, G.; STATZELL-TALLMAN, A. Biodiversity and systematics of basidiomycetous yeasts as determined by large-subunit rDNA D1/D2 domain sequence analysis. **International journal of systematic and evolutionary microbiology**, v. 50 Pt 3, n. 2000, p. 1351–71, 2000. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/10843082>>. .

GADANHO, M.; SAMPAIO, J. P. Polyphasic taxonomy of the basidiomycetous yeast genus *Rhodotorula* : Rh . glutinis sensu stricto. , v. 2, 2002.

GOLDEMBERG, J.; GUARDABASSI, P. The potential for first-generation ethanol production. , p. 17–24, 2009.

GRIGORIEV, I. V; NORDBERG, H.; SHABALOV, I.; et al. The genome portal of the Department of Energy Joint Genome Institute. **Nucleic acids research**, v. 40, n. Database issue, p. D26–32, 2012.

GUTIÉRREZ-ALONSO, P.; FERNÁNDEZ-ARROJO, L.; PLOU, F. J.; FERNÁNDEZ-LOBATO, M. Biochemical characterization of a beta-fructofuranosidase from *Rhodotorula dairenensis* with transfructosylating activity. **FEMS yeast research**, v. 9, n. 5, p. 768–73, 2009. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19486164>>. Acesso em: 20/3/2014.

HECTOR, R. E.; MERTENS, J. A.; BOWMAN, M. J.; et al. *Saccharomyces cerevisiae* engineered for xylose metabolism requires gluconeogenesis and the oxidative branch of the pentose phosphate pathway for aerobic xylose assimilation. , n. August, p. 645–660, 2011.

HO, N. W. Y.; CHEN, Z.; BRAINARD, A. P. Genetically Engineered *Saccharomyces* Yeast Capable of Effective Cofermentation of Glucose and Xylose. Genetically Engineered *Saccharomyces* Yeast Capable of Effective Cofermentation of Glucose and Xylose. , 1998.

JEFFRIES, T. W. Engineering yeasts for xylose metabolism. **Current opinion in biotechnology**, v. 17, n. 3, p. 320–6, 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16713243>>. Acesso em: 2/2/2014.

JEPPSSON, M.; TRAFF, K.; JOHANSSON, B.; HAHNHAGERDAL, B.; GORWAGRAUSLUND, M. Effect of enhanced xylose reductase activity on xylose consumption and product distribution in xylose-fermenting recombinant. **FEMS Yeast Research**, v. 3, n. 2, p. 167–175, 2003. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1567135602001861>>. Acesso em: 7/2/2014.

KÄMPER, J.; KAHMANN, R.; BÖLKER, M.; et al. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. **Nature**, v. 444, n. 7115, p. 97–101, 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17080091>>. Acesso em: 24/1/2014.

KATOH, K.; TOH, H. Recent developments in the MAFFT multiple sequence alignment program. **Briefings in bioinformatics**, v. 9, n. 4, p. 286–98, 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18372315>>. Acesso em: 21/1/2014.

KHATTAB, S. M. R.; WATANABE, S.; SAIMURA, M.; KODAKI, T. A novel strictly NADPH-dependent *Pichia stipitis* xylose reductase constructed by site-directed mutagenesis. **Biochemical and biophysical research communications**, v. 404, n. 2, p. 634–7, 2011. Elsevier Inc. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21146502>>. Acesso em: 31/1/2014.

KIM, S. R.; PARK, Y.-C.; JIN, Y.-S.; SEO, J.-H. Strain engineering of *Saccharomyces cerevisiae* for enhanced xylose metabolism. **Biotechnology advances**, v. 31, n. 6, p. 851–61, 2013. Elsevier Inc. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/23524005>>. Acesso em: 7/2/2014.

LEANDRO, M. J.; SPENCER-MARTINS, I.; GONÇALVES, P. The expression in *Saccharomyces cerevisiae* of a glucose/xylose symporter from *Candida intermedia* is affected by the presence of a glucose/xylose facilitator. **Microbiology (Reading, England)**, v. 154, n. Pt 6, p. 1646–55, 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18524919>>. Acesso em: 7/2/2014.

MARTIN, K. J.; RYGIEWICZ, P. T. Fungal-specific PCR primers developed for analysis of the ITS region of environmental DNA extracts. **BMC microbiology**, v. 5, p. 28, 2005. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1156903&tool=pmcentrez&rendertype=abstract>>. Acesso em: 24/1/2014.

MATSUSHIKA, A.; INOUE, H.; KODAKI, T.; SAWAYAMA, S. Ethanol production from xylose in engineered *Saccharomyces cerevisiae* strains: current state and perspectives. **Applied microbiology and biotechnology**, v. 84, n. 1, p. 37–53, 2009. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19572128>>. Acesso em: 24/1/2014.

NAIK, S. N.; GOUD, V. V.; ROUT, P. K.; DALAI, A. K. Production of first and second generation biofuels: A comprehensive review. **Renewable and Sustainable Energy Reviews**, v. 14, n. 2, p. 578–597, 2010. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1364032109002342>>. Acesso em: 20/1/2014.

