

UNIVERSIDADE ESTADUAL PAULISTA “JULIO DE MESQUITA FILHO”
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CAMPUS DE JABOTICABAL

PROSPECÇÃO DE GENES COM INTERESSE
BIOTECNOLÓGICO

Viviane Schuch

Bióloga

JABOTICABAL – SÃO PAULO – BRASIL

Junho de 2011

UNIVERSIDADE ESTADUAL PAULISTA “JULIO DE MESQUITA FILHO”
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CAMPUS DE JABOTICABAL

PROSPECÇÃO DE GENES COM INTERESSE
BIOTECNOLÓGICO

Viviane Schuch

Orientadora: Prof^a. Dr^a. Eliana Gertrudes de Macedo Lemos

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Campus de Jaboticabal, como parte das exigências para a obtenção do título de Doutor em Microbiologia.

JABOTICABAL – SÃO PAULO – BRASIL

Junho de 2011

S384p Schuch, Viviane
Prospecção de Genes com Interesse Biotecnológico / Viviane Schuch. -- Jaboticabal, 2011
v, 80 f. : il. ; 28 cm

Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, 2011
Orientadora: Eliana Gertrudes de Macedo Lemos
Banca examinadora: Gabriel Padilla, Rodrigo Matheus Pereira, L Maria Carareto Alves e Jackson Antonio Marcondes de Souza
Bibliografia

1. Xilose isomerase. 2. Metagenoma. 3. *Burkholderia*. I. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 578.8:66

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação – Serviço Técnico de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.

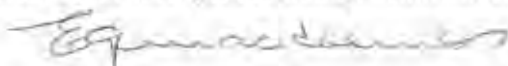
CERTIFICADO DE APROVAÇÃO

TÍTULO: PROSPECÇÃO DE GENES COM INTERESSE BIOTECNOLÓGICO

AUTORA: VIVIANE SCHUCH

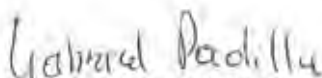
ORIENTADORA: Profa. Dra. ELIANA GERTRUDES DE MACEDO LEMOS

Aprovada como parte das exigências para obtenção do Título de DOUTOR EM MICROBIOLOGIA AGROPECUÁRIA, pela Comissão Examinadora:



Profa. Dra. ELIANA GERTRUDES DE MACEDO LEMOS

Departamento de Tecnologia / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal



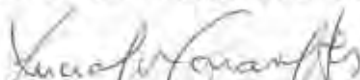
Prof. Dr. GABRIEL PADILLA

Instituto de Ciências Biomédicas / Universidade de São Paulo / São Paulo/SP



Prof. Dr. RODRIGO MATHEUS PEREIRA

Faculdade de Ciências Biológicas e Ambientais / Universidade Federal Da Grande Dourados



Profa. Dra. LUCIA MARIA CARARETO ALVES

Departamento de Tecnologia / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal



Prof. Dr. JACKSON ANTONIO MARCONDES DE SOUZA

Departamento de Biologia Aplicada À Agropecuária / Universidade Estadual Paulista "Júlio de Mesquita Filho"

Data da realização: 21 de junho de 2011,

DADOS CURRICULARES DA AUTORA

VIVIANE SCHUCH - nascida em 9 de novembro de 1976 em São Paulo – SP. Em 1995, ingressou no curso de Ciências Biológicas da UNICAMP – SP, e em 1997 o curso foi transferido para a USP – SP. Concluiu o Bacharelado em 2001 e a Licenciatura em 2002. Neste período lecionou Ciências para alunos da 5^a e 6^a séries do ensino público em Paulínia – SP, trabalhou como Técnica do Laboratório de Microbiologia da Clariant S.A. e como Técnica do Laboratório Didático de Ciências e Biologia do Colégio Visconde de Porto Seguro Unidade I, ambos em São Paulo. Em 2004, mudou-se para a cidade de Jaboticabal e iniciou os trabalhos no Laboratório de Bioquímica de Microrganismos e Plantas do Departamento de Tecnologia da FCAVJ – UNESP, sob supervisão da Prof. Dr. Eliana Gertrudes de Macedo Lemos. Em 2005, ingressou no curso de mestrado em Microbiologia Agropecuária. Em 2007, concluiu o curso de mestrado e ingressou no curso de doutorado no mesmo programa, cujos resultados estão descritos neste trabalho

“somos feitos do mesmo material de que são feitos os sonhos”

William Shakespeare

Aos meus pais, Edward Robert Schuch e Flora Schuch, pelo amor e apoio incondicional.

Ao meu companheiro, Guilherme Gaspar, e ao meu filho Nicolas (que ainda mora na minha barriga), pelo amor, carinho e compreensão.

Aos meus irmãos, Cristiano e Guilherme, pelo amor e apoio incondicional.

A minha professora, Dr^a Eliana Gertrudes de Macedo Lemos, pelo incentivo científico constante.

DEDICO

AGRADECIMENTOS

A FAPESP, pelo auxílio financeiro;

a minha família, pelo apoio incondicional;

a Prof^a Dr^a Eliana Gertrudes de Macedo Lemos, pela confiança em mim depositada;

a todos os amigos do Laboratório de Bioquímica de Microrganismos e Plantas, pela acolhida e amizade;

a amiga Elisângela Soares Gomes, pelo exemplo de esforço e dedicação;

ao programa de Pós-Graduação em Microbiologia Agropecuária, pela oportunidade que me foi dada;

a Faculdade de Ciências Agrárias e Veterinárias – Campus de Jaboticabal e a todos os professores que participaram da minha formação;

Muito Obrigada!

SUMÁRIO

ABREVIATURAS	viii
UNIDADES	ix
RESUMO E PALAVRAS CHAVE	x
SUMMARY AND KEYWORDS	xi
I. INTRODUÇÃO.....	1
II - REVISÃO DA LITERATURA.....	3
2.1. O emprego da enzima xilose isomerase na produção de biocombustíveis.....	3
2.2. Utilização da enzima xilose isomerase na produção de frutose.....	9
2.3. Prospecção de genes utilizando a abordagem metagenômica.....	10
2.4. Prospecção de genes em isolados pertencentes ao gênero <i>Burkholderia</i>	13
2.5. Objetivos.....	14
III. MATERIAIS E MÉTODOS.....	15
3.1. Coleta das amostras de solo.....	15
3.2. Extração de DNA metagenômico.....	16
3.3. Quantificação do DNA metagenômico em gel de agarose.....	16
3.4. Quantificação do DNA metagenômico no NanoDrop.....	16
3.5. Construção da biblioteca metagenômica em vetor cosmídeo e fosmídeo.....	16
3.6. Construção da biblioteca metagenômica em plasmídeo.....	20
3.7. Extração de DNA plasmidial, cosmidial e fosmidial.....	22
3.8. Desenho dos oligonucleotídeos para prospecção de genes de xilose isomerase.....	23
3.9. "Primers" utilizados e reações de PCR.....	23
3.10. Análise dos produtos de PCR obtidos, clonagem e sequenciamento.....	26

3.11. Subclonagem e sequenciamento do clone fosmidial selecionado.....	28
3.12. Obtenção de DNA dos isolados de <i>Burkholderia</i>	30
3.13. Análises computacionais.....	31
3.14. Ensaio de complementação gênica.....	32
IV. RESULTADOS E DISCUSSÃO.....	34
4.1. Descrição das bibliotecas metagenômicas utilizadas neste trabalho	34
4.2. Escolha dos primers utilizados na prospecção da enzima xilose isomerase.....	34
4.3. Extração do DNA plasmidial, cosmidial e fosmidial das bibliotecas metagenômicas.....	39
4.4. Prospecção de genes nas bibliotecas metagenômicas.....	39
4.5. Identificação de genes de xilose isomerase em isolados de <i>Burkholderias</i>	52
4.6. Caracterização dos genes de xilose isomerase através de análises computacionais.....	54
4.7. Teste de complementação gênica.....	65
4.8. Caracterização molecular dos isolados de <i>Burkholderia</i>	67
V. CONCLUSÕES.....	71
VI. REFERÊNCIAS.....	72

ABREVIATURAS

PPP - via das pentose-fosfato

XI - xilose isomerase

XR - xiluloquinase

XD - xilitol desidrogenase

EC - enzyme code

BAC - "bacterial artificial cromossome"

ORF - fase aberta de leitura

DNA – ácido desoxiribonucleico

DNAse – desoxiribonuclease

EDTA – ácido etilenodiaminotetracético

p/v – peso por volume

pH – $\log [H^+]$

RNA - ácido ribonucléico

RNAse – ribonuclease

SDS – lauril sulfato de sódio

TAE – solução tampão Tris-acetato-EDTA

TBE – solução tampão Tris-ácido bórico-EDTA

TE – solução tampão Tris-EDTA

Tris – Tris[hidroximetil]aminometano

v/v – volume por volume

PCR – reação da polimerase em cadeia

UNIDADES

g – aceleração da gravidade

g – grama

kb – kilobase

L – litro

M – molar

mg – miligrama

mL – mililitro

mM – milimolar

μg – micrograma

μL – microlitro

μM – micromolar

ng – nanogramas

pb – pares de bases

kb - kilobases

U – unidades

V - Volt

PROSPECÇÃO DE GENES COM INTERESSE BIOTECNOLÓGICO

RESUMO - Neste trabalho foi realizada a prospecção de novos genes de interesse biotecnológico em bibliotecas metagenômicas e em linhagens de *Burkholderia* através da técnica de reação em cadeia da polimerase. Nas bibliotecas metagenômicas, foi possível identificar genes que codificam álcool desidrogenase, oxirredutase, acil-CoA desidrogenase, luciferase, transportadores de membrana, thiolase, aminoglicosídeo fosfotransferase, desidrogenase de cadeia curta, proteínas repressoras TetR, enoil-CoA hidratase/isomerase, entre outras. Nas treze linhagens de *Burkholderia* testadas, seis apresentaram amplificação positiva para o gene de xilose isomerase. Estes genes foram completamente sequenciados e as sequências foram utilizadas em análises computacionais, que permitiram estabelecer a identidade entre as sequências e a dedução da função das proteínas baseado em similaridades. Foi realizada a medida do índice de adaptação de códons, com a finalidade de se encontrar um hospedeiro onde a expressão seja maximizada, baseando-se na presença de códons preferenciais e raros. Esta análise mostrou que os genes encontrados possuem boas possibilidades de expressão em *Escherichia coli*, mas que podem apresentar uma taxa de expressão ineficiente em *Saccharomyces cerevisiae*. Foi realizado um teste de complementação gênica utilizando um dos genes descobertos e uma linhagem de *Escherichia coli* que possui o gene de xilose isomerase nocauteado. Os transformantes foram capazes de crescer em meio cuja única fonte de carbono era o açúcar xilose, mostrando que o gene é funcional. Estes genes serão utilizados futuramente em ensaios de expressão para caracterização da enzima.

PALAVRAS-CHAVE: *xylA*, xilose isomerase, metagenoma, *Burkholderia*, xarope de milho rico em frutose, biocombustíveis.

PROSPECTION OF GENES WITH BIOTECHNOLOGICAL INTEREST

SUMMARY - In this work was carried out prospection for new genes of biotechnological interest in metagenomic libraries and strains of *Burkholderia* through the technique of polymerase chain reaction. In metagenomics libraries, we could identify genes encoding alcohol dehydrogenase, oxidoreductase, acyl-CoA dehydrogenase, luciferase, membrane transporters, thiolase, aminoglycoside phosphotransferase, short-chain dehydrogenase, TetR repressor protein, enoyl-CoA hydratase/isomerase, among others. Of the 13 strains of *Burkholderia* tested, 6 showed positive amplification for the xylose isomerase gene. The genes were completely sequenced and the sequences were used in computational analysis, which allowed to establish the identity between the sequences and deducing protein function based on similarities. We evaluated the codon adaptation index, with the aim of finding a host in which the expression is maximized, based on the presence of preferred codons and rare. This analysis showed that the genes found have good possibilities of expression in *Escherichia coli*, but that may have an inefficient rate of expression in *Saccharomyces cerevisiae*. We conducted a genetic complementation test using one of the discovered genes and one strain of *Escherichia coli* that has the xylose isomerase gene knocked out. Transformants were able to grow in a medium whose sole source of carbon was the sugar xylose, showing that the gene is functional. These genes will be used in expression assays for characterization of new enzymes.

KEYWORDS: *xylA*, xylose isomerase, metagenome, *Burkholderia*, High-fructose corn syrup, biofuels.

I. INTRODUÇÃO

A proposta deste trabalho de doutorado foi prospectar genes de interesse biotecnológico, particularmente os envolvidos na conversão de xilose em xilulose. A enzima xilose isomerase catalisa a isomerização reversível da D-xilose e D-glicose para D-xilulose e D-frutose, respectivamente (Fig. 1). Na natureza, a interconversão de xilose em xilulose serve uma exigência nutricional de bactérias saprófitas, que se desenvolvem em material vegetal em decomposição. Esta enzima possui imensa importância biotecnológica, pois é empregada na produção de frutose a partir de amido e na produção de biocombustível.

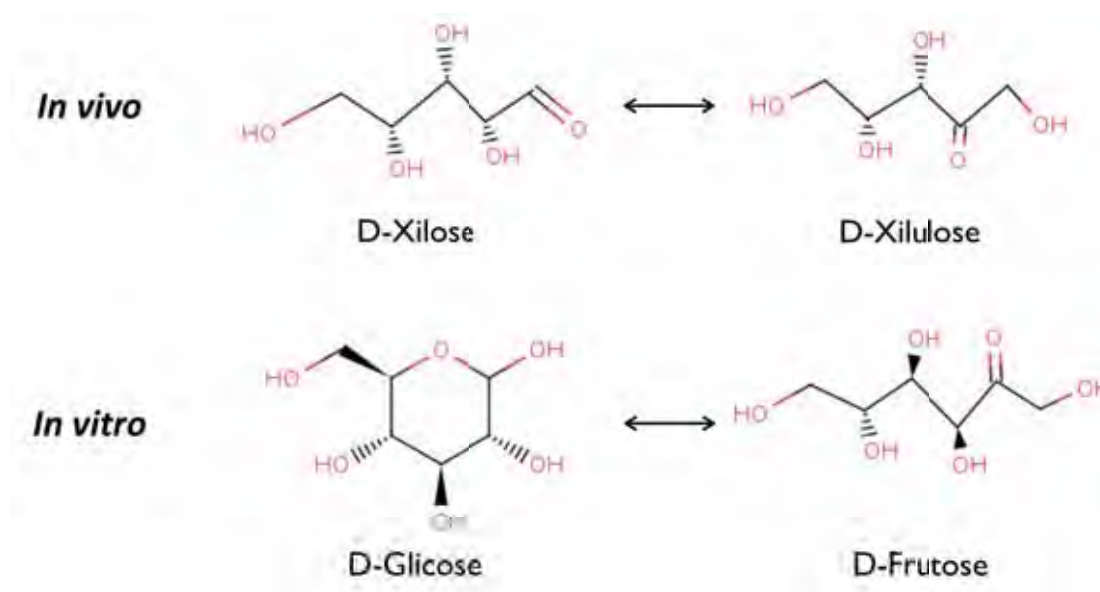


Figura 1. Reação catalisada pela enzima xilose isomerase (EC 5.3.1.5).

Xilose é o segundo carboidrato mais abundante na natureza, e a fermentação comercial deste composto para a produção de etanol pode representar uma alternativa

de aumento da produtividade para o futuro. A levedura selvagem *Saccharomyces cerevisiae* fermenta prontamente glicose, mas não é capaz de metabolizar xilose. Apenas uma pequena parcela de bactérias, leveduras e fungos filamentosos são naturalmente capazes de fermentar xilose. A prospecção dos genes foi realizada em bibliotecas metagenômicas e em isolados de *burkholderias*. Estes genes serão futuramente utilizados em experimentos de engenharia metabólica e engenharia evolutiva, visando obtenção de linhagens capazes de fermentar xilose, para uma conversão mais eficiente da biomassa em etanol.

II - REVISÃO DA LITERATURA

2.1. O emprego da enzima xilose isomerase na produção de biocombustíveis

A descoberta de fontes alternativas de energia tem despertado interesse de cientistas ao redor do planeta em virtude tanto da já prevista escassez de petróleo para as próximas décadas, como também com o aumento da preocupação a respeito da poluição do meio ambiente ocasionada pelos derivados da indústria petrolífera.

A maior contribuição para o aquecimento global é creditada aos altos níveis de emissão de gases “efeito-estufa”, principalmente o dióxido de carbono (CO₂), liberado na queima de combustíveis fósseis. Assim, para mitigar a emissão de CO₂, fontes de energia renováveis, como o etanol, são uma promissora alternativa ao consumo de derivados de petróleo.

O Brasil foi primeira nação a conduzir um programa em larga escala utilizando o etanol como combustível. O Programa Nacional Brasileiro de Etanol (Proálcool) foi originalmente criado em 1975, devido a crise de energia ocorrida em 1973, e propiciou uma redução na dependência do país em combustíveis fósseis. O programa representa uma das maiores aplicações comerciais da biomassa para a produção e utilização de energia no mundo, e demonstrou a viabilidade técnica de produção em larga escala de etanol de cana e seu uso para combustível motores de automóveis (ROVERE et al, 2011). Atualmente, o etanol contribui com 40% do combustível utilizado no país. Assim, o Brasil se tornou o segundo maior produtor de etanol combustível, depois da USA, com 27,5 bilhões de litros produzidos na safra de 2008/2009, segundo dados da União da Indústria de Cana-de-Açúcar (UNICA), a maior organização representativa do setor de açúcar e bioetanol do Brasil.

O aumento da demanda de óleo bruto, que elevou o preço do barril a US\$ 110 em 2008, renovou o interesse na exploração de resíduos de lignocelulose para produção de etanol combustível. Nas décadas passadas, já era consenso que a

produção de etanol a partir da biomassa de plantas deveria ocorrer não somente dos açúcares fermentáveis tais como o amido e a sacarose, mas também partindo de hidrolisados de lignocelulose. Isto porque existe uma grande variedade de matéria prima: bagaço de cana de açúcar, as árvores de crescimento rápido, grama, plantas inteiras, resíduos da produção industrial, plantas aquáticas, entre outros (HAHN-HAGERDAL et al., 2007). A utilização completa do substrato é um dos pré-requisitos para tornar economicamente competitivo o processo de produção de etanol a partir de hemicelulose, cujo constituinte mais abundante é a xilose (tabela 1).

Tabela 1-Composição em polímeros da lignocelulose (INGRAM et al., 1999).

<i>Fração</i>	<i>Conteúdo na lignocelulose</i>	<i>Monômeros principais</i>
<i>Celulose</i>	33-51%	Glicose
<i>Hemicelulose</i>	19-34%	Glicose, manose, galactose Xilose e arabinose
<i>Lignina</i>	20-30%	Alcoóis aromáticos
<i>Pectina</i>	2-20%	Ácido galacturônico e ramnose

Celulose é o principal constituinte da lignocelulose, um polímero linear composto de milhares de subunidade de D-glicoses unidas através de ligações glicosídicas do tipo β -(1→4). A celulose tipicamente ocorre como polímeros paralelos com grande quantidade de pontes de hidrogênio. Esta conformação estrutural, bem como associação com lignina, hemicelulose, amido, proteína e minerais torna a celulose altamente resistente à hidrólise. (VAN MARIS et al., 2006).

Hemicelulose, o segundo maior constituinte da lignocelulose, é um complexo heteropolímero altamente ramificado que contem hexoses (D-glicose, D-galactose, D-manose, L-ramnose e L-fucose), pentoses (D-xilose e D-arabinose) e ácidos urônicos (D-glicurônico e D-galacturônico). A composição deste polímero, entretanto, é

dependente da planta do qual ele se originou. Por outro lado, em contraste com a celulose, a hemicelulose é facilmente hidrolisada em seus monossacarídeos constituintes (VAN MARIS et al., 2006).

Lignina, constituindo de 10-20% da biomassa seca, é um polímero aromático contendo resíduos fenólicos tais como: álcool trans-p-coumaril; trans-p-coniferil e álcool trans-p-sinapil. Além de não contribuir com carbono fermentável este composto também é uma relevante fonte de inibição microbiana (VAN MARIS et al., 2006).

Tabela 2: Composição dos principais açúcares da lignocelulose em restos de vegetais.
(Adaptado de VAN MARIS et al., 2006)

	milho	trigo	bagaço	algodão	beterraba
Carboidratos (%)					
Glicose	34,6	32,6	39,0	37,1	24,1
Manose	0,4	0,3	0,4	1,1	4,6
Galactose	1,0	0,8	0,5	2,4	0,9
Xilose	19,3	19,2	22,1	9,4	18,2
Arabinose	2,5	2,4	2,1	2,3	1,5
Ácidos urônicos	3,2	2,2	2,2	NA	20,7
Não carboidratos (%)					
Lignina	17,7	16,9	23,1	28,8	1,5
Extrativos	7,7	13,0	3,8	7,7	NA
Cinza	10,4	10,2	3,7	10,5	8,2

A levedura selvagem *Saccharomyces cerevisiae* fermenta prontamente glicose, manose e frutose, via Embden-Meyerhof, glicólise, enquanto a D-galactose é fermentada pela ação combinada da via Leloir e glicólise. Entretanto, este microrganismo não fermenta pentoses tais como a xilose.

Junto com a glicose, xilose é um dos principais componentes monossacarídeos da biomassa derivada de plantas (Tabela 2). Para viabilizar economicamente os processos industriais de produção de etanol a partir de hidrolisados lignocelulósicos derivados de biomassa é necessária a fermentação de todos os açúcares presentes. A fermentação da xilose a etanol é um dos maiores desafios tecnológicos do momento.

D-xilose e L-arabinose entram na via das pentose-fosfato (PPP) através da D-xilulose. Basicamente, há duas vias de metabolização da xilose (Fig. 2).

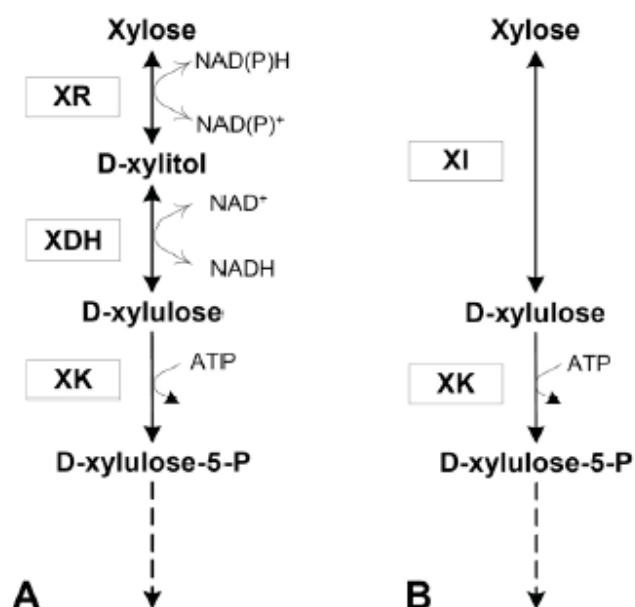


Figura 2. Vias metabólicas para a utilização de xilose. A. via XR-XDH. B. via XI (Karhumaa *et al.*, 2007). XR: xilose redutase. XD: xilitol desidrogenase. XI: xilose isomerase. XK: xilulokinase.

Poucas espécies de bactérias, leveduras e fungos filamentosos são naturalmente capazes de fermentar xilose em etanol. Levando em conta o substrato lignocelulósico e

considerando as estratégias moleculares modernas para o desenvolvimento de linhagens, cada grupo de micro-organismos tem suas vantagens e desvantagens (HAHN-HAGERDAL et al., 2007).

Em algumas linhagens de leveduras, a conexão entre o metabolismo de xilose e a PPP ocorre pela sua conversão em xilulose 5-fosfato. Duas oxiredutases estão envolvidas neste processo, a xilose redutase (XR) e a xilitol desidrogenase (XDH). Essas oxiredutases possuem diferentes especificidades de cofatores, pois a conversão de xilose em xilulose rende um NADP^+ e um NADH . NADPH e NAD^+ precisam ser regenerados, para manutenção do balanço redox (JEFFRIES, 2006). Em condições anaeróbicas, quando não há aceptor de elétrons disponível, a célula não consegue manter o balanço redox e não fermenta a xilose. Mas algumas leveduras, como *Pichia stipitis* e *Pachysolen tannophilus*, são capazes de fermentar xilose sob condições anaeróbicas, pela habilidade de sua XR utilizar tanto o NADH como o NADPH (GOVINDASWAMY & VAME, 2007, KARHUMAA et al., 2007).

A conversão de xilose em xilulose também é realizada pela xilose isomerase (XI), uma enzima que não depende de cofatores e, portanto, não altera o balanço redox da levedura. A clonagem destes genes em leveduras industriais representa uma promissora possibilidade de conversão da xilose. Porém, há informações na literatura de que as XI procarióticas simplesmente não funcionam bem em *S. cerevisiae*. Surpreendentemente, a existência desta enzima foi relatada em uma levedura anaeróbica denominada *Piromyces* sp, que foi isolada a partir de excremento de elefante asiático (BOXMA et al., 2004).

As pesquisas mais recentes se concentram na engenharia metabólica e evolutiva de *S. cerevisiae* visando obtenção de novas linhagens estáveis, com a capacidade de fermentar todos os açúcares derivados da lignocelulose, principalmente xilose, a etanol (Fig. 3). A clonagem de genes de XI ou de XR e XDH é mundialmente considerada de fundamental importância para conduzir a xilose para a PPP (JEFFRIES 2006; KARHUMAA et al., 2007).

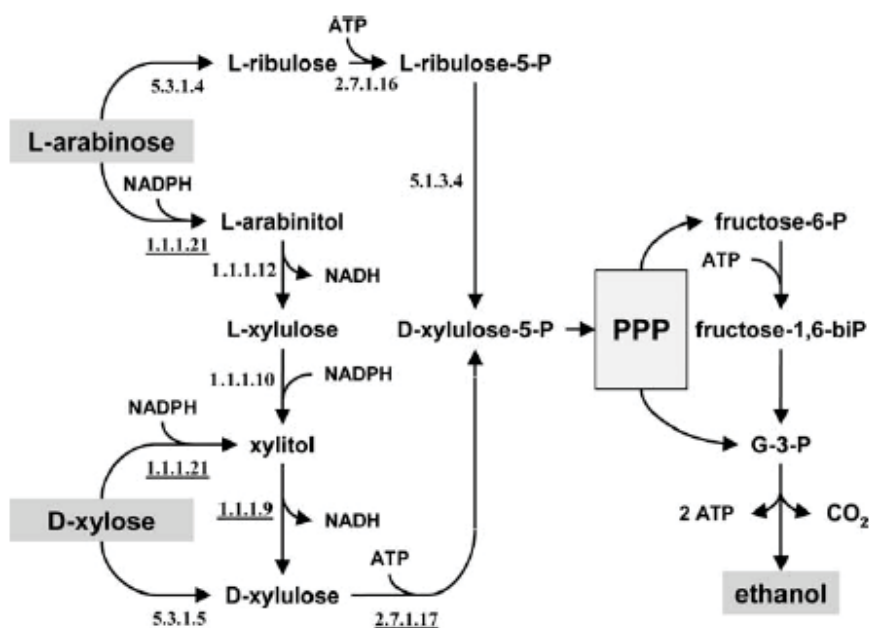


Figura 3. Catabolismo de L-xilose e D-arabinose em linhagens de *S. cerevisiae* metabolicamente engenheiradas. Números de "Enzyme Code" (EC) sublinhados representam enzimas/passos presentes no metabolismo da linhagem selvagem de *S. cerevisiae*. O nome dos genes correspondentes às enzimas são dados nos parênteses da legenda desta figura. 1.1.1.21, aldose/xilose redutase (GRE3/*xy1*); 1.1.1.9, xilitol desidrogenase (XYL2/*xy2*); 2.7.1.17, xilulokinase (XKS1/*xy3*); 5.3.1.5, xilose isomerase (*xy1A*); 1.1.1.12, arabinitol-4-desidrogenase (*lad1*); 1.1.1.10, L-xilulose redutase (*lxr1*); 5.3.1.4, L-arabinose isomerase (*araA*); 2.7.1.16, L-ribulokinase (*araB*); 5.1.3.4, L-ribulose-5-fosfate-4-epimerase (*araD*). G-3-P, gliceraldeido-3-fosfato; PPP, pentose-fosfato "pathway" (VAN MARIS et al., 2006).

A primeira tentativa de clonagem e expressão do gene de xilose isomerase (*xy1A*) em *S. cerevisiae* bem sucedida ocorreu com o gene proveniente da bactéria termófila *Thermus thermophilus*. Xilose isomerase de *T. thermophilus* pôde ser expressa em *S.*

cerevisiae em uma forma ativa, mas a atividade dessa enzima termófila, com uma temperatura ótima de 85 °C foi muito baixa a 30 °C (WALFRIDSSON et al., 1996). Em 2003, KUYPER et al. expressaram com sucesso um gene *xylA* do fungo anaeróbico *Piromyces* sp. em *S. cerevisiae*, obtendo atividade enzimática. No entanto, uma desvantagem desta enzima foi a sua forte inibição por xilitol. Em 2005, VORONOVSKY et al. reportaram a expressão dos genes que codificam xilose isomerases de *Escherichia coli* e *Streptomyces coelicolor* na levedura mutante *Hansenula polymorpha* *Dxyl1*, cujo gene que codifica xilose redutase estava suprimido. A expressão dos genes bacterianos de xilose isomerase restabeleceu a capacidade do mutante *H. polymorpha* *Dxyl1* de crescer em meio com xilose como a única fonte de carbono. Em 2009, BRAT et al. relataram êxito na clonagem e expressão de uma xilose isomerase proveniente da bactéria *Clostridium phytofermentans* com atividade em *S. cerevisiae*. Além disso, esta enzima se mostrou menos sensível à inibição por xilitol do que a enzima proveniente de *Piromyces*. Houve todavia inúmeras tentativas de expressão de genes de *xylA* provenientes de diferentes espécies que não foram bem sucedidas, seja pela impossibilidade da levedura expressar o gene, seja devido à síntese da enzima em uma forma não ativa (SARTHY et al., 1987; AMORE et al., 1989; MOES et al., 1996; BRAT et al., 2009).

2.2. Utilização da enzima xilose isomerase na produção de frutose

Em 1957, MARSHALL e KOOI descobriram que a enzima xilose isomerase proveniente de *Pseudomonas hydrophila* possuía capacidade de isomerização da glicose. Este foi o ponto inicial da exploração desta enzima para a fabricação de frutose a partir de glicose como substituto para o açúcar de cana de açúcar e de beterraba. A isomerização enzimática da glicose foi realizada pela primeira vez em escala industrial em 1967 pela indústria americana Clinton Corn Processing Co. A enzima glicose isomerase tornou-se disponível comercialmente em 1974, e em 1980 praticamente todas as companhias do mundo já utilizavam a tecnologia de isomerização enzimática

da glicose. Hoje, esta enzima comanda o maior mercado na indústria de alimentos e é uma das três que possuem o maior valor por tonelada, sendo amilase e protease as outras duas (BHOSALE et al, 1996).

A principal aplicação desta enzima é na produção de xarope de milho rico em frutose (HFCS - "high fructose corn syrup"). HFCS é uma mistura de equilíbrio de glicose e frutose (1:1), 1,3 vezes mais doce que a sacarose e 1,7 vezes mais doce que a glicose. HFCS é fabricado a partir de amido, e seu preço é 10-20% menor que o da sacarose sobre a base de seu poder adoçante e é preferido pela indústria de alimentos porque não apresenta problema de cristalização como é o caso da sacarose. Além disso, D-frutose desempenha um papel importante como um edulcorante para diabéticos, porque é lentamente reabsorvido pelo estômago e não influencia o nível de glicose no sangue. As principais utilizações do HFCS estão na produção de bebida, panificação, conservas, e nas indústrias de confeitaria. Por ser uma enzima com relevante importância na indústria de alimentos, a maior parte das informações a seu respeito estão documentadas na forma de patentes (BHOSALE et al, 1996).

2.3. Prospecção de genes utilizando a abordagem metagenômica

A descoberta de novas enzimas de conversão de xilose com atividade catalítica em *S. cerevisiae* requerer um escrutínio intensivo de coleções de culturas. Neste sentido, a abordagem metagenômica pode representar uma promissora alternativa para a descoberta de novas enzimas com propriedades ótimas (VAN MARIS et al., 2006).

Até a década de 90, a identificação microbiana requeria o isolamento de culturas puras e sua caracterização era realizada principalmente por testes nutricionais, bioquímicos e fisiológicos. Isso limitou a avaliação taxonômica e filogenética e a estimativa da diversidade microbiana de um ambiente, devido à incapacidade de cultivo da maioria dos micro-organismos pelos métodos convencionais (PACE, 1997).

O acesso molecular individual aos tipos de bactérias existentes na natureza tem sido baseado na determinação das seqüências 16S rRNA amplificadas e clonadas a partir deste DNA obtido do solo. Estes estudos mostraram que a diversidade genética de micro-organismos na natureza é muito maior que a refletida nas coleções e bancos de linhagens, porque apenas uma pequena parcela dos micro-organismos, em torno de 1%, são cultiváveis pelas técnicas padrões de cultivo (AMMAN et al., 1995; BORNEMAN et al., 1996; GESOLMINO et al., 1999; SILVEIRA et al., 2006; PEREIRA et al., 2006). Essas descobertas proporcionaram base científica para o desenvolvimento de novas abordagens de acesso ao potencial metabólico dos micro-organismos do solo sem a necessidade de cultivá-los. A Metagenômica, ou análise genômica da comunidade de micro-organismos de um determinado habitat por técnicas independentes de cultivo, se consolida como uma estratégia interessante e promissora (SCHLOSS & HANDELSMAN, 2003). Essa abordagem consiste na extração de DNA diretamente do ambiente e construção de uma biblioteca de DNA com este genoma misto, para posterior clonagem e expressão de genes de interesse (Fig. 4).

A metodologia de construção de bibliotecas genômicas, que envolve a geração e clonagem de fragmentos de DNA de alto peso molecular em vetores apropriados e a prospecção de genes de interesse, tem sido utilizada exaustivamente por mais de três décadas. A clonagem de DNA metagenômico, entretanto, foi pela primeira vez reportada na metade da década de 90, com a construção de uma biblioteca de metagenoma marinho (STEIN et al., 1996).

A ampla diversidade microbiana do solo constitui-se de um enorme “pool” genético e bioquímico que pode ser explorado para a descoberta de novos genes, vias metabólicas inteiras e seus produtos. A metagenômica oferece um caminho para acessar mais completamente a diversidade microbiana total de uma amostra ambiental. Este acesso permite examinar a genômica dos membros de sua microflora, mesmo sem meios de cultivo de muitos dos organismos presentes.

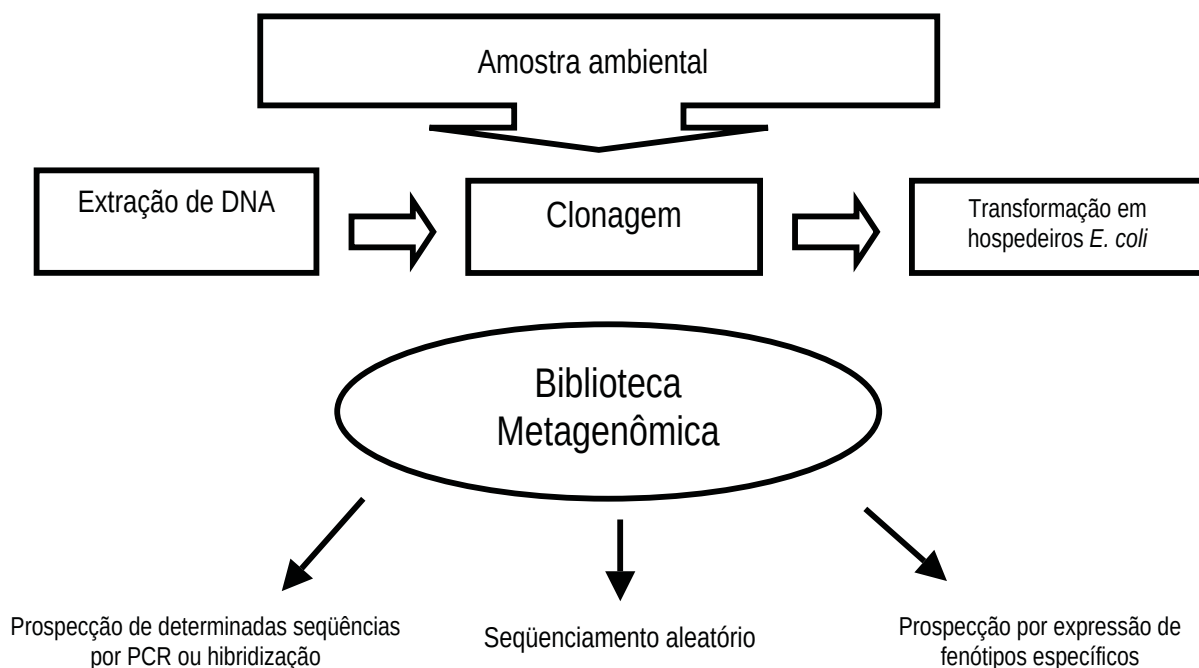


Figura 4. Metagenômica envolve a construção de uma biblioteca de DNA de uma população microbiana de um ambiente e análise das funções e seqüências contidas na biblioteca (adaptado de RIESENFELD et al., 2004).

Essa abordagem tem sido utilizada com sucesso em vários estudos: para análises filogenéticas, para descrição e análise de genes específicos, para determinação de vias biossintéticas complexas ou para a descrição do metabolismo de bactérias desconhecidas. Além disso, essa metodologia pode ser utilizada para estudo de comunidades microbianas contidas em diversas amostras específicas como as que constituem biofilmes ou consórcio de bactérias contidas em culturas de enriquecimento.

Muitos ambientes são focos da metagenômica, incluindo solos, cavidade oral, rumem, habitats aquáticos, entre outros. Em diferentes tipos de amostras, diferentes estratégias podem ser utilizadas para obtenção do DNA metagenômico. Para a

utilização em técnicas de clonagem, é necessário que o DNA extraído seja de alta qualidade, puro e não fragmentado (ZHOU et al., 1996).

Bibliotecas que contem grandes insertos de DNA ambiental são construídas principalmente em vetores cosmídeos (COURTOIS et al., 2003), fosmídios (GINOLHAC et al., 2004) e BAC (RONDON et al., 2000), e permitem identificação e exploração de genes e vias biossintéticas complexas. Vetores que se mantêm na célula em alto número de cópias, e que aceitam insertos pequenos também são utilizados pela facilidade de obtenção de DNA plasmidial suficiente para análises baseadas em sequência (TYSON et al., 2004).

Os elementos genéticos clonados podem ser identificados com base na sequência de DNA, por PCR ou hibridização, ou na função expressada pelo hospedeiro. Os métodos de prospecção podem ser conduzidos em pequena escala, onde cada clone é testado separadamente para uma determinada característica, ou em larga escala, geralmente de forma automatizada, que permite uma identificação muito rápida dos transformantes desejados (HANDELSMAN, 2005). De forma geral, intensivas análises de clones individuais ou de “pools” de clones da biblioteca são necessárias tanto para a prospecção baseada em atividade como em sequência.

2.4. Prospecção de genes em isolados pertencentes ao gênero *Burkholderia*

O gênero *Burkholderia* é atualmente composto por mais de 60 espécies, isoladas de diferentes nichos ecológicos, como o solo, rizosfera, água, plantas, fungos, animais, ambientes hospitalares e humanos infectados (COENYE e VANDAMME, 2003). É um grupo bacteriano intrigante, de grande importância em vários ecossistemas, devido à sua habilidade em degradar poluentes, produzir compostos antimicrobianos, promover o crescimento vegetal, fixar nitrogênio atmosférico e ocupar uma grande variedade de nichos ecológicos (SALLES, 2005).

Este gênero é formado por bactérias Gram-negativas, aeróbias em forma de bastonete pertencentes ao filo das Proteobactérias e a classe das Betaproteobactéria (HOLMES et al., 1998).

A criação do gênero *Burkholderia* foi sugerida em 1992 por YABUUCHI et al., para agrupar o grupo II em homologia do rRNA de *Pseudomonas*, com *Burkholderia cepacia* (antiga *Pseudomonas cepacia*) como espécie padrão (PALLERONI, 1997). Esta classificação foi baseada em características fenotípicas, na composição celular, na sequência do gene 16S rDNA e na homologia do DNA-DNA (YABUUCHI et al., 1992).

O nome deste gênero provém do pesquisador Burkholder, que foi quem observou esta espécie padrão (antiga *Pseudomonas cepacia*) e a descreveu como sendo o agente causal da podridão mole da cebola (BURKHOLDER, 1950).

2.5. Objetivos

O projeto de pesquisa original se apoiava em dois objetivos principais: a construção de um banco de diversidade genética microbiana através da abordagem metagenômica e a prospecção de enzimas de interesse biotecnológico envolvidas na produção de etanol.

A utilização completa do substrato é um dos pré-requisitos para tornar economicamente competitivo o processo de produção de etanol a partir de lignocelulose. Dessa forma, a descoberta de novos genes de xilose isomerase é uma promissora alternativa para aumento da eficiência do processo de produção de etanol. Novos genes envolvidos na produção de etanol serão futuramente usados em experimentos de clonagem, biossíntese combinatória e expressão heteróloga. Os resultados deste trabalho vão fornecer material biológico, conhecimento, capacitação técnica e ferramentas para a manipulação de genes responsáveis pela produção de etanol.

III. MATERIAIS E MÉTODOS

O presente trabalho foi executado no Laboratório de Bioquímica de Micro-organismos e Plantas do Departamento de Tecnologia da Faculdade de Ciências Agrárias e Veterinárias (FCAV) da UNESP de Jaboticabal. Os protocolos, quando não especificados pelo fornecedor dos kits ou reagentes utilizados neste trabalho, foram baseados em SAMBROOK et al. (1989). Tais técnicas não oferecem risco para o meio ambiente e saúde humana, encaixando-se nos padrões de biossegurança (NB-1).

3.1. Coleta das amostras de solo

Vinte amostras simples ao acaso em ziguezague foram coletadas. Para cada uma das áreas, as amostras foram reunidas e homogeneizadas resultando em uma amostra composta. Estas foram imediatamente conduzidas ao Laboratório de Bioquímica de Micro-organismos e Plantas (LBMP), no Departamento de Tecnologia, para extração do DNA metagenômico. A Fig. 5 ilustra o esquema de coletas das amostras de solo.