NARDI, T.; CORICH, V.; GIACOMINI, A.; BLONDIN, B. A sulphite-inducible form of the sulphite efflux gene SSU1 in a *Saccharomyces cerevisiae* wine yeast. **Microbiology (Reading, England)**, v. 156, n. Pt 6, p. 1686–96, 2010. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/20203053>>. Acesso em: 7/2/2014.

PUNTA, M.; COGGILL, P. C.; EBERHARDT, R. Y.; et al. The Pfam protein families database. **Nucleic acids research**, v. 40, n. Database issue, p. D290–301, 2012.

RODRIGUES, J. A. R. Revisão. , v. 34, n. 7, p. 1242–1254, 2011.

SHI, W.; DING, S.-Y.; YUAN, J. S. Comparison of Insect Gut Cellulase and Xylanase Activity Across Different Insect Species with Distinct Food Sources. **BioEnergy Research**, v. 4, n. 1, p. 1–10, 2010. Disponível em: <<http://link.springer.com/10.1007/s12155-010-9096-0>>. Acesso em: 7/2/2014.

SIMS, R. E. H.; MABEE, W.; SADDLER, J. N.; TAYLOR, M. An overview of second generation biofuel technologies. **Bioresource technology**, v. 101, n. 6, p. 1570–80, 2010. Elsevier Ltd. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19963372>>. Acesso em: 24/1/2014.

SOCOL, C. R.; VANDENBERGHE, L. P. D. S.; MEDEIROS, A. B. P.; et al. Bioethanol from lignocelluloses: Status and perspectives in Brazil. **Bioresource technology**, v. 101, n. 13, p. 4820–5, 2010. Elsevier Ltd. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/20022746>>. Acesso em: 21/1/2014.

SONNENBERG, R.; NOLTE, A. W.; TAUTZ, D. An evaluation of LSU rDNA D1-D2 sequences for their use in species identification. **Frontiers in zoology**, v. 4, p. 6, 2007. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1805435&tool=pmcentrez&rendertype=abstract>>. Acesso em: 3/2/2014.

SOUZA, A. P.; LEITE, D. C. C.; PATTATHIL, S.; HAHN, M. G.; BUCKERIDGE, M. S. Composition and Structure of Sugarcane Cell Wall Polysaccharides: Implications for Second-Generation Bioethanol Production. **BioEnergy Research**, v. 6, n. 2, p. 564–579, 2012. Disponível em: <<http://link.springer.com/10.1007/s12155-012-9268-1>>. Acesso em: 22/1/2014.

STEPHEN, J. D.; COLUMBIA, B. ethanol be able to compete with first-generation ethanol? Opportunities for cost. , p. 159–176, 2012.

SUH, S.-O.; MCHUGH, J. V.; POLLOCK, D. D.; BLACKWELL, M. The beetle gut: a hyperdiverse source of novel yeasts. **Mycological Research**, v. 109, n. 3, p. 261–265, 2005. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0953756208614112>>. Acesso em: 7/2/2014.

VELASCO DE CASTRO OLIVEIRA, J.; CORRÊA, R. A.; BORGES, T. A.; RIAÑO-PACHÓN, D. M. Draft Genome Sequence of *Pseudozyma brasiliensis* sp. nov. Strain GHG001, a High Producer of Endo-1,4-Xylanase Isolated from an Insect Pest of Sugarcane. , , n. December, p. 4–5, 2013.

WAHLBOM, C.; VANZYL, W.; JONSSON, L.; HAHNHAGERDAL, B.; OTERO, R. Generation of the improved recombinant xylose-utilizing TMB 3400 by random mutagenesis and physiological comparison with CBS 6054. **FEMS Yeast**

Research, v. 3, n. 3, p. 319–326, 2003. Disponível em:
<<http://linkinghub.elsevier.com/retrieve/pii/S1567135602002064>>. Acesso em:
31/1/2014.

WANG, Q.-M.; JIA, J.-H.; BAI, F.-Y. *Pseudozyma hubeiensis* sp. nov. and *Pseudozyma shanxiensis* sp. nov., novel ustilaginomycetous anamorphic yeast species from plant leaves. **International journal of systematic and evolutionary microbiology**, v. 56, n. Pt 1, p. 289–93, 2006. Disponível em:
<<http://www.ncbi.nlm.nih.gov/pubmed/16403900>>. Acesso em: 7/2/2014.

XU, P.; BURA, R.; DOTY, S. L. Genetic analysis of D-xylose metabolism by endophytic yeast strains of *Rhodotorula graminis* and *Rhodotorula mucilaginosa*. **Genetics and molecular biology**, v. 34, n. 3, p. 471–8, 2011. Disponível em:
<<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3168190&tool=pmcentrez&rendertype=abstract>>. .

YANG, V. W.; JEFFRIES, T. W. Regulation of phosphotransferases in glucose- and xylose-fermenting yeasts. **Biotechnology for Fuels and Chemicals.**, p. 97-108, 1997.