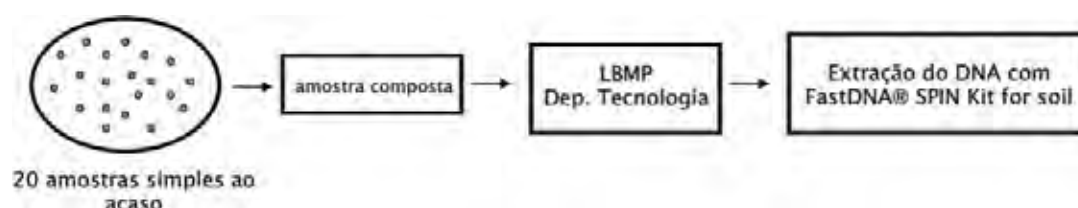


Figura 5. Esquema de coleta das amostras de solo.

3.2. Extração de DNA metagenômico

A extração do DNA das amostras foi realizada com a utilização do FastDNA® SPIN Kit for soil (MP Bio), de acordo com as instruções do fabricante. No final, as amostras foram mantidas a 20°C negativos.

3.3. Quantificação do DNA metagenômico em gel de agarose.

A quantificação foi feita através de eletroforese em gel de agarose 0,8% (p/v), contendo brometo de etídio (0,5 µg/ml). A eletroforese foi realizada com uma alíquota de 3 µL de DNA adicionada de 3 µL de tampão de carregamento [0,025% (p/v) de azul de bromofenol e 50% de glicerol v/v] e conduzida em tampão TEB 1X [Tris 89 mM, Ácido Bórico 89 mM e EDTA 2,5 mM, pH 8,3], a 100 V. A visualização do DNA foi realizada sob luz UV e a imagem foi documentada em um aparelho fotodocumentador (Bio Rad – Gel Doc 1000), através do “software” Quantity OneR (Bio Rad™, Hercules, CA, USA).

3.4. Quantificação do DNA metagenômico no NanoDrop

O aparelho foi ajustado para leitura de fita dupla de DNA, e este foi calibrado com uma amostra de água MILLI-Q filtrada. Posteriormente, foram lidas no NanoDrop 1 µl das soluções contendo o DNA metagenômico

3.5. Construção da biblioteca metagenômica em vetor cosmídeo e fosmídeo

A metodologia para construção das bibliotecas metagenômicas em vetor

cosmídeo e foscídeo é a mesma. A Fig. 6 resume os passos necessários para construção da biblioteca em vetor foscídeo e cosmídeo.

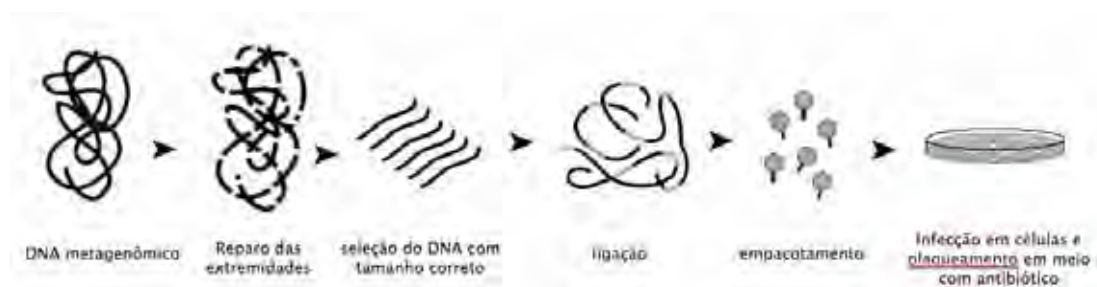


Figura 6. Resumo dos passos necessários para construção da biblioteca em vetor cosmídeo e foscídeo.

Como o vetor estava aberto no sítio de restrição da enzima *Sma*I, cujo corte gera extremidades abruptas, fez-se necessário realizar o reparo das extremidades do inserto. Para cada amostra contendo até 15 µg de DNA metagenômico, em um volume de 52 µl, foram acrescentados 8 µl de 10x concentrado “end repair buffer” (Epicentre), 8 µl de uma mistura de dNTPs, 8 µl de ATP 10 mM e 4 µl de “end repair enzyme” (Epicentre). A reação foi mantida a temperatura ambiente por 45 minutos e em seguida a 75°C, por 20 minutos, para inativação das enzimas. As amostras foram estocadas a 4°C até serem utilizados.

A seleção de tamanho dos insertos reparados de DNA metagenômico foi realizada através de eletroforese preparativa em gel de agarose de baixo ponto de fusão (Sigma) 1% (p/v), isento de brometo de etídio. Para isso, cuba de eletroforese foi previamente descontaminada com hipoclorito 4% (v/v).

Uma alíquota de 20 µL de DNA metagenômico foi aplicada no gel e ao lado 100 ng do marcador de 36 Kb T7 “control” DNA (Epicentre). Pulou-se uma canaleta e na próxima foi aplicado todo o volume da solução de DNA. A eletroforese foi conduzida em

tampão TAE 1X [40 mM Tris-Acetato, 1 mM EDTA] isento de brometo de etídio, a voltagem constante 30 V por aproximadamente 8 horas.

Após a eletroforese descrita, foi realizado um corte vertical no gel de agarose resultando duas porções de gel, uma contendo os insertos a serem recuperados e outra com o marcador e a amostra comparativa para corar. A porção do gel que continha o marcador foi corada com brometo de etídio e observada sob um transiluminador. O padrão de bandejamento serviu como guia para obtenção dos insertos na porção não corada do gel, colocando-se as duas partes do gel lado a lado. A região do gel contendo insertos de tamanho entre 20 e 40 kb foi marcada e cortada com o auxílio de uma lâmina de bisturi, sendo depositada em tubos eppendorfs de 1,5 ml, aproximadamente 400 mg de gel/tubo. Os pedaços de agarose que continham o DNA de tamanho selecionado foram armazenados a 4°C até o momento do uso. O restante do gel foi corado, visualizado em luz UV e documentado em fotodocumentador (Bio Rad – Gel Doc 1000), através do “software” Quantity OneR (Bio Rad™, Hercules, CA, USA).

Os tubos contendo os fragmentos de gel de agarose foram incubados em banho a 65°C durante 15 minutos. Cada amostra teve seu volume igualado a 500µl, completando o volume com tampão TE (10:1) pH 8,0, quando necessário. O tampão da enzima gelase (10 µl) e a enzima gelase (3µl) (Epicentre) foram adicionados aos tubos, que foram agitados e centrifugados rapidamente. As amostras foram incubadas em banho-maria a 45°C por 1 hora. Em seguida, a reação foi transferida para banho-maria a 70°C por 15 minutos para inativação da enzima. Igual volume de acetato de amônio 5 M foi adicionado, e em seguida o material foi centrifugado a 14.000 g por 10 minutos para precipitação do material insolúvel.

Cuidadosamente, foram transferidos 95% (v/v) do sobrenadante para um tubo novo. O DNA foi precipitado pela adição de 2 volumes de etanol absoluto, incubação de 20 minutos em ultrafreezer (-80°C) e recuperado por centrifugação a 14.000 g por 30 minutos a 4°C. Em seguida, foi lavado com 1 ml de etanol 70% (v/v) e centrifugado 14.000 g por 15 minutos a 4°C. O DNA precipitado foi ressuspenso num volume total de 40 µl e quantificado em gel de agarose 1% (p/v), contendo brometo de etídio (0,5

µg/ml). A eletroforese foi conduzida em tampão TBE 1X, a voltagem constante de 90 V por 2 horas.

A reação de ligação dos insertos no vetor foi realizada mantendo uma relação de concentração de insertos de DNA metagenômico com DNA vetor na proporção [1:10]. A reação foi executada em um volume de 20 µl contendo 2 µl de inserto DNA metagenômico ($\pm 0,2$ µg, ± 30 Kb), 1 µl de vetor pCC2FOS, 2 µl de tampão “fast-link ligation buffer” 10X (Epicentre), 1 µl de ATP 10 mM, 1 µl da enzima “fast link DNA ligase” (Epicentre) e 13 µl de H₂O Milli-Q estéril. A reação foi mantida a temperatura ambiente (25°C) por 3 horas e depois a 70°C por 10 minutos para inativação da enzima. Os fosmídios ligados ao inserto foram armazenados a 20°C negativos.

Para transformação das células fago-competentes, um tubo de extrato de empacotamento (Epicentre) com 50 µl foi retirado do freezer 80°C negativos, e colocado no gelo. Em cada tubo contendo os fosmídios ligados ao inserto foi adicionado 25 µl de extrato de empacotamento. As amostras foram agitadas cuidadosamente, centrifugadas rapidamente, e depois incubadas em banho-maria a 30°C por 90 minutos. Passado esse período, foram adicionados mais 25 µl de extrato de empacotamento em cada tubo e as amostras foram novamente incubadas a 30°C por mais 90 minutos. Em seguida foi adicionado 500 µl de tampão de diluição de fagos [10mM Tris-HCl (pH 8,3), 100mM NaCl, 10mM MgCl₂] e 25 µl de clorofórmio. Os tubos foram agitados por inversão e estocados a 4°C.

Colônias isoladas da célula EPI-300 (Epicentre) foram obtidas através do cultivo dos estoques em placas de petri com meio de cultura LB suplementado com 10 mM MgSO₄. No dia anterior a realização da reação de empacotamento, foram inoculados 50 ml de caldo LB suplementado com uma colônia isolada de células EPI-300. O cultivo foi mantido durante a noite a 37°C e agitação de 230 rpm. No dia da reação de empacotamento, 50 ml de caldo LB suplementado foi inoculado com 5 ml do pre-inóculo e cultivado sob a mesma agitação a 37°C até atingir a densidade ótica a 600 nm entre 0,8 e 1,0. A densidade óptica foi determinada no Biophotometer (Eppendorf). As amostras foram mantidas a 4°C até o momento do uso.

Para infecção das células EPI-300, cada 10 µl da solução com os fosmídios empacotados foi misturado, em condições estéreis, a uma alíquota de 100 µL de células do inoculo preparado. Os tubos foram gentilmente agitados e incubados em estufa a 37°C por 20 minutos sem agitação. Após esse período, os 110 µl de cada um dos tubos foi aplicado em placas de petri contendo meio de cultura LB com cloranfenicol (12,5 µg/ml). As placas foram incubadas durante a noite em estufa a 37°C. Como controle positivo, células EPI-300 sem vetor foram cultivadas em meio sem antibiótico, e como controle negativo as células EPI-300 sem vetor foram plaqueadas em meio contendo antibiótico.

Os clones transformados foram repicados com palitos de madeira esterilizados para placas estéreis de poliestireno (96 “well assay plate”, 250 µl) contendo 125 µl de meio Circle Grow (BIO-101) acrescido de 12,5 µg/ml de cloranfenicol. As placas foram incubadas em estufa B.O.D. por 22 horas. Após esse período, foram adicionados 125 µl de solução de glicerol 40% (v/v). Estas placas constituem o estoque de clones das bibliotecas metagenômicas e estão estocadas à 80°C negativos.

3.6. Construção da biblioteca metagenômica em plasmídio.

Esta biblioteca foi construída com DNA metagenômico obtido a partir de amostras de terra de cupinzeiro. Foi realizado o reparo das extremidades do inserto com as enzimas T4 DNA polimerase, Klenow e T4 polinucleotídeo quinase (Fermentas), de acordo com as instruções do fabricante. A reação foi mantida a temperatura ambiente por 45 minutos e em seguida a 75°C, por 20 minutos, para inativação das enzimas. Os tubos foram estocados a 4°C até serem utilizados.

A seleção de tamanho dos insertos reparados de DNA metagenômico foi realizada através de eletroforese preparativa em gel de agarose de baixo ponto de fusão, da mesma forma descrita no item "construção da biblioteca metagenômica em vetor fosmídio". Foram selecionados insertos com tamanho variando entre 5 e 10 kb..

O vetor utilizado foi o pUC19DNA/*Sma*I (Fermentas) (Fig. 7). Foi efetuada ligação com relação de concentração DNA insertos:DNA vetor na proporção [4:1]. A reação foi executada em um volume de 20 μ l contendo 400 ng do DNA selecionado purificado, 100 ng de vetor pUC19DNA/*Sma*I (Fermentas), 2 μ l de tampão ligase [10x] e 20U de T4 DNA ligase (N.E. Biolabs), misturados gentilmente, incubando-se a 16°C durante a noite, aproximadamente 16 h. Como controle, uma reação similar contendo apenas 100 ng do DNA plasmidial do vetor foi realizada nas mesmas condições.

A linhagem DH5 α F⁻ [*supE44*, Δ *lacU169* (ϕ 80 *lacZ* Δ M15), *hsdR17*, *recA1*, *endA1*, *gyrA96*, *thi-1*, *relA1*] sensível a ampicilina, foi utilizada como hospedeira dos vetores recombinantes. Corresponde a uma linhagem supressiva deficiente em recombinação. ϕ 80 *lacZ* Δ M15 permite a α complementação com o N-terminal da β -galactosidase codificada pelo vetor (SAMBROOK, 1989).

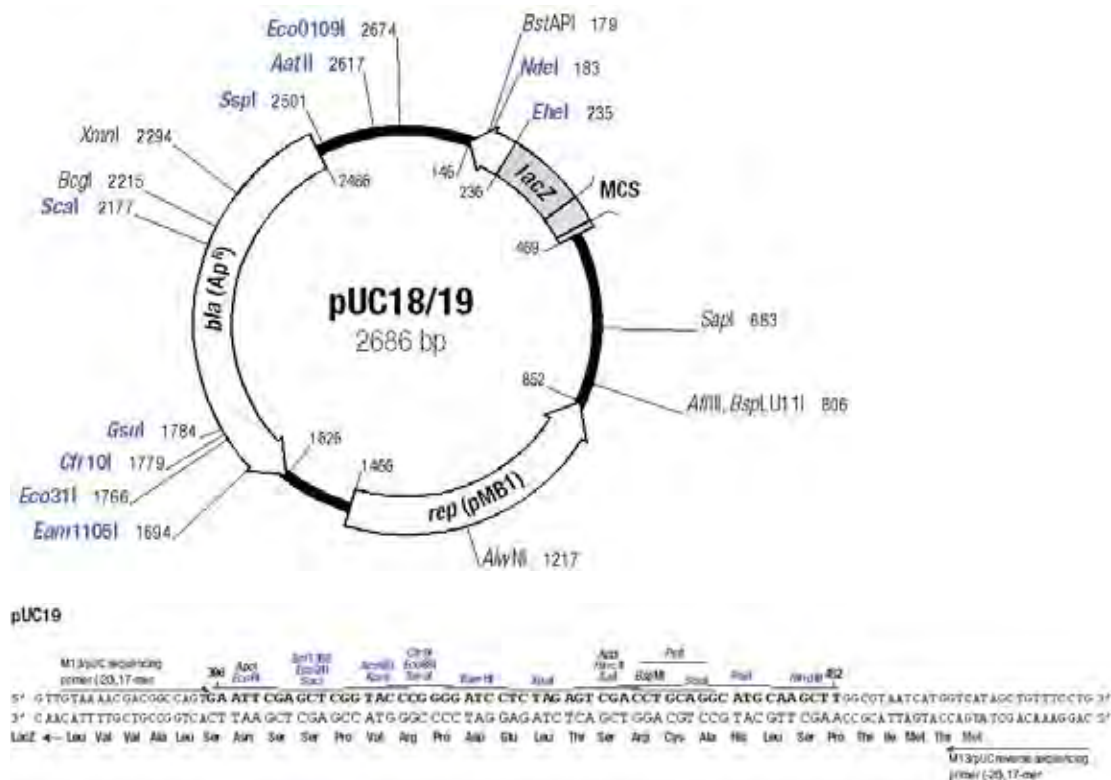


Figura 7. Mapa do vetor pUC19.

A transformação das células competentes foi realizada por choque térmico. Em tubos (falcon) de 15 ml, foram adicionados volumes entre 4 – 20 µl dos produtos das reações de ligação. Nesse momento, as células competentes (*E. coli* DH5α) foram removidas da temperatura -80°C e foram colocadas em banho frio por ±5 minutos. Foram adicionadas 50 µl da suspensão de células aos tubos com os produtos de ligação, misturando-se gentilmente e incubando-se em gelo por 20 minutos. Os tubos foram então imediatamente aquecidos em banho a 42°C, sem agitação, por 90 segundos, para proceder-se o choque térmico. Após esse período, os tubos retornaram rapidamente ao gelo, onde permaneceram por 2 minutos. Em seguida, adicionou-se 950 µl de meio SOC à suspensão, incubando-a a 37°C, com agitação de 220 rpm, durante 1,5 hora. A cultura de células do tubo foi então cultivada em placas preparadas com meio LB sólido (100 µl/placa), acrescido em 50 µg/ml do antibiótico ampicilina, 100 µl/placa do indutor IPTG 100 mM e 50 µl/placa do substrato X-Gal (40 mg/ml em dimetil-formamida).

Após o período de incubação, foram selecionadas as colônias recombinantes, de coloração branca. Os clones transformados foram repicados em placas estéreis de poliestireno (96 *well assay plate*, 250 µl) contendo 125 µl de meio LB contendo 50 µg/ml de ampicilina. As placas foram incubadas em estufa B.O.D. por 24 horas. Após esse período, foram adicionados 125 µl de solução de glicerol 40% (v/v). Estas placas estão estocadas à 80°C negativos.

3.7. Extração de DNA plasmidial, cosmidial e fosmidial

Para extração em placas de 96 clones, os clones da biblioteca foram cultivados em 1,0 ml de meio LB contendo 50 µg/ml de ampicilina, em placas de cultivo de bactérias (*Mega Titer*, 96 *wells*, 1,5 ml). As células foram incubadas com agitação de 280 rpm, a 37°C, por 22 horas. Após a incubação, as amostras foram centrifugadas por 10 minutos, a 3220 x *g*, a 20°C, o sedimento foi ressuspendido em 60µl de solução GTE

[Glicose 50 mM; Tris-HCl 23 mM (pH 8,0); EDTA 10 mM] que foi submetido à extração pelo método de lise alcalina (SAMBROOK *et al.*, 1989). O DNA obtido foi armazenado a 4°C até o momento do uso.

3.8. Desenho dos oligonucleotídeos para prospecção de genes de xilose isomerase

Para desenho dos “primers” de prospecção de genes de xilose isomerase, inicialmente foi realizado um levantamento de todas as sequências de aminoácidos correspondentes a enzima xilose isomerase (EC 5.3.1.5). Este levantamento foi realizado na seção Swiss-Prot do banco de dados UniProtKB (www.uniprot.org). Foram selecionadas 50 sequências de aminoácidos de diferentes espécies. Estas sequências em formato fasta foram alinhadas pelo programa ClustalX (LARKIN *et al.*, 2007). O alinhamento foi utilizado para construção de uma árvore pelo método de distância, utilizando os programas PROTDIST e NEIGHBOR do pacote PHYLIP (FELSENSTEIN, 2005). A árvore foi visualizada utilizando o programa Njplot (PERRIERE & GOUY, 1996). Baseado nas análises filogenéticas realizadas, dois pares de “primers” foram desenhados. Um específico para sequências de genes provenientes de bactérias com alto conteúdo de bases C e G, desenhado com auxílio do programa PRIMER3 (ROZEN & SKALETSKY, 2000), e outro degenerado desenhado com auxílio do programa iCODEHOP (ROSE *et al.*, 1998).

Os demais “primers” utilizados neste trabalho foram desenhados com auxílio do programa PRIMER3 (ROZEN & SKALETSKY, 2000) ou obtidos da literatura.

3.9. "Primers" utilizados e reações de PCR

3.9.1. “Primer” para prospecção de genes de xilose isomerase de bactérias com alto

conteúdo de bases C e G : SxylAF: TTT GGA TCC ATG AAC TAC CAG CCC ACC CCC GA e SxylAR: TAT AAG CTT TCA GCC GCG GGC GCC CAG CAG. A reação de PCR utilizada foi a seguinte: 1x Taq *Buffer*, 0,2 mM dNTP mix, 3mM de MgCl₂, 25pmol de “primer” “forward” e “reverse” 1.25 U de Taq DNA polimerase (Fermentas) e 100 ng de DNA molde, para uma reação com volume final de 50µl. As reações foram conduzidas em termociclador sob as seguintes condições: um passo de desnaturação inicial a 95°C por 5 min, seguido de 1 min a 94°C, 1 min a 64°C e 1 min a 72°C por 35 vezes e um passo de extensão final a 72°C por 15 minutos.

Para a determinação das condições ideais de amplificação, foi utilizado como molde o DNA da linhagem *Streptomyces galilaeus* CCT 4839, fornecida pela Coleção de Culturas Tropical da Fundação André Tosello. A linhagem foi cultivada em meio de cultura apropriado e em seguida foi realizada a extração de DNA genômico. Este foi utilizado para padronização da reação de amplificação. Uma série de reações de PCR foram realizadas, visando à otimização da reação frente aos seguintes parâmetros: temperatura de pareamento do “primer”; concentração de DNA; concentração de Taq DNA polimerase e concentração de cloreto de magnésio. As análises foram feitas por avaliação do perfil eletroforético dos produtos de PCR em géis de agarose 1%.

3.9.2. “Primer” degenerado para prospecção de genes de xilose isomerase de bactérias com baixo conteúdo de bases C e G. B22F: 5'- GGC TAC GAG ACC CTG YTN AAY ACN GA -3' e B38R: 5'- AYC ART AYG AYD ACG ACG TGG CCA C-3'. A reação de PCR utilizada foi a seguinte: 1x Taq *Buffer*, 0,2 mM dNTP mix, 3mM de MgCl₂, 25pmol de “primer” “forward” e “reverse” 1.25 U de Taq DNA Polimerase (Fermentas) e 100 ng de DNA molde, para uma reação com volume final de 50µl. As reações foram conduzidas em termociclador sob as seguintes condições: um passo de desnaturação inicial a 95°C por 5 min, seguido de 1 min a 94°C, 1 min a 66°C e 1 min a 72°C por 30 vezes e um passo de extensão final a 72°C por 15 minutos. Como controle positivo, foi utilizado como molde o DNA da linhagem o microrganismo *Bacillus subtilis* (ATCC 6633). A linhagem foi cultivada em meio de cultura apropriado e em seguida foi

realizada a extração de DNA genômico. Este foi utilizado para padronização da reação de amplificação. Uma série de reações de PCR foram realizadas, visando à otimização da reação frente aos seguintes parâmetros: temperatura de pareamento do “*primer*”; concentração de DNA; concentração de Taq DNA polimerase e concentração de cloreto de magnésio. As análises foram feitas por avaliação do perfil eletroforético dos produtos de PCR em géis de agarose 1%.

3.9.3. “Primer” degenerado para prospecção de genes de xilose isomerase de *Burkholderias*: F-xylA: 5'-GCTGCGCRATTTGCAATTGCCG-3' e R-xylA: 5'-TCAGCGCAVCSCGTAGATCGCCTG-3' (LOPES et al., 2009). A reação de PCR continha tampão 1x concentrado (20mM Tris-HCl (pH 8.4), 50mM KCl), 200 µM de cada desoxinucleotideo, 1.0 mM de MgCl₂, 40 pmols de cada “primer”, 2.5 U Taq DNA polimerase (Invitrogen™, São Paulo, Brazil) e 50 ng de DNA genômico, em um volume final de 50 µL. A reação de amplificação do gene de xilose isomerase foi conduzida em termociclador MJ Research Inc, modelo PTC - 200 e o programa utilizado foi o seguinte: 95 °C por 1 min, 59 °C por 1 min e 72 °C por 3 min, para um total de 30 ciclos. Este “primer” teve as suas condições de funcionamento determinadas utilizando como DNA molde o obtido de *Burkholderia thailandensis* (DSM 13276, ATCC 700388). A linhagem foi cultivada em meio de cultura apropriado e em seguida foi realizada a extração de DNA genômico. Este foi utilizado para padronização da reação de amplificação. Uma série de reações de PCR foram realizadas, visando à otimização da reação frente aos seguintes parâmetros: temperatura de pareamento do “*primer*”; concentração de DNA; concentração de Taq DNA polimerase e concentração de cloreto de magnésio. As análises foram feitas por avaliação do perfil eletroforético dos produtos de PCR em géis de agarose 1%.

3.9.4. “Primer” utilizado para a amplificação do gene 16S rRNA: Pa (5'- AGA GTT TGA TCC TGG CTC AG - 3') e pc5B (5'-TAC CTT GTT ACG ACT T-3'). A reação de PCR continha tampão 1x concentrado (20mM Tris-HCl (pH 8.4), 50mM KCl), 200 µM de cada

desoxinucleotideo, 1.5 mM de $MgCl_2$, 50 pmols de cada "primer", 2.5 U Taq DNA polimerase (InvitrogenTM, São Paulo, Brazil) e 50 ng de DNA genômico, em um volume final de 50 μ L. A reação de amplificação do gene foi conduzida em termociclador MJ Research Inc, modelo PTC - 200 e o programa utilizado foi o seguinte: 94 °C por 2 min, 35 ciclos a 94 °C por 30 seg, 55 °C por 50 seg 72 °C por 2 min e uma extensão final de 72 °C por 5 min.

3.9.5. "Primer" utilizado para a amplificação do gene de xilose isomerase do isolado B3, para expressão em vetor pET28(a) e pGEN -T easy: "primer forward" - B3 (usando a enzima *SacI*) 5' -GCA GAG CtC ATA ACC GGA GAC - 3', "primer reverse" 1 - B3 (usando a enzima *HindIII*) 5'- ATA AGC TTG GTC GGT TCC T - 3', "primer reverse" 2 - B3 (usando a enzima *HindIII*) 5'- ATA AGC TTG CCT GAT TCA CGA C - 3'. A reação de PCR continha tampão 1x concentrado (20mM Tris-HCl (pH 8.4), 50mM KCl), 200 μ M de cada desoxinucleotideo, 1.0 mM de $MgCl_2$, 40 pmols de cada "primer", 2.5 U *Pfu* DNA polimerase (Fermentas) e 50 ng de DNA genômico, em um volume final de 50 μ L. A reação de amplificação do gene de xilose isomerase foi conduzida em termociclador e o programa utilizado foi o seguinte: desnaturação inicial de 3 min a 95 °C, trinta ciclos de 95 °C por 30 seg, 56 °C por 30 seg e 72 °C por 1 min, e um passo de extensão final de 10 min a 72 °C. Para clonagem em vetor pGEN-T, é necessário realizar a adição de um dATP nas extremidades do produto amplificado, que foi realizada da seguinte forma: 7 μ L do produto de PCR amplificado, tampão 1x concentrado (20mM Tris-HCl (pH 8.4), 50mM KCl) com $MgCl_2$, 200 μ M de dATP e 1U de Taq DNA polimerase (InvitrogenTM, São Paulo, Brazil). A reação foi incubada a 70 °C por 30 min.

3.10. Análise dos produtos de PCR obtidos, clonagem e sequenciamento.

Os produtos de PCR obtidos foram purificados com o GFX PCR DNA and gel band purification kit (GE Lifesciences) de acordo com as instruções do fabricante. Os

produtos purificados foram clonados utilizando o pGEM®-T Easy Vector Systems (Promega). O mapa do vetor pGEM-T easy está representado na Fig. 8. Para transformação de células competentes, hospedeiras de vetores recombinantes, foram utilizadas as linhagens DH5 α (genótipo: F⁻ [*supE44*, Δ *lacU169* (ϕ 80 *lacZ* Δ M15), *hsdR17*, *recA1*, *endA1*, *gyrA96*, *thi-1*, *relA1*]), e DH10B (genótipo: F⁻ [*mcrA* Δ (*mrr-hsdRMS-mcrBC*) ϕ 80d *lacZ* Δ M15 Δ *lacX74* *endA1* *recA1* *deoR* Δ (*ara,leu*)7697 *araD139* *galU* *galK* *nupG* *rpsL* λ]), sensíveis a ampicilina. A transformação das células hospedeiras foi realizada por choque térmico. Os clones transformados foram repicados em tubos falcon de 15 ml estéreis contendo 3 ml de meio LB + 100 μ g/ml de ampicilina. Os tubos foram incubados em agitador rotativo por 18 horas a 37°C e 230 rpm. Após esse período, foi extraído o DNA plasmidial pelo método da lise alcalina (SAMBROOK *et al.*, 1989).

Para confirmar a presença de inserto, realizou-se um ensaio de restrição dos plasmídios utilizando a enzima *EcoRI* (Biolabs) conforme a seguinte reação: 300 ng de DNA plasmidial; 0,5 μ l de *EcoRI* 10 U/ μ l (Biolabs); 1,8 μ l de Tampão T4 e 8 μ l de H₂O Milli-Q. Essa preparação foi mantida a 37°C por 2 horas. Os clones que continham insertos foram submetidos ao sequenciamento automático de DNA pelo método de Sanger.

As reações de PCR de seqüenciamento foram realizadas em micro tubos de 0,2 ml, nas seguintes condições: 2 μ l de Dynamic (GE Lifesciences); 10 pmoles dos "primers", M13f ou M13r; 2 μ l de tampão [400 mM Tris-HCl, pH 9; 10 mM MgCl₂]; 80 - 100 ng de DNA; água milli-Q esterilizada para completar o volume final para 10 μ l. As reações foram conduzidas em termociclador seguindo o programa: 1 ciclo a 96°C por 1 min, 25 ciclos (96°C por 10 s, 50°C por 5s e 60°C por 4 min). Os produtos de PCR foram precipitados com isopropanol, secos e ressuspensos em formamida deionizada. A eletroforese foi realizada utilizando um sequenciador automático ABI3100, de acordo com as instruções do fabricante (Applied Biosystems).

Os dados brutos dos eletroferogramas de seqüências de DNA foram analisados pelo programa PhredPhrap (EWING *et al.*, 1998; EWING & GREEN, 1998) e Consed

(GORDON et al., 1998) a fim de gerar sequências formato fasta e avaliar a qualidade e alinhamento semi global das mesmas. Valores de phred acima de 20 foram considerados de alta qualidade. As sequências foram submetidas ao GenBank do National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/BLAST>), para comparação com sequências depositadas no banco de dados através da ferramenta Blastn.

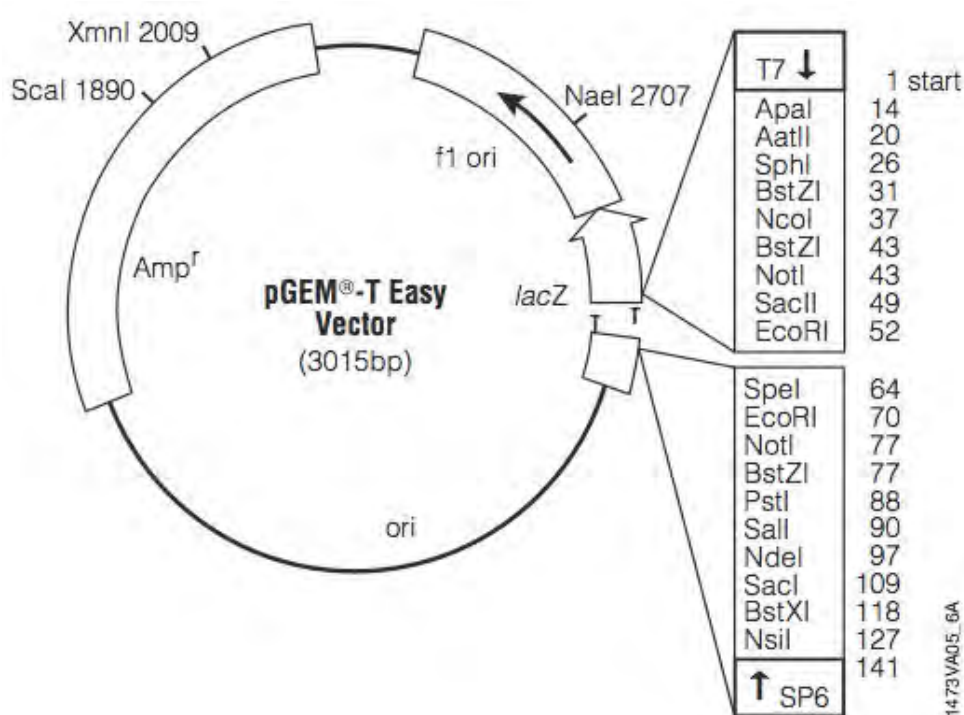


Figura 8. Mapa do vetor pGEM-T easy.

3.11. Subclonagem e sequenciamento do clone fosmidial selecionado

O DNA fosmidial do clone selecionado foi submetido à fragmentação mecânica por nebulização. Após a fragmentação do DNA, a reação de reparo das extremidades

dos insertos foi realizada para que este pudesse ser ligado ao vetor. Foram usadas as enzimas T4 DNA polimerase, Klenow e T4 polinucleotídeo kinase (Fermentas), de acordo com as instruções do fabricante.

Para obtenção dos insertos de tamanho selecionado, foi realizada uma eletroforese preparativa em gel de agarose de baixo ponto de fusão (Sigma) 1% (p/v) e a recuperação dos fragmentos de DNA do gel de agarose foi realizada com a utilização da enzima Gelase (Epicentre), de acordo com a descrição do fabricante.

O DNA obtido foi analisado em gel de agarose 1,0% (p/v) em tampão TBE 1X, adicionado de brometo de etídio (0,5 µg/ml), com a finalidade de quantificação e visualização da qualidade deste DNA recuperado. A quantificação foi realizada com o uso do software Quantity One (Bio Rad™, Hercules, CA, USA). O DNA obtido também foi analisado em espectrofotômetro NanoDrop ND-100 para a determinação da concentração e pureza da amostra.

O vetor utilizado foi o pUC19DNA/*Sma*I (Fermentas) (Fig. 6). Foi efetuada ligação com relação de concentração DNA insertos:DNA vetor na proporção [4:1]. A reação foi executada em um volume de 20 µl contendo 400 ng do DNA purificado, 100 ng de vetor pUC19DNA/*Sma*I (Fermentas), 2 µl de tampão ligase [10x] e 20U de T4 DNA ligase (N.E. Biolabs), misturados gentilmente, incubando-se a 16°C durante a noite, aproximadamente 16 h. Como controle, uma reação similar contendo apenas 100 ng do DNA plasmidial do vetor foi realizada nas mesmas condições.

Para transformação de células competentes, hospedeiras de vetores recombinantes, foi utilizada a linhagem DH10B, (genótipo: F⁻ [*mcrA* Δ (*mrr-hsdRMS-mcrBC*) ϕ 80d *lacZ* Δ M15 Δ *lacX74* *endA1* *recA1* *deoR* Δ (*ara,leu*)7697 *araD139* *galU* *galK* *nupG* *rpsL* λ]) sensível a ampicilina. A transformação das células hospedeiras foi realizada por choque térmico. Os clones transformados foram repicados em tubos falcon de 15 ml estéreis contendo 3 ml de meio LB + 100 µg/ml de ampicilina. Os tubos foram incubados em agitador rotativo por 18 horas a 37°C e 230 rpm. Após esse período, foi extraído o DNA plasmidial pelo método da lise alcalina (SAMBROOK *et al.*,

1989).

As reações de PCR de seqüenciamento foram realizadas em micro tubos de 0,2 ml, nas seguintes condições: 2 µl de Dynamic (GE Lifesciences); 10 pmoles dos “primers”, M13f ou M13r; 2µl de tampão [400 mM Tris-HCl, pH 9; 10 mM MgCl₂]; 80 - 100 ng de DNA; água milli-Q esterilizada para completar o volume final para 10 µl. As reações foram conduzidas em termociclador seguindo o programa: 1 ciclo a 96°C por 1 min, 25 ciclos (96°C por 10 s, 50°C por 5s e 60°C por 4 min). Os produtos de PCR foram precipitados com isopropanol, secos e ressuspensos em formamida deionizada. A eletroforese foi realizada utilizando um sequenciador automático ABI3100, de acordo com as instruções do fabricante (Applied Biosystems).

Os dados brutos dos eletroferogramas de seqüências de DNA foram analisados pelo programa PhredPhrap (EWING *et al.*, 1998; EWING & GREEN, 1998) e Consed (GORDON *et al.*, 1998) a fim de gerar seqüências formato fasta e avaliar a qualidade e alinhamento das mesmas. Valores de phred acima de 20 foram considerados de alta qualidade. As seqüências foram submetidas ao *GenBank* do *National Center for Biotechnology Information* (NCBI) (<http://www.ncbi.nlm.nih.gov/BLAST>), para comparação com seqüências homólogas depositadas no banco de dados através da ferramenta Blastn.

3.12. Obtenção de DNA dos isolados de *Burkholderia*

Os isolados de *Burkholderia* utilizados neste trabalho são oriundos da coleção de micro-organismos do Laboratório de Bioquímica de Micro-organismos e Plantas, da Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal - UNESP, e foram isolados pela equipe do LBMP. A extração de DNA genômico foi realizada segundo MARMUR, 1961. A quantificação do DNA foi realizada por espectrofotometria utilizando aparelho NanoDrop 1000 (Thermo Fisher Scientific).

A visualização dos amplicons foi feita através de eletroforese em gel de agarose

1,0 % (p/v), contendo brometo de etídio (0,5 µg/ml). A eletroforese foi realizada com uma alíquota de 5 µL da reação de amplificação adicionada de 1 µL de tampão de carregamento [0,025% (p/v) de azul de bromofenol e 50% de glicerol v/v] e conduzida em tampão TEB 1X [Tris 89 mM, Ácido Bórico 89 mM e EDTA 2,5 mM, pH 8,3], a 80 V. A visualização da eletroforese foi realizada sob luz UV e a imagem foi documentada em um aparelho fotodocumentador (Bio Rad – Gel Doc 1000), através do “software” Quantity OneR (Bio Rad™, Hercules, CA, USA).

3.13. Análises computacionais

As fases abertas de leitura ("open reading frames" - ORFs) foram determinadas utilizando o programa ORFfinder (NCBI). A anotação dos genes foi realizada a partir da sequência de aminoácidos das ORFs encontradas, através da procura por similaridades com proteínas presentes em diferentes bancos de dados, utilizando a ferramenta de alinhamento local Blastp.

As análises filogenéticas foram realizadas utilizando os programas do pacote PHYLIP, versão 3.6 (FELSENSTEIN, 2005). Para a análise da sequência de aminoácidos dos genes de xilose isomerase identificados, foi realizado um levantamento de sequências correspondentes a enzima xilose isomerase (EC 5.3.1.5) presentes nos bancos de dados públicos. Estas sequências em formato fasta foram alinhadas pelo programa ClustalX (LARKIN *et al*, 2007). O alinhamento foi utilizado para construção de uma árvore filogenética pelo método de distância, utilizando os programas PROTDIST e NEIGHBOR. Os 1000 conjuntos de dados foram gerados utilizando o programa SEQBOOT. Uma árvore consenso foi construída para cada replicata e então os "bootstraps" foram computados com o programa CONSENSE. As árvores obtidas foram visualizadas através do programa NJplot (PERRIERE & GOUY, 1996).

Para a análise das sequências de nucleotídeos dos genes de xilose isomerase e

16S rRNA, as sequências dos genes foram alinhadas pelo programa ClustalX (LARKIN *et al.*, 2007). O alinhamento foi manualmente editado quando necessário utilizando o programa eBioX. O alinhamento foi utilizado para construção de uma árvore filogenética pelo método de distância, utilizando os programas DNADIST e NEIGHBOR. Os 1000 conjuntos de dados foram gerados utilizando o programa SEQBOOT. Uma árvore foi construída para cada replicata e então os "bootstraps" foram computados com o programa CONSENSE. A árvore foi visualizada utilizando o programa Njplot (PERRIERE & GOUY, 1996).

Para a determinação do índice de adaptação de códons, foi utilizada a ferramenta CodonW, disponível em http://mobyli.pasteur.fr/cgi-bin/MobyliPortal/portal.py?form_codonw (RICE *et al.*, 2000).

3.14. Ensaio de complementação gênica

O gene de xilose isomerase do isolado B3 foi amplificado utilizando a enzima *Pfu* DNA polimerase. Os produtos de ligação foram clonados utilizando o pGEM®-T Easy Vector Systems (Promega) (Fig. 7), de acordo com as instruções do fabricante. Para transformação de células competentes, hospedeiras de vetores recombinantes, foi utilizada a linhagem hospedeira *E. coli* HB101, que possui o gene de xilose isomerase nocauteado (genótipo: *supE44*, $\Delta(mcrC-mrr)$, *recA13*, *ara-14*, *proA2*, *lacY1*, *galk2*, *rpsL20*, *xyl-5*, *mtl-1*, *leuB6*, *thi-1*). Os transformantes foram crescidos em meio BHB (Bushnell-Hass) ($MgSO_4$, 0,2g/l, KH_2PO_4 1,0 g/l, $CaCl_2$ 0,02g/l, $(NH_4)_2HPO_4$ 1,0g/l, KNO_3 1,0 g/l, $FeCl_3$ 0,05g/l, Bactoagar 9,0 g/l), suplementado com 1% de D-xilose, 50 μ g/ml do antibiótico ampicilina, 100 μ l/placa do indutor IPTG [100 mM] e 50 μ l/placa do substrato X-Gal (40 mg/ml em dimetil-formamida). Os transformantes foram incubados a 37°C por um período de 4 dias. Após esse período, foi extraído o DNA plasmidial pelo método da lise alcalina (SAMBROOK *et al.*, 1989).

Para confirmar a presença de inserto, realizou-se um ensaio de restrição dos

plasmídios utilizando a enzima *EcoR* I (Biolabs) conforme a seguinte reação: 300 ng de DNA plasmidial; 0,5 µl de *EcoR* I 10 U/ µl (Biolabs); 1,8 µl de Tampão T4 e 8 µl de H₂O Milli-Q. Essa preparação foi mantida a 37°C por 2 horas. Os clones que continham insertos foram submetidos ao sequenciamento automático de DNA.

As reações de PCR de seqüenciamento foram realizadas em micro tubos de 0,2 ml, nas seguintes condições: 2 µl de Dynamic (GE Lifesciences); 10 pmoles dos "primers", T7 ou SP6; 2µl de tampão [400 mM Tris-HCl, pH 9; 10 mM MgCl₂]; 80 - 100 ng de DNA; água milli-Q esterilizada para completar o volume final para 10 µl. As reações foram conduzidas em termociclador seguindo o programa: 1 ciclo a 96°C por 1 min, 25 ciclos (96°C por 10 s, 50°C por 5s e 60°C por 4 min). Os produtos de PCR foram precipitados com isopropanol, secos e ressuspensos em formamida deionizada. A eletroforese foi realizada utilizando um sequenciador automático ABI3100, de acordo com as instruções do fabricante (Applied Biosystems).

Os dados brutos dos eletroferogramas de seqüências de DNA foram analisados pelo programa PhredPhrap (EWING *et al.*, 1998; EWING & GREEN, 1998) e Consed (GORDON *et al.*, 1998) a fim de gerar seqüências formato fasta e avaliar a qualidade e alinhamento das mesmas. Valores de phred acima de 20 foram considerados de alta qualidade. As seqüências foram submetidas ao GenBank do "National Center for Biotechnology Information" (NCBI) (<http://www.ncbi.nlm.nih.gov/BLAST>), para comparação com seqüências homólogas depositadas no banco de dados através da ferramenta Blastn.

IV. RESULTADOS E DISCUSSÃO

4.1. Descrição das bibliotecas metagenômicas utilizadas neste trabalho

Para execução deste trabalho foram utilizadas as bibliotecas que compõem o banco de DNA metagenômico do Laboratório de Bioquímica de Micro-organismos e Plantas (LBMP), e foram construídas junto com a equipe de metagenôma do LBMP. O banco de DNA metagenômico é consistido por: uma biblioteca metagenômica de solo sob arboreto de eucaliptos, construída em vetor cosmídio, que possui 9320 clones e insertos de tamanho variando entre 20 e 40 kb, uma biblioteca metagenômica em vetor fosmídeo, construída a partir de DNA obtido de um consórcio bacteriano degradador de hidrocarbonetos, que possui 4.250 clones e insertos de tamanho médio de 30kb; e uma biblioteca metagenômica construída em vetor plasmídeo, utilizando DNA obtido a partir de solo de cupinzeiro, que possui 3.600 clones e insertos com tamanho variando entre 6 e 10 kb.

4.2. Escolha dos “primers” utilizados na prospecção da enzima xilose isomerase

Inicialmente, foi realizado um levantamento de todas as sequências de aminoácidos correspondentes a enzima xilose isomerase ("Enzyme Code" 5.3.1.5). Este levantamento foi realizado na seção Swiss-Prot do banco de dados UniProtKB (www.uniprot.org). Este banco de dados foi escolhido por conter apenas sequências anotadas e revistas manualmente, sendo assim de alta confiabilidade.

Foram selecionadas 50 sequências de aminoácidos de diferentes espécies. Estas sequências em formato fasta foram alinhadas pelo programa ClustalX, e o alinhamento foi utilizado para construção de uma árvore pelo método de distância. A árvore obtida está representada na Fig. 9.

A análise filogenética permitiu identificar dois grandes grupos de enzimas. O primeiro grupo (grupo I) está representado por bactérias cujo genoma contém baixo conteúdo de bases C e G, e no segundo (grupo II) estão representados micro-organismos com alto conteúdo de bases C e G. De forma geral, cada um dos dois grandes grupos identificados contém pelo menos um representante cujo gene *xylA* foi clonado e expressado em leveduras. Baseado na análise filogenética realizada, definiu-se a estratégia de desenhar “primers” para regiões conservadas de cada um dos grupos evidenciados na árvore construída.

Outros dois “primers” foram escolhidos. Um foi retirado do artigo de LOPES *et al*, 2009, é degenerado e baseado no alinhamento de sequências do gene *xylA* de *Burkholderia*. O outro foi desenhado com base na sequência do gene de *xylA* da levedura *Piromyces*, utilizando o software Primer3.

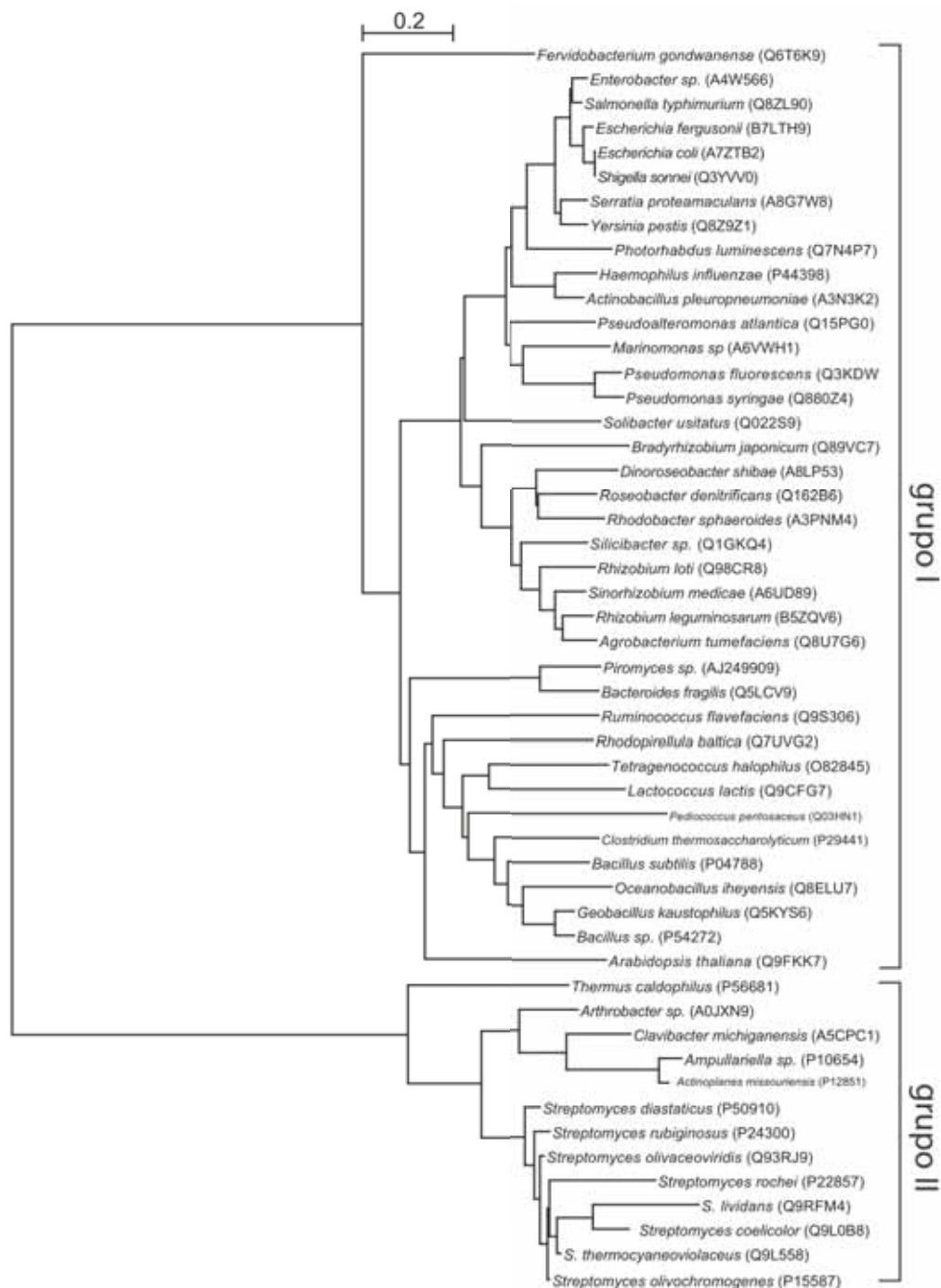


Figura 9. Árvore filogenética baseada na matriz de distância elaborada a partir do alinhamento de sequências da proteína xilose isomerase. Os números de acesso estão em parênteses ao lado do nome das linhagens.

Com o primer desenhado para bactérias de alto conteúdo de bases C e G foram realizadas reações de PCR utilizando como DNA molde os mesmos DNAs utilizados para construção das bibliotecas metagenômica de solo de eucalipto e de terra de cupinzeiro. Foi obtida uma amplificação positiva com o DNA de terra de cupinzeiro.

Os amplicons obtidos foram purificados e clonados em vetor pGEN-T easy. Foram coletados 96 clones. Uma análise de restrição destes clones com a enzima *EcoR* I mostrou que 32 apresentavam insertos com tamanho compatível ao que foi clonado. Estes 32 clones foram submetidos ao sequenciamento de DNA utilizando *primers foward e reverse*.

Os dados brutos dos eletroferogramas de seqüências de DNA foram analisados pelo programa PhredPhrap e as seqüências foram submetidas ao programa BlastX, para comparação com seqüências de proteínas homólogas depositadas no banco de dados. Das 64 seqüências com os *primers foward e reverse* obtidas, 17 apresentaram homologia com seqüências de xilose isomerase. Estas foram submetidas ao programa BlastN, cujos resultados estão descritos na tabela 3.

Tabela 3: Descrição da homologia obtida para cada um dos amplicons dos genes de *xylA* sequenciados.

Produto de PCR	Description	Accession	Max score	Total score	Query coverage	E value	Max ident
H6	<i>Streptomyces</i> sp. SK strain <i>xylA</i> gene, strain SK	Y15518.2	246	246	83%	4e-62	86%
D6	<i>Streptomyces chibaensis</i> D-xylose isomerase (<i>xylA</i>) gene	AF170068.1	377	377	88%	2e-101	88%
G6	<i>Streptomyces avermitilis</i> MA-4680 DNA, complete genome	BA000030.3	388	388	99%	6e-105	89%
D2	<i>Streptomyces avermitilis</i> MA-4680 DNA, complete genome	BA000030.3	521	521	100%	7e-145	94%
E2	<i>Streptomyces avermitilis</i> MA-4680 DNA, complete genome	BA000030.3	370	370	94%	2e-99	91%
A6	<i>Streptomyces avermitilis</i> MA-4680 DNA, complete genome	BA000030.3	486	486	97%	3e-134	90%
B8	<i>Streptomyces avermitilis</i> MA-4680 DNA, complete genome	BA000030.3	483	483	94%	4e-133	89%
F8	<i>Streptomyces violaceoniger xylA</i> gene	X12816.1	390	390	89%	2e-105	88%
C5	<i>Streptomyces violaceoniger xylA</i> gene	X12816.1	424	424	92%	2e-115	90%
C4	<i>Streptomyces violaceoniger xylA</i> gene	X12816.1	457	457	99%	2e-125	88%
C3	<i>Streptomyces violaceoniger xylA</i> gene	X12816.1	457	457	99%	2e-125	88%
A4	<i>Streptomyces violaceoniger xylA</i> gene	X12816.1	449	449	100%	4e-123	88%
B4	<i>Streptomyces violaceoniger xylA</i> gene	X12816.1	549	549	99%	4e-153	91%
E6	<i>Streptomyces rubiginosus</i> xylose isomerase gene	M73789.1	67.6	67.6	13%	3e-08	97%
D4	<i>Streptomyces rubiginosus</i> xylose isomerase gene	M73789.1	350	350	100%	3e-93	87%
C8	<i>Streptomyces rubiginosus</i> xylose isomerase gene	M73789.1	375	375	81%	6e-101	88%
A8	<i>Streptomyces rubiginosus</i> xylose isomerase gene	M73789.1	411	411	89%	2e-111	87%

4.3. Extração do DNA plasmidial, cosmidial e fosmidial das bibliotecas metagenômicas

O DNA plasmidial de todos os clones da biblioteca metagenômica de solo de cupinzeiro foi extraído em placas de 96 poços e quantificado em gel de agarose e em nanodrop. Foi considerado adequado o DNA que possuía concentração acima de 20ng/μl e relação de absorbância a 260 nm e 280 nm entre 1,7 e 2,0. Depois que todos os “pools” de DNA foram obtidos e quantificados, iniciou-se a prospecção dos genes de xilose isomerase utilizando os diferentes “primers” elegidos neste trabalho.

4.4. Prospecção de genes de xilose isomerase nas bibliotecas metagenômicas

Depois que todos os “pools” de DNA foram obtidos e quantificados, iniciou-se a prospecção dos genes de xilose isomerase utilizando os diferentes “primers” elegidos neste trabalho. O primeiro conjunto de “primers” testados foi para bactérias de alto conteúdo de bases C e G.

4.4.1. Prospecção de genes na biblioteca metagenômica de solo de cupinzeiro e o conjunto de “primers” bactérias de alto conteúdo de bases C e G.

A Fig. 10 ilustra a análise da amplificação por PCR utilizando como molde DNA plasmidial de “pools” de 12 clones da biblioteca metagenômica de solo de cupinzeiro. De todos os “pools” testados, 8 apresentaram sinal de amplificação positiva.

Uma nova reação de PCR foi realizada nos clones individuais das placas cujos “pools” apresentaram amplificação positiva e um clone isolado foi obtido, conforme ilustrado na Fig. 11. Foi realizada a purificação e o sequenciamento do amplicom, e a análise da sequência obtida utilizando as ferramentas de bioinformática.

O produto de PCR possui aproximadamente 429 pb. A sequência obtida foi comparada com as disponíveis no banco de dados Genbank, através das ferramentas Blastn e Blastx. O produto de PCR apresenta similaridade com proteínas da família DoxX, que são proteínas de função desconhecida.

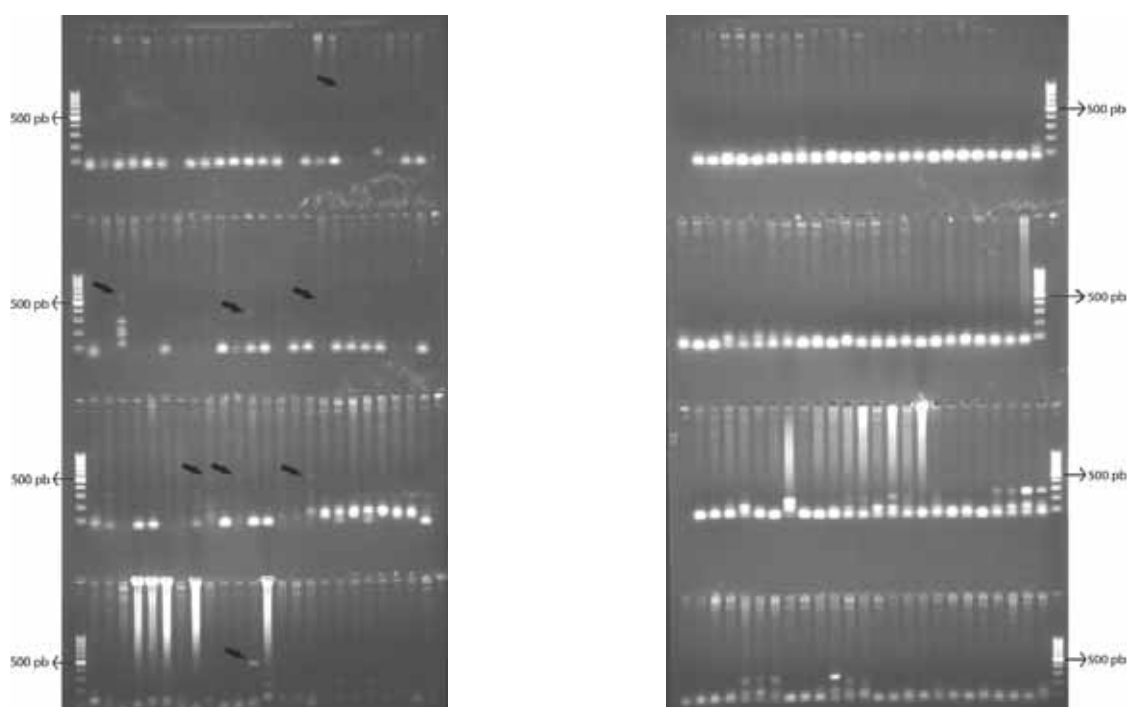


Figura 10. Perfil eletroforético em gel de agarose 1% (p/v) da amplificação por PCR utilizando como molde DNA plasmidial da biblioteca metagenômica de terra de cupinzeiro em “pools” de 12 clones. Marcador utilizado: 100 bp DNA “ladder” (Fermentas). As setas indicam sinal de amplificação positiva.



Figura 11. Perfil eletroforético em gel de agarose 1,5% (p/v) da amplificação por PCR utilizando como molde DNA plasmidial dos clones da fileira E da placa 16 da biblioteca metagenômica de terra de cupinzeiro. Em destaque o amplicom obtido, com tamanho de 429 pb. Marcador utilizado: 100 bp DNA “ladder” (Fermentas).

Visto isso, optou-se por realizar o sequenciamento das extremidades do inserto contido no plasmídio do clone E8 da placa 16. A sequência obtida foi comparada com as disponíveis no banco de dados Genbank, através das ferramentas Blastn. A sequência de uma das extremidades apresentou similaridade com a enzima 3-phosphoshikimate 1-carboxyvinyltransferase. Esta enzima catalisa a reação entre o shikimato-3-fosfato (S3P) e fosfoenolpiruvato (PEP) para formar 5-enolpiruvishikimato-3-fosfato (EPSP), um intermediário na via do shikimato levando a biossíntese dos aminoácidos aromáticos. A reação é fosfoenolpiruvato + 3-fosfoshikimato = fosfato + 5-O-(1-carboxivinil)-3-fosfoshikimato. Pode ser encontrada em bactérias e plantas, mas não animais. A enzima é o alvo do amplamente utilizado herbicida glifosato, que foi mostrado para ocupar o sítio ativo. A sequência da outra extremidade apresentou baixa similaridade com várias proteínas diferentes.

4.4.2. Prospecção de genes na biblioteca metagenômica de solo de eucaliptos e o conjunto de “primers” bactérias de alto conteúdo de bases C e G.

Um clone da biblioteca metagenômica de solo de eucaliptos, o clone E3P16,

apresentou um forte sinal de amplificação positiva, conforme ilustrado na figura 12.

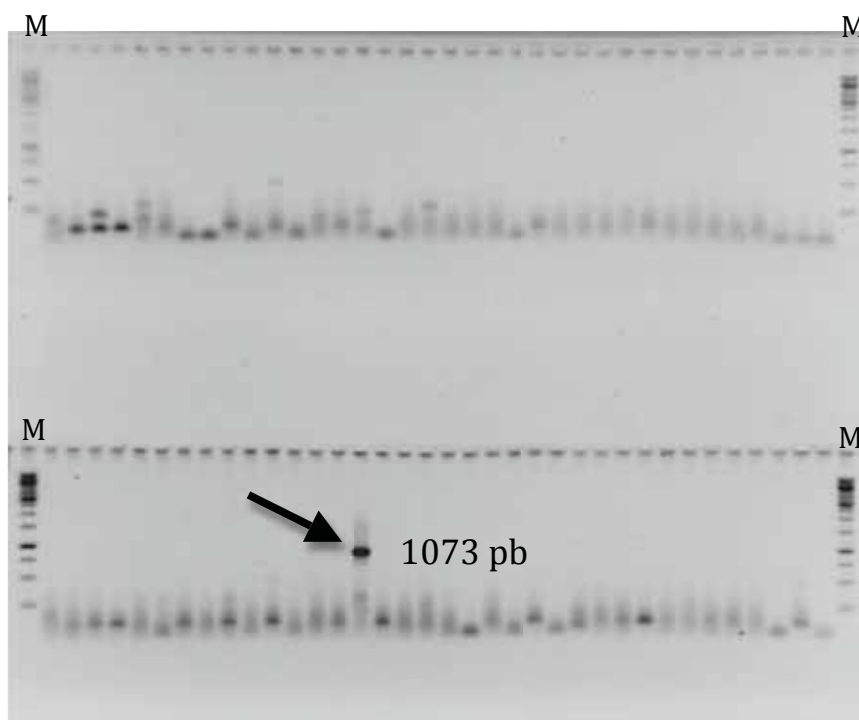


Figura 12: Perfil eletroforético em gel de agarose 1% do DNA cosmidial de clones da placa 16 da biblioteca metagenômica de solo de eucaliptos. M. Marcador 1 kb DNA “ladder” (Fermentas).

Foi realizada a purificação, clonagem em vetor pGEN-T easy e sequenciamento do amplicom, e a análise da sequência obtida utilizando as ferramentas de bioinformática. O produto de PCR possui 1073 bases. A sequência obtida foi comparada com as disponíveis no banco de dados Genbank, através da ferramenta Blastx. O produto de PCR apresenta 72% de similaridade e 55% de identidade com proteínas da família DNA ligase D, do organismo *Anaeromyxobacter* sp.

DNA ligase D (LigD) é uma grande proteína polifuncional que participa de um caminho recentemente descoberto de ligação de extremidades não homólogas (“Non-

homologous end joining - NHEJ") em bactérias. LigD consiste em um domínio ligase ATP-dependente fundido a um domínio da polimerase (Pol) e um módulo fosfoesterase. A atividade Pol é notável pela sua dependência de manganês, sua capacidade de executar reações de extensão com ou sem molde, e sua preferência para a adição de ribonucleotídeos na extremidade do DNA (Zhu et al, 2006)

4.4.3. Prospecção de genes na biblioteca de DNA do consórcio degradador de hidrocarbonetos e o conjunto de "primers" bactérias de alto conteúdo de bases C e G.

Na biblioteca do consórcio degradador de hidrocarbonetos, o "pool" correspondente a placa 14 apresentou sinal de amplificação positiva. Foi realizada uma PCR utilizando como molde os clones individuais desta placa, e dois apresentaram sinal de amplificação positiva: o clone E2PL14 e o clone A9PL14 (Fig. 13). Foi realizada a purificação, clonagem, sequenciamento dos amplicons e análise das sequências obtidas utilizando as ferramentas de bioinformática.

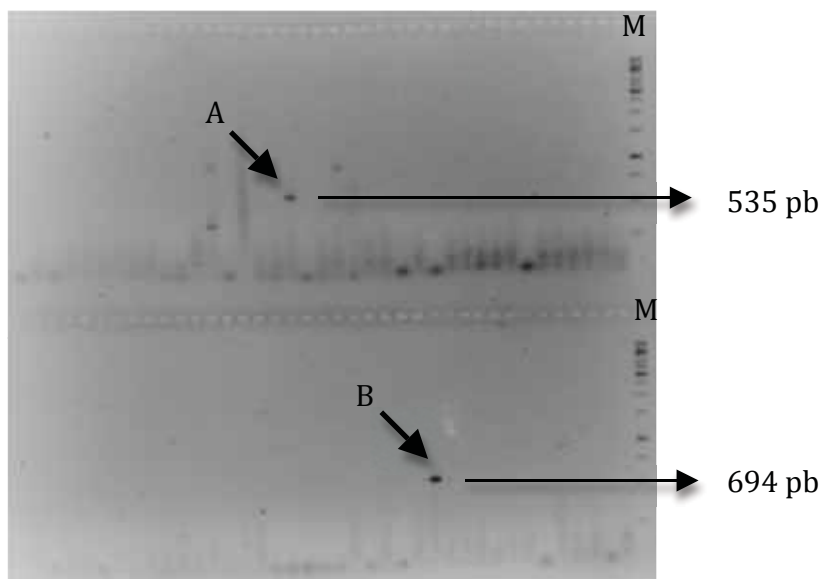


Figura 13: Perfil eletroforético em gel de agarose 1% da amplificação utilizando o conjunto de “primers” para gene de xilose isomerase de bactérias de alto conteúdo de bases C e G e DNA molde de “pools” de 96 clones da biblioteca metagenômica do consórcio degradador de hidrocarbonetos. A. Sinal de amplificação positiva detectado para o clone E2PL14. B. Sinal de amplificação positiva detectado para o clone A9PL14 M. Marcador 1 kb DNA “ladder” (Fermentas).

O produto de PCR do clone A9PL14 possui 694 bases. A sequência obtida foi comparada com as disponíveis no banco de dados Genbank, através da ferramenta Blastx. O produto de PCR apresenta 97% de similaridade e 93% de identidade com proteínas da família de transportadores ABC de *Pseudomonas aeruginosa*. Estes transportadores efetuam o transporte ativo através da membrana de diferentes tipos de moléculas, tais como açúcares, aminoácidos oligopeptídeos, íons, drogas, etc.

O produto de PCR do clone E9PL14 possui 535 bases. As análises utilizando a

ferramenta Blastx revelaram que o produto de PCR apresenta 65% de similaridade e 55% de identidade com proteínas hipotéticas de *Pseudomonas aeruginosa*.

4.4.4. Prospecção de genes na biblioteca de DNA do consórcio degradador de hidrocarbonetos e o conjunto de “primers” específico para genes de xilose isomerase de *Burkholderias*

O “primer” específico para genes de xilose isomerase de *Burkholderias* teve as suas condições de funcionamento determinadas utilizando como DNA molde o obtido de *Burkholderia thailandensis* (DSM 13276, ATCC 700388) (Fig. 14).

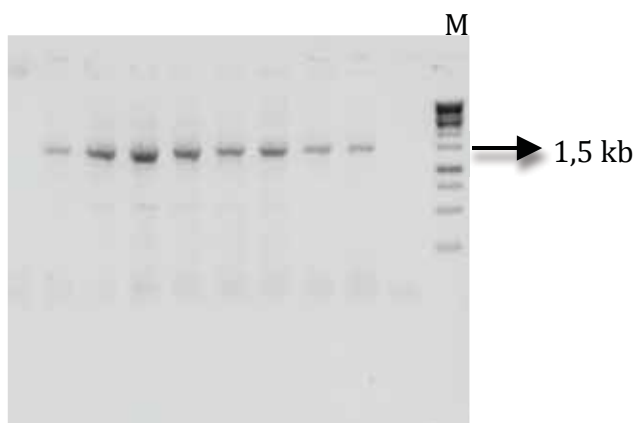


Figura 14: Perfil eletroforético em gel de agarose 1% da amplificação utilizando o conjunto de “primers” para gene de xilose isomerase de *Burkholderia* e DNA molde obtido de *B. thailandensis*. M. Marcador 1 kb DNA “ladder” (Fermentas).

Realizada esta etapa, iniciou-se a prospecção do gene nas bibliotecas metagenômicas. O único sinal de amplificação positiva obtido foi para o clone A8 da placa 18 da biblioteca metagenômica do consórcio degradador de hidrocarbonetos, com

o tamanho do amplicom compatível com o esperado para este “primer” (Fig. 15 e 16).

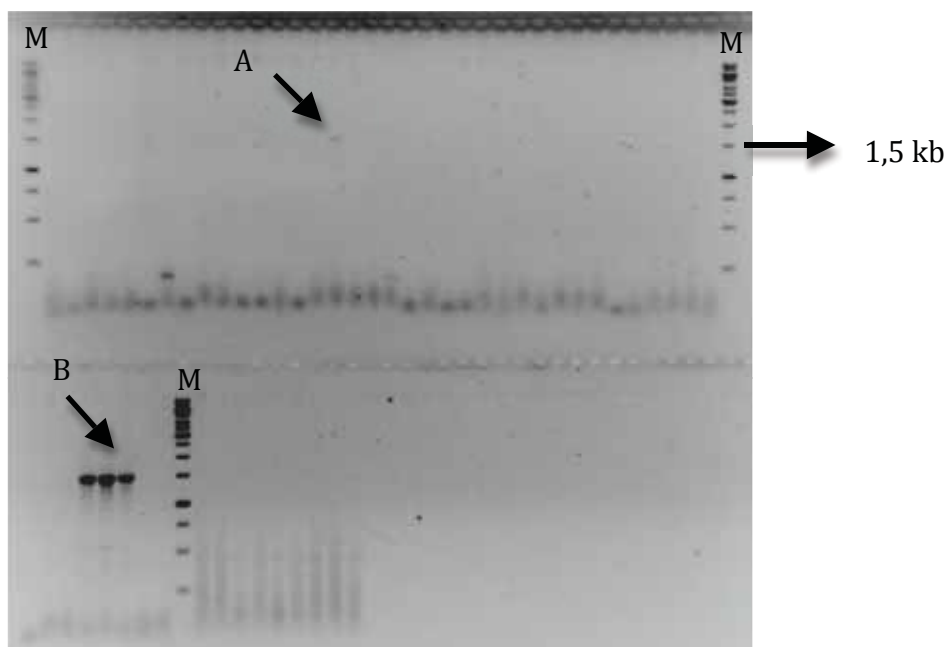


Figura 15. Perfil eletroforético em gel de agarose 1% da amplificação utilizando o conjunto de “primers” para gene de xilose isomerase de *Burkholderia* e DNA molde “pools” de 96 clones da biblioteca metagenômica do consórcio degradador de hidrocarbonetos. A. Sinal de amplificação positiva detectado para o “pool” da placa 18. B. Controle positivo da reação de PCR. M. Marcador 1 kb DNA “ladder” (Fermentas).

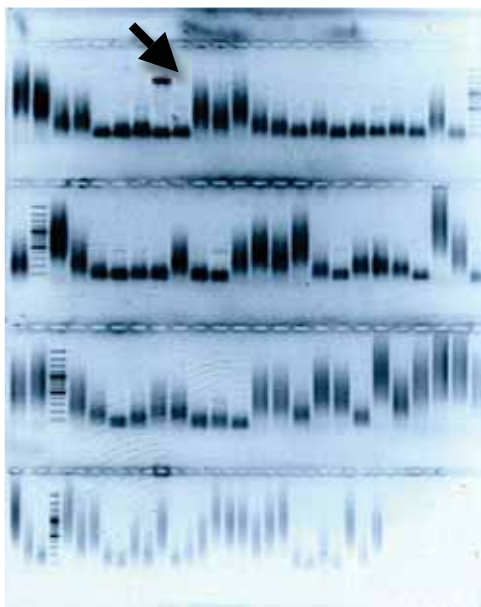


Figura 16: Perfil eletroforético em gel de agarose 1% da amplificação utilizando o conjunto de “primers” para gene de xilose isomerase de *Burkholderia* e DNA molde clones da placa 18 da biblioteca metagenômica do consórcio degradador de hidrocarbonetos. M. Marcador 100 bp DNA “ladder” (Fermentas).

O clone A8PL18 teve o seu cosmídio subclonado e totalmente sequenciado. O inserto de DNA metagenômico obtido possui tamanho de 25.269 pb. A sequência foi analisada para a presença de possíveis fases abertas de leitura (ORFs) utilizando o programa ORF finder, e as atribuições funcionais preliminares das ORFs individuais foi realizada de forma manual pela comparação dos possíveis produtos gênicos com proteínas de função conhecida depositadas nos bancos de dados. Vinte e uma ORFs foram identificadas. A possível função dos genes contidos nas ORFs encontradas foi determinada através da comparação com sequências similares presentes nos bancos de dados, e da análise dos domínios funcionais (tabela 4).

A ORF1 apresentou similaridade com proteínas da família de metanol/etanol desidrogenases dependentes de pirroloquinolina quinona (PQQ). Esta enzima é uma

quinoproteína, encontrada no periplasma de bactérias autotróficas e metilotróficas. Ela cataliza a oxidação do metanol ou outros alcoois primários em seus aldeídos correspondentes. Além do cofator pirroloquinolina quinona (PQQ), esta enzima requer um cátion de cálcio divalente para a sua atividade catalítica.

A ORF2 apresentou similaridade com 3-hidroxi-CoA desidrogenase. Esta enzima é classificada como uma oxidoredutase, e está envolvida em processos metabólicos de ácidos graxos. Especificamente catalisa a terceira etapa da beta-oxidação, a oxidação do L-3-CoA hidroxiacil por NAD⁺. A reação converte o grupo hidroxila em um grupo ceto.

As ORF 3 e a ORF 6 apresentaram similaridade com L-carnitina desidratase. Esta enzima pertence a família das liases, especificamente hidro-liases, as quais clivam ligações carbono-oxigênio.

As ORF 4 e a ORF 8 apresentaram similaridade com proteínas contendo domínios de acil-CoA desidrogenases. A acil-CoA desidrogenases são uma família enzimas que contém flavina adenina dinucleotídeo, que catalisam o primeiro passo na β -oxidação de ácidos graxos e catabolismo de alguns aminoácidos. O

A ORF 5 apresentou similaridade com sintetases e ligases dependentes de AMP cíclico. Esta família de enzimas inclui ácidos graxos-CoA de cadeia longa ligase, luciferase, acetil-CoA sintetase e várias outras sintetases estreitamente relacionadas.

A ORF 7 apresentou similaridade com proteínas "major facilitator superfamily MFS_1". Esta família inclui transportadores de membrana que realizam transporte de pequenos solutos em resposta a gradientes quimiosmóticos de íons.

As ORFs 9 e 10 apresentaram similaridade com enzimas da família thiolases, que são também conhecidas como acetil-CoA A acetiltransferases, que são enzimas que convertem duas unidades de acetil-CoA em acetoacetil-CoA na via biossintética do mevalonato.

A ORF 11 apresentou similaridade com aminoglicosídeo fosfotransferases, que são proteínas bacterianas que conferem resistência a antibióticos aminoglicosídeos via

fosforilação. A ORF 12 apresentou similaridade com luciferase bacteriana. esta proteína é uma flavina monooxigenases que catalisa a oxidação de aldeídos de cadeia longa e libera energia na forma de luz visível, e que utiliza como substrato flavina ao invés de um cofator. As ORFs 13 e 18 apresentaram similaridade com uma proteína hipotética

As ORFs 14 e 15 apresentaram similaridade com short-chain dehydrogenase/reductase SDR. Esta é uma família muito grande de enzimas, a maioria das quais são conhecidos por serem oxidorreduções NAD ou NADP-dependente.

A ORF 16 apresentou similaridade com álcool desidrogenases, que são um grupo de enzimas desidrogenases que ocorrem em muitos organismos e facilitam a interconversão entre álcoois e aldeídos ou cetonas com a redução de NAD + a NADH.

O ORF 17 apresentou similaridade com proteínas repressoras TetR, que são proteínas que regulam a expressão de genes de resistência a tetraciclina. A ORF 19 apresentou similaridade com proteínas de função desconhecida.

O ORF 20 e 21 apresentou similaridade com Enoil-CoA isomerase, que é uma enzima que catalisa a conversão da dupla ligação cis ácidos graxos na posição 3 a duplas ligações trans na posição 2. Ele tem uma importância especial no metabolismo de ácidos graxos insaturados.

Tabela 4: Possível função das ORFs contidas no inserto do clone fosmidial CRP A8P18

ORF	"Frame"	Posição (pb)	Tamanho (aa)	Proteína	Organismo	Sim/ Id
1	+2	14525-16615	696	<i>PQQ-dependent dehydrogenase, methanol/ethanol family</i>	<i>Sphingobium chlorophenolicum</i> L-1	88/79
2	+3	16602-18680	692	<i>3-hydroxyacyl-CoA dehydrogenase NAD-binding</i>	<i>Sphingobium chlorophenolicum</i> L-1	90/82
3	-1	23005-24948	660	<i>L-carnitine dehydratase/bile acid-inducible protein F</i>	<i>Sphingobium chlorophenolicum</i> L-1	87/79
4	+3	4776-6461	561	<i>acyl-CoA dehydrogenase domain-containing protein</i>	<i>Sphingobium chlorophenolicum</i> L-1	91/87
5	+3	12936-14519	527	<i>AMP-dependent synthetase and ligase</i>	<i>Sphingobium chlorophenolicum</i> L-1	92/86
6	+3	1755-3101	448	<i>L-carnitine dehydratase/bile acid-inducible protein F</i>	<i>Sphingobium chlorophenolicum</i> L-1	90/85
7	+3	11600-12943	447	<i>major facilitator superfamily MFS_1</i>	<i>Sphingobium chlorophenolicum</i> L-1	85/77
8	+2	10313-11578	421	<i>acyl-CoA dehydrogenase domain-containing protein</i>	<i>Sphingobium chlorophenolicum</i> L-1	96/90
9	-3	6500-7672	390	<i>acetyl-CoA acetyltransferase</i>	<i>Sphingobium chlorophenolicum</i> L-1	93/86
10	+2	18677-19828	383	<i>thiolase</i>	<i>Sphingobium chlorophenolicum</i> L-1	92/84
11	+3	8571-9554	327	<i>aminoglycoside phosphotransferase</i>	<i>Sphingobium chlorophenolicum</i> L-1	86/77
12	-3	21959-22939	326	<i>Luciferase-like, subgroup</i>	<i>Sphingobium chlorophenolicum</i> L-1	92/85
13	-3	248-1144	298	<i>hypothetical protein Plav_2409</i>	<i>Parvibaculum lavamentivorans</i> DS-1	52/31
14	+2	3098-3880	260	<i>short-chain dehydrogenase/reductase SDR</i>	<i>Sphingobium chlorophenolicum</i> L-1	92/84
15	+1	9598-10332	244	<i>short-chain dehydrogenase/reductase SDR</i>	<i>Sphingobium chlorophenolicum</i> L-1	89/80
16	-2	20265-20990	241	<i>alcohol dehydrogenase</i>	<i>Pseudomonas mendocina</i> ymp	77/57
17	-3	7718-8437	239	<i>regulatory protein TetR</i>	<i>Sphingobium chlorophenolicum</i> L-1	83/70
18	-1	21256-21903	215	<i>hypothetical protein SphchDRAFT_0152</i>	<i>Sphingobium chlorophenolicum</i> L-1	86/78
19	+3	19758-20249	163	<i>protein of unknown function DUF35</i>	<i>Sphingobium chlorophenolicum</i> L-1	91/81
20	+1	4318-4779	153	<i>Enoyl-CoA hydratase/isomerase</i>	<i>Sphingobium chlorophenolicum</i> L-1	95/90
21	+2	3905-4318	137	<i>Enoyl-CoA hydratase/isomerase</i>	<i>Sphingobium chlorophenolicum</i> L-1	85/77

"Frame": sentido de leitura da ORF no contig. Posição: posição da ORF no contig, representado em pares de bases. Tamanho: tamanho da proteína codificada, em aminoácidos. Proteína: proteína na qual a sequência estudada apresentou maior similaridade. Organismo: organismo cuja sequência apresentou maior homologia. Sim/Id (Similaridade e Identidade), percentagem de similaridade e identidade entre as sequências obtidas e o organismo relacionado.

Utilizando os “primers” para bactérias de baixo conteúdo de bases C e G e o “primer” para xilose isomerase fúngica, obtido a partir da sequência do gene de *Piromyces*, nenhum clone obteve sinal de amplificação positiva. A figura 17 mostra um gel de prospecção onde não foi possível identificar sinal de amplificação.

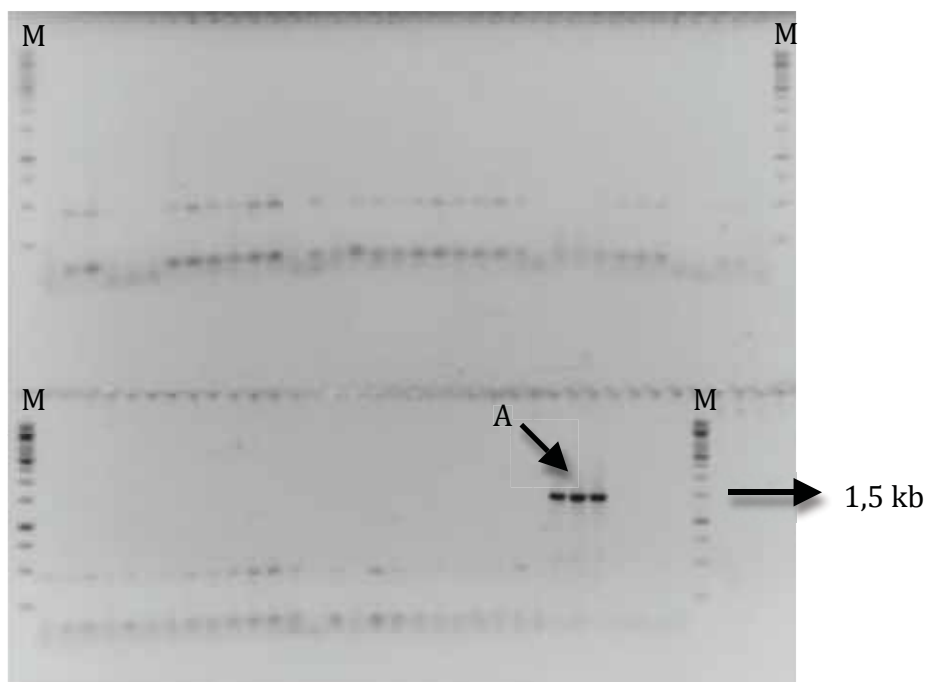


Figura 17. Perfil eletroforético em gel de agarose 1% da amplificação utilizando o conjunto de “primers” para gene de xilose isomerase de *Burkholderia* e DNA molde “pools” de 96 clones da biblioteca metagenômica do solo de eucaliptos. A. Controle positivo da reação de PCR M. Marcador 1 kb DNA “ladder” (Fermentas).

4.5. Identificação de genes de xilose isomerase em isolados de *Burkholderias*

A prospecção do gene de xilose isomerase foi realizada em isolados de *Burkholderia* pertencentes a coleção de micro-organismos do Laboratório de Micro-organismos e Plantas, através da técnica da reação em cadeia da polimerase (PCR), utilizando os “primers” degenerados propostos por LOPES *et al.*, 2009. O oligonucleotídeo foi desenhado com base na organização do locus *xyl* para diferentes espécies de *Burkholderia* (Fig. 18), de forma que o gene *xylA* estava completamente representado no amplicom.

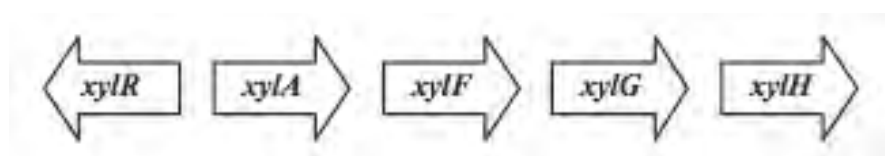


Figura 18. Organização do locus *xyl* para diferentes espécies de *Burkholderia* e *Pseudomonas*. O gene *xylA* codifica a enzima xilose isomerase. Os componentes de transporte *xylFGH* são representados por *xylF* (que codifica *xylose-binding protein*), *xylG* (*ATP-binding protein*), e *xylH* (transportador de membrana). O gene *xylR* está envolvido na regulação transcricional dos genes *xyl* (Song e Park, 1998).

Dos 13 isolados testados, 6 apresentaram amplificação positiva (Fig. 19). O amplicom obtido possuía aproximadamente 1,4 kb. Estes foram purificados (Fig. 20 e 21), clonados e submetidos ao sequenciamento automático de DNA. Cada um dos transformantes foi ressequenciado diversas vezes, para obtenção de uma sequência de alta qualidade e longa o suficiente para permitir a união das sequências direta e reversa

sem a necessidade de utilizar um oligonucleotídeo iniciador interno. A verificação do tamanho e qualidade das sequências, bem como a montagem dos contigs, foi realizada utilizando os programas Phred/Phrap/Consed.

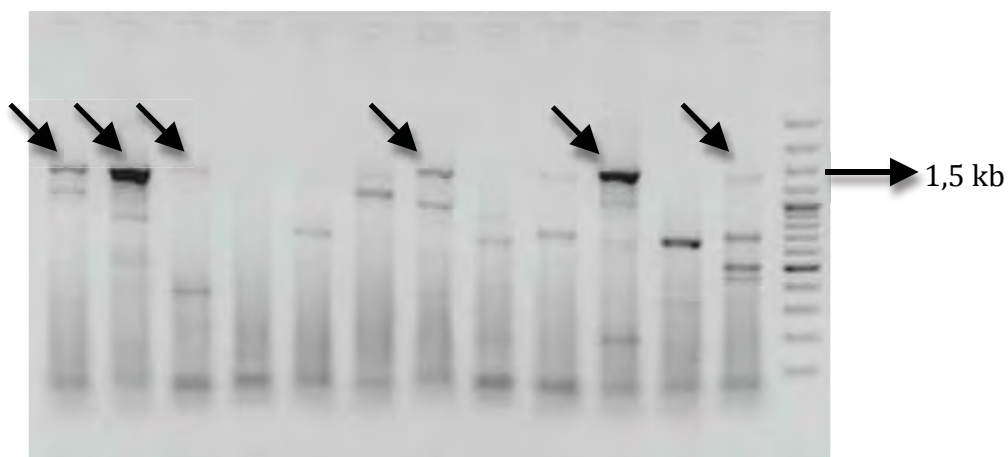


Figura 19. Perfil eletroforético em gel de agarose 1% do produto da amplificação utilizando o conjunto de “primers” para gene de xilose isomerase de *Burkholderia* e DNA molde obtido dos isolados de *Burkholderias*. M. Marcador 1 kb DNA “ladder” (Fermentas).

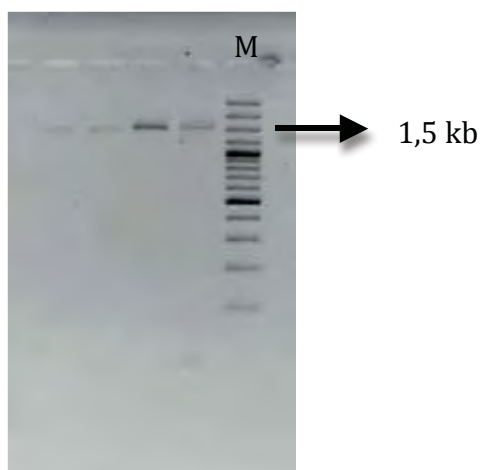


Figura 20. Perfil eletroforético em gel de agarose 1% do produto purificado da amplificação utilizando o conjunto de “primers” para gene de xilose isomerase de *Burkholderia* e DNA molde obtido dos isolados de *Burkholderias*. M. Marcador 1 kb DNA “ladder” (Fermentas).

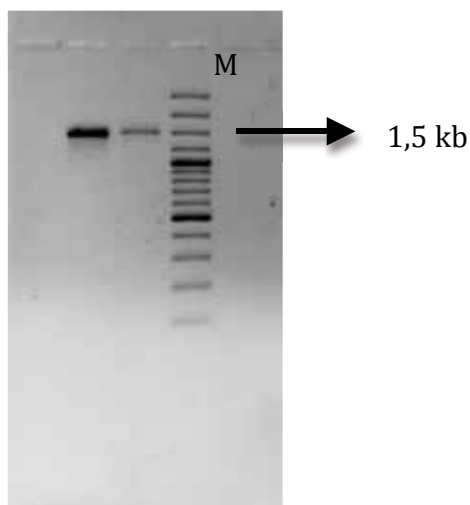


Figura 21. Perfil eletroforético em gel de agarose 1% do produto purificado da amplificação utilizando o conjunto de “primers” para gene de xilose isomerase de *Burkholderia* e DNA molde obtido dos isolados de *Burkholderias*. M. Marcador 1 kb DNA “ladder” (Fermentas).

4.6. Caracterização dos genes de xilose isomerase através de análises computacionais

O sequência de nucleotídeos obtida foi submetida ao programa OrfFinder, para identificação das ORFs presentes. A sequência de nucleotídeos e de aminoácidos das ORFS encontradas estão mostradas a seguir.

4.6.1. Sequência do isolado B1

>Orf_isolado_B1

```
ATGTCGTATTTTGAACAGATTCCCGCGATTGCTACGAAGGTCCGCAATCGGACAACCCGCTTGC
CTACCGCCATTACGACCGCACGAAGCAGGTGCTCGGCAAGACGCTCGAGGAACACTTGCGCATCG
```

CGGTGTGCTACTGGCACACGTTTCGTGTGGCCGGGGCACGACATCTTCGGCCAGGGCGCGTTCCGG
 CGGCCGTGGCAGCAGCCGGGCGACCCGCTCGAGCGGGCGCGGCAGAAGGCGGATGCGGGCGTTCCG
 GTTCTTCACGAAGCTCGGCACGCCGTTCTATACGTTCCACGACACCGATGTCGCGCCGGAAGGCG
 ACAGCCTGCGCGAGTACGTCGAGAATTTTCGCGCGGATGACGGACTACCTGGGCGAGCGTCAGCAG
 GCCACCGGCGTGC GGCTGCTGTGGGGCACCGCGAACCTGTTCTCGCATCCGCGCTTCGCGGGCGGG
 CGCGGCGACCAACCCGAATCCCGACGTGTTTCGCGTGGGCGGGCGACCCAGGTGCGCCACGCGCTCG
 ACGCGACGCACCGGCTCGGCGGGCGAGAACTACGTGCTGTGGGGCGGGCCGCGAAGGCTACGAGACG
 CTGCTCAATACCGATCTCGTGCGCGAGCGCGAGCAGTTTCGCGCGCTTCCTGTCGATGGTCGTCGA
 GCACAAGCACCGGATCGGTTTCAAGGGCGCGCTGCTGATCGAGCCGAAGCCGCAGGAGCCGACCA
 AGCACCAGTACGACTACGACGTCGCGACCGTGCACGGTTTCTCACGCAATACGGATTGCAGAAC
 GAGATCCGCGTGAACATCGAGGCGAACACGCGACGCTCGCCGGCCATTTCGTTCCATCACGAGAT
 CGCGAACGCGTTCGCGCTCGGCGTGTTCGGCAGCGTCGATGCGAATCGCGGGCGATCCGCGAAGC
 GCTGGGACACCGACAGTTCCCGAACAGCGTCGAGGAGCTGACCCTCGCGTTCTACGAGATCCTG
 CGCCACGGCGGCTTCACTACCGGCGGCATGAACTTCGACGCGAAGGTGCGGGCGCCAGAGCGTCGA
 TCCGGAAGACCTGTTCCACGGCCATGTGGGCGCGATCGACGTGCTCGCGCTCGCGCTCGAACGCG
 CCGCCGTGCTGGTCGCCAACGATCGGCTCGACGCGCTGCGTCGGCAGCGCTACGCGCAGTGGGAC
 AGCGCGTTTCGGCCGCGAGATCCTGTGCGGGCGGCTATACGCTGGAGTCGCTCGCCGCCGACGCGCT
 CGCGCGCGGCGTGAACCCGCAGCATGCGAGCGCGGGCAGGAGCGGCTCGAGAACATCGTGAACC
 AGGCGATCTACGCGCTGCGCTGA

>proteina_isolado_B1

MSYFEQIPAIRYEGPQSDNPLAYRHYDRTKQVLGKTLEEHLRIAVCYWHTFVWPGHDFGQGAFR
 RPWQQPGDPLERARQKADAAFGFFTKLGTPTFYTFHDTDVAPEGDSLREYVENFARMTDYLGERQQ
 ATGVRLLWGTANLFSHPFAAGAATNPNDVF AWAATQVRHALDATHRLGGENYVLWGGREGYET
 LLNTDLVREREQFARFLSMVVEHKHRIGFKGALLIEPKPQEPTKHQYDYDVATVHGFLTQYGLQN
 EIRVNI EANHATLAGHSFHHEIANAFALGVFGSVDANRGDPQNGWDTDQFPNSVEELTLAFYEIL
 RHGGFTTGGMNFDKVRQSVDPEDLFHGHVGAIDVLALALERA AVL VANDRLDALRRQRYAQWD
 SAFGREILSGGYTLES LAADALARGVNPQHASGGQERLENIVNQAIYALR*

4.6.2. Sequência do isolado B2

>ORF_isolado_B2

ATGTCGTATTTTGAATCCATCCCCGCCATTTCGCTACGAAGGTCCGCAGTCGGACAATCCACTCGC
 TTTCCATTACTACGACCGCTCGAAAAAGGTGCTCGGCAAGACGATGGAAGAGCATCTGCGCGTGG
 CCGTGTGCTACTGGCACACGTTTCGTGTGGCCGGGCATCGATATTTTCGGGCGAGGGCACCTTCCAT
 CGTCCGTGGCACCAGCCCGGCGACGCGATGGAGCGGGCACGCCAGAAGGCCGACGCCGCGTTTCA
 ATTTTTCTCGAAGCTCGGCGCGCGTTCTATACGTTCCACGACACCGATGTAGCGCCCGAAGGCA
 ACGATCTGAAGGACTACACAGAAAATTTGCCCCGATGGTCGATTACCTGGGCGAGCGCCAGCAG
 GCGAGCGGCGTGAAGCTTCTGTGGGGCACGGCGAATCTCTTTTTCGAATCCACGCTACGCGGGCGG
 CGCGGCGACCAACCCGAACCCCGAGGTGTTTCGCTATGCGGCGACCCAGGTGTGCCACGCTCTCG
 ACGCCACGCGCAAACTGGGTGGCGAAAACCTACGTGCTGTGGGGCGGGCCGCGAAGGCTATGAAACG

CTGCTCAATACCGACCTCGTGCGCGAGCGCGAACAGTTCGCGCGCTTTCTTTTCGATGGTGGTGGA
 GCACAAGCACAGCACGGGCTTCACGGGCACGCTGCTGATCGAACCCAAGCCGAGGAGCCACCA
 AGCATCAGTACGACTACGACGTGGCGACGGTGCAAGGCTTTCTCACGCAGTACGGCTTGCAGAAC
 GAGATTGCGGTGAACATCGAGGCGAATCACGCGACGCTCGCCGGCCACTCGTTTCATCACGAAAT
 CGCCATGGCCTACGCATTGGGCGTGTTTCGGCAGTGTGGACGCCAACC GCGGCGATCAGCAGAACG
 GCTGGGATACCGACCAGTTCGCCAACAGCGTGGAAGAACTCACGCTCGCGTTCTACGAGATCCTG
 CGTCACGGCGGCTTCACGACCGGCGGCATGAACCTTCGACGCGAAGGTGCGGCGCCAGAGCGTCGA
 TCCCGAAGACCTGTTCTACGGGCACGTCGGCGCGATGGACGTGCTGGCGATCGCGCTGGAGCGCG
 CCGCGTCTTGGTCGAGAACGAGCGCCTCGACCAGTTGCGCAAGCAACGCTATGCGGGCTGGGAT
 ACGGAGCTCGGCCGCAAGATTCTGGCGGGCGGCTACTCGCTCTCGTCCCTTGCCGCCGATGCCAC
 TCGCATGGTGTGAACCCGCATCACGCAAGCGGGCAGCAAGAACGGTTCGAAAACATCGTCAATC
 AGGCGATCTACGGGCTGCGCTGA

>proteína_isolado_B2

MSYFESIPAIRYEGPQSDNPLAFHYDRSKKVLGKTMEEHLRVAVCYWHTFVWPGIDIFGQGTFFH
 RPWHQPGDAMERARQKADAAFEFFSKLGAPFYTFHDTDVAPGNDLKDYTENFARMVDYLGERQQ
 ASGVKLLWGTANLFSNPRYAAGAATNPNEVFAYAATQVCHALDATTRKLGGENYVLWGGREGYET
 LLNTDLVREREQFARFLSMVVEHKHSTGFTGTLLEPKPQEPTKHQYDYDVATVQGFLLTQYGLQN
 EIRVNIENHATLAGHSFHHEIAMAYALGVFGSVDANRGDQQNGWDTDQFPNSVEELTLAFYEIL
 RHGGFTTGGMNFDKVRQSVDPEDLFYGHVGAMDVLAIALERAALVENERLDQLRKQRYAGWD
 TELGRKILAGGYSLSLAADATAHGVPNPHASGQQRERFENIVNQAIYGLR*

4.6.3. Sequência do isolado B3

>ORF_isolado_B3

ATGTCGTATTTTGAAGCCATTCCCGCGATTTCGTTACGAAGGTCCGCAATCGGATAATCCGCTTGC
 CTATCACCATTACGACCGCACGAAGAAGGTGCTCGGCAAGACGCTCGAAGAGCATTTGCGCATTG
 CCGCGTGTTACTGGCATTTCGTTTCGTGTGGCCGGGCGTCGATATCTTCGGGCAGGGCACGTTCAA
 CGTCCGTGGCAGCAACCGGGCGATGCAATGGAACGCGCGCATCAGAAAGCCGACGCCGCGTTTCA
 ATTCTTCACGAAGCTCGGCGTGCCGTTCTACACCTTCCACGACACCGACGTGTCGCCCCGAAGGCA
 AGGACCTGAAGGACTACACCGAAAACCTTCGCGCGCATGGTCGATTACCTCGGCGCGAAGCAGCAG
 TCGAGCGGCATGAAGCTGTTGTGGGGCACGGCGAATCTGTTTTTCGATCCGCGCTATGCGGGCGG
 CGCGGCGACGAGCCGAATCCGGAAGTGTTTCGCTTATGCGGCGACGCAGGTGTGCCACGCGCTCG
 ACGCCACGCACAAGCTCGGCGGCGAAAACCTATGTGTTGTGGGGCGGTTCGCGAAGGTTACGACACA
 TTGCTGAACACCGATCTCGTACGTGAGCGCGAACAGTTCGCGCGCTTTCTTTTCGATGGTGGTGAG
 CACAAGGAGCGCATCGGCTTCAAGGGCGCGCTGCTGA

>proteína_isolado_B3

MSYFEAIPAIRYEGPQSDNPLAYHHYDRTKKVLGKTLEEHLRIAACYWHSFVWPGVDIFGQGTFFK
 RPWQQPGDAMERAHQKADAAFEFFTKLGVPFYTFHDTDVSPGKDLKDYTENFARMVDYLGAQQ

SSGMKLLWGTANLFSHPRYAGGAATSPNPEVFAYAATQVCHALDATHKLGGENYVLWGGREGYDT
LLNTDLVREREQFARFLSMVVSTRSASASRARC*

4.6.4. Sequência do isolado B7

>ORF_isolado_B7

ATGTCGTATTTTGAACACATTCCCGCGATTTCGCTACGAAGGCCCGCAGTCGGACAATCCGCTCGC
GTACCACCATTACGACCCCGAGCGGCGCGTGCTGGGCAAGACACTCGCCGAGCATCTGCGGATCG
CGGTGTGCTACTGGCACACGTTTCGTGTGGCCGGGGCACGACATCTTCGGGCAGGGCGCGTTCCAG
CGGCCGTGGCAGCAGCCGGGCGACGCGCTCGAGCGGGCGCGGATGAAGGCCGATGCGGCGTTTCA
GTTCTTCACGAAGCTCGGCACGCCTTTCTATACGTTCCATGACACCGACGTCGCACCGGAAGGCG
ACAGCCTGCGGGAATATGCGGCGAATTTTCGCGCGGATGGTCGACTATCTCGGCGAGCGTCAGCAG
GCGAGCGGCGTGCGGCTCCTGTGGGGCACCGCTAACCTGTTCTCGCACCCGCGCTTCGCGGCCGG
CGCCGCGACGAACCCGAACCCCGACGTGTTTCGCGTGGGCTGCGACCCAGGTATGTCATGCGCTCG
ATGCGACGCACCGGCTCGGCGGCGAGAACTATGTGCTGTGGGGCGGGCGGGAAGGGGTACGAACG
CTGCTCAATACGGACCTGAAGCGCGAACGCGAGCAGTTGCCCGCTTCTGTGCGATGGTCGTCGA
GCACAAGCACCGGATCGGCTTCACGGGCGCGCTGCTGATCGAGCCGAAGCCGCGAGGAGCCGACCA
AGCACCAGTACGACTACGACGTCGCGACGGTGCACGGCTTTCTCGTGACGTACGGATTGCAGAAC
GAAATTCGCGTGAACATCGAGGCGAATCACGCGACGCTCGCCGGCCACTCGTTCCATCACGAGAT
CGCGAACGCGTTTCGCGCTCGGCGTGTTTCGGCAGCGTCGACGCGAACCGCGGCGATCCGCGAAGC
GCTGGGACACCGACAGTTCCCGAACAGCGTCGAGGAGCTGACGCTCGCGTTCTACGAAATCCTG
CGTCACGGCGGCTTCACGACCGGCGGGATGAACTTCGACGCGAAGGTGCGCCGCCAGAGCATCGA
TCCCGAAGATCTGTTCTACGGCCACGTCGCGCGGATCGACGTGCTCGCGCTCGCGCTCGAACGCG
CGGCGGTGCTGGTCGAGAACGACCGGCTCGACGCACTGCGCCGGCAACGCTATGCGCAGTGGGAC
GACGCGTTTCGGCCGCAAGATCCTGTGCGGGCGACTATACGCTCGAGTCGCTCGCGGCCGATGCGCT
CGCGCGCGGTGTGAACCCGCGACGCGAGCGGCGCGCAGGAGCGGTGCGAGAACATCGTGAACC
AGGCGATCTACGCGCTGCGCTGA

>proteína_isolado_B7

MSYFEHIPAIRYEGPQSDNPLAYHHYDPERRVLGKTLAEHLRIAVCYWHTFVWPGHDIFGQGAFFQ
RPWQQPGDALERARMKADAAFEFFTKLGTPFYTFHDTDVAPEGDSLREYAANFARMVDYLGERQQ
ASGVRLWGTANLFSHPRFAAGAATNPNDVFAWAATQVCHALDATHRLGGENYVLWGGREGVRT
LLNTDLKREREQLPRFLSMVVEHKHRIGFTGALLIEPKPQEPTKHQYDYDVATVHGFLVQYGLQN
EIRVNIENHATLAGHSFHHEIANAFALGVFGSVDANRGDPQNGWDTQFPNSVEELTLAFYEIL
RHGGFTTGGMNFDKAVRRQSIDPEDLFYGHVGAIDVLALALERA AVL VENDORLDALRRQRYAQWD
DAFGRKILSGDYTTLESLAADALARGVNPQHASGAQERCENIVNQAIYALR*

4.6.5. Sequência do isolado B10

>ORF_isolado_B10

ATGTCGTATTTTCGAGCACATTCCCGCGATTTCGCTACGAAGGCCCGCAGTCGGACAACCCGCTTGT
 TTATCACCATTACGACCCCGAGAAGCGCGTGCTCGGCAAGACGCTCGCCGAGCATCTGCGGATCG
 CGGTCTGCTATTGGCACACGTTCTGTCTGGCCGGGCAACGACATCTTCGGGCAGGCCGCGATTTCGAG
 CGGCCGTGGCACCAGCCGGGCGACGCGCTCGAGCGCGCGCGGATGAAGGCCGATGCCGCGATTTCGA
 ATTCTTCACGAAGCTCGGCACGCCGTTCTATACGTTCCATGACACCGACGTCGCGCCGGAAGGCG
 ACAGCCTGCGCGACTACGCAGCCAATTTTCGCGCGGATGGTCGACTATCTCGGCGAGCGTCAGCAG
 GCAAGCGGCGTGCGGCTCCTGTGGGGCACCGCGAACGCTGTTCTCGCACCCGCGCTTCGCGGCCG
 GCGCGCCGCGACCAACCCGAACCCCGACGTGTTTCGCATGGGCGGCGACCAAGGTGTGCCACGCGCTC
 GACGCGACCCACCGGCCCGGGGAGGAACTACGTCCTGTGGGGCGGACGGGAAGGCTAGAACCC
 TGCTCATACCGACCTGAAGCGCGAGCGCGACCAAGTCGCCCCGCTTCTGTGCGATGTCGTCGAGCAC
 AGCACCGGATCGCTTCAAGGGCGCGCTGCTGATCGAGCCGAAGCCGAGGAGCCGACCAAGCACC
 AGTACGACTACGACGTCGCGACCGTGCACGGCTTCTCTGTCGAGTACGGACTCCAGAACGAGATT
 CGCGTGAACATCGAGGCGAACACGCGACGCTCGCCGGTCATTCTGTTCCATCAGAGATCGCGAG
 CGCCTTCGCGCTCGGCGTGTTTCGGCAGCGTCGACGCGAACCGCGGCGATCCGCGAGAACGGCTGGG
 ACACCGACCAAGTTCCCGAACAGCGTCGAGGAGCTGACGCTCGCGTTCTACGAGATCCTGCGCCAC
 GCGGCTTCACGACCGGCGGGATGAACTTCGACGCGAACGTTGCGACGCCAGAGCATCGATCCGGA
 AGACCTGTTCTACGGCCACGTCGGCGCGATCGACGTGCCGCGCTCGCGCTCGAACGCGCGGCGAG
 TGCTGGTCGAGAACGACCGGCTCGACGCGCTGCGCCGGCAGCGCTACGCGCAGTGGGACGACGCG
 TTCGGCCGCAAGATCCTGTCCGGCGGTTATACGCTGGAATCGCTGGCGGCCGACGCGCTCGCGCG
 TGGGGTGAACCCGCGGCATGCGAGCGGCGCGCAGGAGCGGCTGGAGAACATCGTGAACCAGGCGA
 TCTACGCACTTCGCTGA

>proteina_isolado_B10

MSYFEHIPAIRYEGPQSDNPLVYHHYDPEKRVLGKTLAEHLRIAVCYWHTFVWPGNDFGQAAFE
 RPWHQPGDALERARMKADAAFEFFTKLGTPTFYTFHDTDVAPEGDSLRYAANFARMVDYLGERQQ
 ASGVRLWGTANAVLAPALRGRAPRPTPTCSHGRRPVPRARRDPPARGRKLRPVGRTRLEP
 CSYRPEARARPVARFLSMSSSTAPDRFKGALLIEPKPQEPTKHQYDYDVATVHGFLVQYGLQNEI
 RVNIEANHATLAGHSFHHEIASAFALGVFGSVDANRGDPQNGWTDQFPNSVEELTLAFYEILRH
 GGFTTGGMNFDKVRQRQSIDPEDLFYGHVGAIDVPALALERA AVL VENDORLDALRRQRYAQWDDA
 FGRKILSGGYTLES LAADALARGVNPRHASGAQERLENIVNQAIYALR*

4.6.6. Sequência do isolado B20**>ORF_isolado_B20**

ATGTCGTATTTTCGAACACATTCCCGCGATTTCGCTACGAAGGCCCGCAATCGGACAATCCGCTCGC
 GTACCGGCATTACGACCGCACGAAGCGGGTGCTCGGCAAGACGCTCGAGGAGCACCTGCGGATCG
 CCGTCTGCTATTGGCACACGTTCTGTGTGGCCGGGGCAAGACGTGTTTCGGCCAGGGCGCGTTGCGG
 CGGCCGTGGCAGCAGCCGGGCGAGCCGCTCGAGCGCGCGCGGCGAGAAGGCCGATGCGGCGTTTCG
 GTTCTTCACGAAGCTCGGCACGCCGTTCTATACGTTTCACGACACCGACGTCGCGCCGGAAGGGG
 CGAACCTGCACGAGTACGTCGAGAACTTCGCGCGGATGGTCGACTACCTCGGCGAGCACCAGCAG

GCGAGCGGCGTGCGGCTGCTGTGGGGCACCGCGAACCTGTTCTCGCATCCGCGCTTCGCGGCCGG
 CGCGGCGACGAGCCCGAATCCGATGCGTTCGCGTGGGCGGCGACCCAGGTGTGCCACGCGCTCGA
 CGCGACGCACCGGTTTCGGCGGCGAGAACTACGTGCTGTGGGGCGCCCCGGAAGGGTACGAGACG
 CTGCTCAACACGGATCTCGCGCGCGAGCGCGAGCATTGGGCCAGGTTTTTTGGCGATGGTCGTCTGA
 GCACAAGCACCGGATCGGCTTCACCGGCGCGCTGCTGATCGAGCCGAAACCGCAGGAGCCGACGA
 AGCACCAGTACGACTACGACGTGCGGACCGTGCACGGCTTGCTCACGCAATACGGATTGCAGAAC
 GAGATCCGCGTGAACATCGAGGCGAACACGCGACGCTCGCCGGCCATTTCGTTCCATCACGAGAT
 CGCGAACGCGTTCGCGCTCGGCGTGTTCGGCAGCGTCGACGCGAACCGCGGCGATCCGCAGAACG
 GCTGGGACACCGACCAAGTTCCCGAACAGCGTCGAGGAGCTGACGCTCGCGTTCTACGAGATCCTG
 CGCCACGGCGGTTTTACGACGGGCGGCATGAAC TTCGACGCGAAGGTACGGCGCCAGAGCGTCGA
 TCCGGAGGACCTGTTTTACGGCCATATCGGCGCGATCGACGTGCTCGCGCTCGCGCTCGAGCGCG
 CGGCCGTGCTGGTGGAGAACGACCGGCTCGACGCGCTGCGCCGGCGCCGCTACGCGCAATGGGAT
 TCGGAGTTCGGCCGCAAGATACTCGCGGGCGGTTATTCGCTGCAGTCGCTCGCGGCCGACGCGCT
 CGCGCGCGGCGTGAACCCGCGACGCAAGCGGCGCGCAGGAGCGGCTCGAGAACATCGTGAACC
 AGGCGATCTACGCACTGCGCTGA

>proteina_isolado_B20

MSYFEHIPAIRYEGPQSDNPLAYRHYDRTKRVLGKTL EEHLRIAVCYWHTFVWPGQDVFGQGALR
 RPWQQPGEPLERARQKADAAFAFFTKLGPFTFHD TDVAPEGANLHEYVENFARMVDYLGEHQQ
 ASGVRLWLGTANLFSHPFAAGAATSPNPMRSRGR RPRCATRSTRRTGSAARTTCCGAPREGYET
 LLNTDLAREREHWARFLAMVVEHKHRIGFTGALLIE PKPQEPTKHQYDYDVATVHGLLTQYGLQN
 EIRVNI EANHATLAGHSFHHEIANAFALGVFGSVDANRGDPQNGWDTDQFPNSVEELTLAFYEIL
 RHGGFTTGGMNFDAKVRQSVDPEDLFHGHIGAIDV LALALERA AVL VENDORLDALRRRRYAQWD
 SEFGRKILAGGYSLQSLAADALARGVNPQHASGAQER LENIVNQAIYALR*

A anotação dos genes foi realizada a partir da sequência de aminoácidos das ORFs encontradas, através da procura por similaridades com proteínas presentes em diferentes bancos de dados, utilizando a ferramenta de alinhamento local Blastp. O resumo dos relatórios obtidos no Blastn e Blastp estão mostrados nas tabelas 5 e 6, respectivamente.

Tabela 5. Identificação dos genes de xilose isomerase dos isolados utilizando o algoritmo Blastn.

Linhagem	Verossimilhança ⁽¹⁾	S ⁽²⁾	ID ⁽³⁾	Frag ⁽⁴⁾	GenBank ⁽⁵⁾	e-value ⁽⁶⁾
B1	<i>Burkholderia multivorans</i>	1934	93%	1323	AP009386.1	0,0
B2	<i>Burkholderia vietnamiensis</i>	1303	85%	1323	CP000616.1	0,0
B3	<i>Burkholderia sacchari</i>	665	84%	687	FJ374785.1	0,0
B7	<i>Burkholderia cenocepacia</i>	2034	94%	1323	AM747722.1	0,0
B10	<i>Burkholderia cenocepacia</i>	2194	97%	1317	CP000460.1	0,0
B20	<i>Burkholderia thailandensis</i>	2394	100%	1323	CP000086.1	0,0

⁽¹⁾Verossimilhança, organismo que possui a sequência com a qual a sequência estudada apresentou maior similaridade. ⁽²⁾S (Score), pontuação do alinhamento. ⁽³⁾Id (Identidade), percentagem de identidade entre a sequência das linhagens estudadas e o organismo relacionado. ⁽⁴⁾Frag. (Tamanho do fragmento), tamanho da sequência de nucleotídeos. ⁽⁵⁾GenBank, número de acesso da sequência do organismo relacionado, ⁽⁶⁾e-value: expect value.

Tabela 6. Identificação das proteínas de xilose isomerase dos isolados utilizando o algoritmo Blastp.

Linhagem	Verossimilhança ⁽¹⁾	S ⁽²⁾	Sim/Id ⁽³⁾	Tam ⁽⁴⁾	GenBank ⁽⁵⁾	e-value ⁽⁶⁾
B1	<i>Burkholderia ubonensis</i>	885	99%/98%	441	ZP_02379535.1	0,0
B2	<i>Burkholderia ubonensis</i>	811	94%/88%	441	ZP_02379535.1	0,0
B3	<i>Burkholderia sacchari</i>	426	94%/91%	229	ACJ11237.1	0,0
B7	<i>Burkholderia cenocepacia</i>	883	98%/97%	441	YP_626338.1	0,0
B10	<i>Burkholderia cenocepacia</i>	709	85%/83%	439	YP_626338.1	0,0
B20	<i>Burkholderia thailandensis</i>	809	95%94	441	YP_442859.2	0,0

⁽¹⁾Verossimilhança, organismo que possui a sequência com a qual a sequência estudada apresentou maior similaridade. ⁽²⁾S (Score), pontuação do alinhamento. ⁽³⁾Sim/Id (Similaridade e Identidade), percentagem de similaridade e identidade entre a sequência das linhagens estudadas e o organismo relacionado. ⁽⁴⁾Frag. (Tamanho do fragmento), tamanho da sequência de aminoácidos. ⁽⁵⁾GenBank, número de acesso da sequência do organismo relacionado, ⁽⁶⁾e-value: expect value.

Uma árvore filogenética foi construída, baseado nas sequências de nucleotídeos para o gene de xilose isomerase obtidas neste estudo e outras selecionadas de bancos de dados públicos. A árvore obtida está mostrada na Fig. 22.

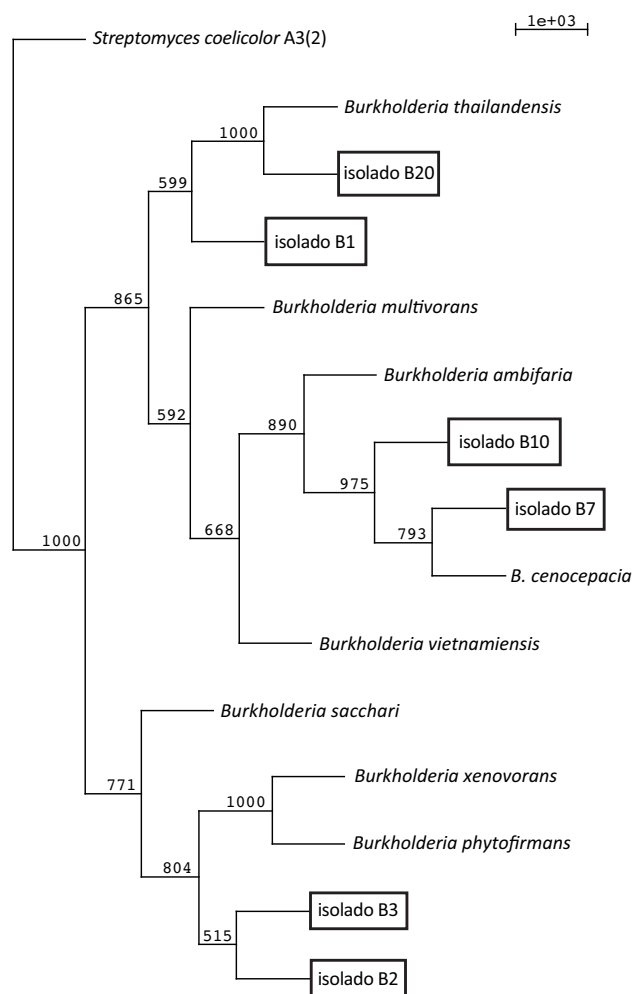


Figura 22. Análise filogenética do gene de xilose isomerase. A reconstrução filogenética foi computada pelo método de distância, utilizado o algoritmo Jukes-Cantor e o método neighbor-joining de construção de filogramas, com 1000 "bootstraps". O número de substituições de nucleotídeos é proporcional ao comprimento da escala. As sequências obtidas neste trabalho estão ilustradas em vermelho. Os números de acesso das sequências oriundas dos bancos de dados são: *B. cenocepacia* (geneID: 6804251), *B. thailandensis* (geneID: 3848180), *B. multivorans* (geneID: 6362527), *B. ambifaria* (geneID: 6181849), *B. vietnamiensis* (geneID: 4949355), *B. xenovorans* (geneID: 4006662), *B. phytofirmans* (geneID: 6279032) e *B. sacchari* (Genbank: FJ374785.1).

A árvore filogenética do gene de xilose isomerase mostra que a sequência de alguns dos genes encontrados possuem estreita relação com os disponíveis nos bancos de dados, como é o caso do isolado B20, que forma um grupo monofilético com *B. thuylandensis*, e do isolado B7, que forma um grupo monofilético com *B. cenocepacia*. Os isolados B2 e B3 formaram juntos um grupo monofilético.

A sequência de aminoácidos dos isolados de *Burkholderia* e mais 50 sequências de aminoácidos que codificam a enzima xilose isomerase disponíveis nos bancos de dados públicos foram utilizados para construir uma análise filogenética (Fig. 23). Dentre as sequências elegidas para compor a árvore, estão representadas sequências de xilose isomerase fúngicas, bacterianas e de plantas. O objetivo desta análise foi situar as sequências identificadas neste trabalho dentro do universo de sequências para enzimas de xilose isomerase. As sequências obtidas neste trabalho formaram um grupo monofilético composto, e estão incluídas dentro do grupo que agrupa as sequências provenientes de proteobactérias, como era esperado.

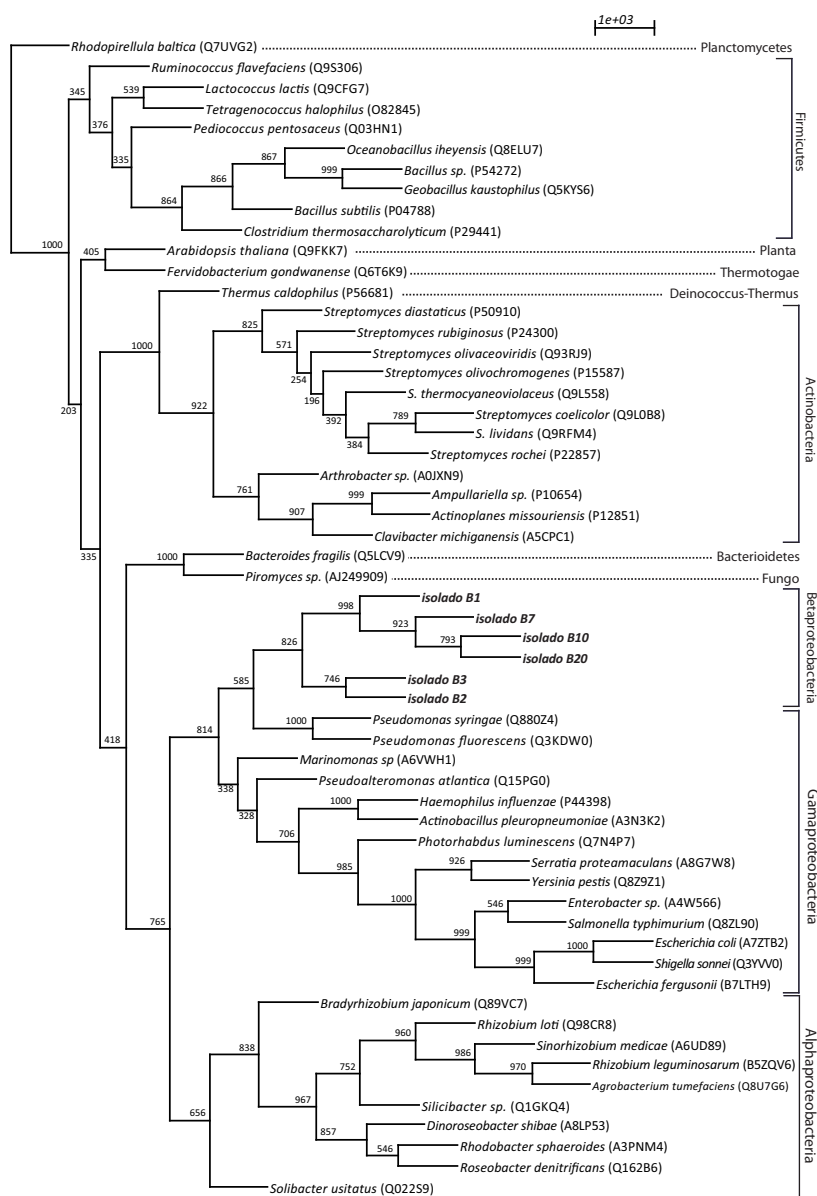


Figura 23. Análise filogenética da sequência de aminoácidos da proteína xilose isomerase. A reconstrução filogenética foi computada pelo método de distância, utilizado a matriz PAM e o método *neighbor-joining* de construção de filogramas, com 1000 *bootstraps*. O número de substituições de aminoácidos é proporcional ao comprimento da escala. As sequências obtidas neste trabalho estão ilustradas em negrito. Os números de acesso das sequências oriundas dos bancos de dados estão em parênteses ao lado do nome das linhagens.

Foi realizada uma análise para prever o quanto os genes identificados neste trabalho serão bem sucedidos quando expressos em determinado sistema de expressão, baseando-se na presença de códons preferenciais e raros. Existem diversas ferramentas em bioinformática disponíveis para a realização dessa predição baseado na sequência gênica codificadora da proteína heteróloga. Baseando-se nisso, as sequências dos genes de xilose isomerase obtidos foram analisadas no "European Molecular Biology Open Software Suite" (EMBOSS), sítio que disponibiliza várias ferramentas de bioinformática para diversos tipos de análises. Essa ferramenta foi utilizada para definir um índice de adaptação de códon ("Codon Adaptation Index" - CAI).

O índice de adaptação dos códons (SHARP & LI, 1987) é uma simples e efetiva medida do códon preferencial e é usado para inferir o nível de expressão do gene heterólogo, assim como demonstrar a probabilidade de uma ORF (fase de leitura aberta) ser codificadora de proteína. Ele resume a adaptação do uso dos códons na direção dos códons utilizados pelos genes altamente expressos. Foi realizada a predição do nível de expressão dos genes encontrados nas espécies *E. coli*, *Bacillus subtilis* e *S. cerevisiae*, com a finalidade de se encontrar um hospedeiro onde a expressão seja maximizada. Os resultados desta análise estão mostrados na tabela 7. Esta análise revelou que, comparado com *S. cerevisiae*, o índice de adaptação de códons da sequência dos genes dos isolados é sempre muito baixo, variando entre 0.036 e 0.055 (o valor máximo possível é 1,0), os quais podem resultar em uma taxa de expressão gênica ineficiente em *S. cerevisiae*. Porém, a utilização de códons ("codon usage") pode ser adaptada para aquela de genes que codificam enzimas glicolíticas em *S. cerevisiae* para melhorar a conversão de xilose em leveduras (WIEDEMANN & BOLES, 2008). Por outro lado, estes genes demonstraram possuir um bom índice de adaptação de codon para a expressão em *E. coli*.

Tabela 7: Índice de Adaptação de Códon (CAI) das sequências obtidas.

	<i>E. coli</i>	<i>B. subtilis</i>	<i>S. cerevisiae</i>
B1	0,358	0,275	0,038
B2	0,343	0,291	0,055
B3	0,323	0,292	0,055
B7	0,327	0,272	0,036
B10	0,287	0,264	0,038
B20	0,285	0,251	0,032

4.7. Teste de complementação gênica

O isolado B3 foi selecionado para realização de ensaios complementação gênica e de expressão. Para isso, foi desenhado “primers” com sítios de restrição para clonagem em vetor pGEN-T (Promega) e pET28a(+) (Novagen). Para amplificação do gene de xilose isomerase do isolado B3, foram desenhados três “primers”. Os “primers” foram utilizados em reações para amplificação do gene de xilose isomerase utilizando como molde o DNA do isolado B3. O tamanho esperado do amplicom obtido com os “primers foward”-B3 e “reverse”1-B3 é de aproximadamente 750pb, enquanto que para os “primers foward”-B3 e “reverse”-2-B3 é de aproximadamente 1400 pb. Os amplicons foram purificados e clonados em vetor pGEN-T easy. Para a transformação, foi utilizada a linhagem hospedeira *E. coli* HB101, que possui o gene de xilose isomerase nocauteado (genótipo: *supE44*, $\Delta(mcrC-mrr)$, *recA13*, *ara-14*, *proA2*, *lacY1*, *galK2*, *rpsL20*, *xyl-5*, *mtl-1*, *leuB6*, *thi-1*). Os transformantes foram crescidos em meio mineral mínimo, suplementado com 1% de D-xilose. Através deste ensaio, foi possível verificar que o gene encontrado possui capacidade de restaurar a função de uma bactéria cujo gene de xilose isomerase estava nocauteado. A Fig. 24 mostra os transformantes que cresceram em meio mínimo, suplementado com de xilose, ampicilina, IPTG e X-gal,

este último foi adicionado ao meio de cultura para facilitar a visualização dos transformantes no meio de cultura.

Foram coletados e sequenciados 6 transformantes, e as sequências obtidas foram comparadas com as disponíveis no banco de dados Blastn, para confirmar a presença e identidade dos genes. Todos os transformantes sequenciados possuem o plasmídeo com o gene de xilose isomerase clonado.

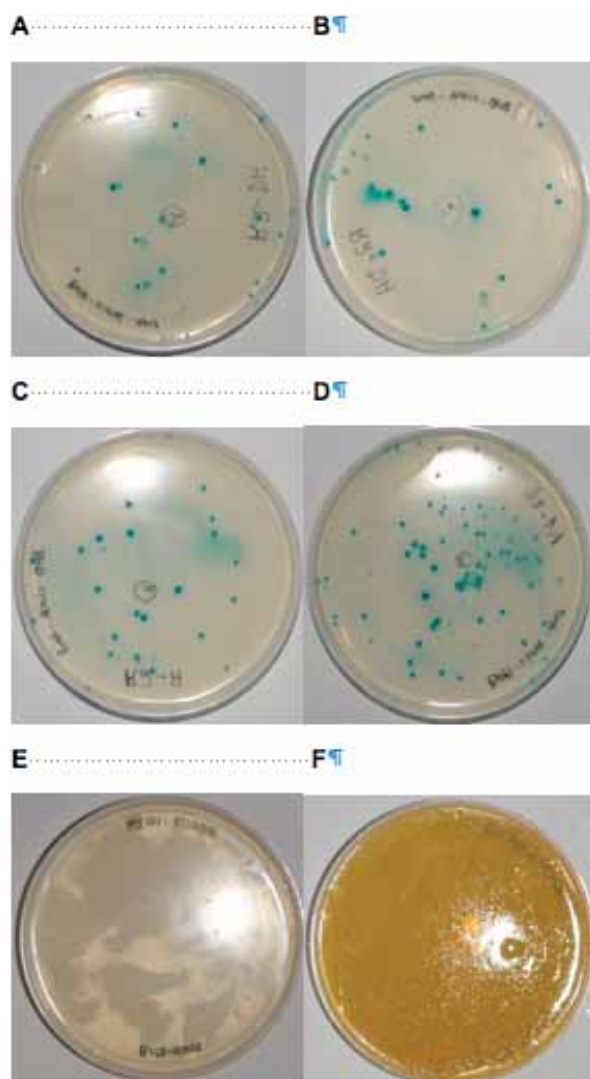


Figura 24. A, B, C e D: transformantes que cresceram em meio mínimo suplementado com xilose, ampicilina, IPTG e X-gal. E: controle negativo, com a célula HB101 e meio mínimo suplementado com xilose. F: controle positivo, com a célula HB101 e meio LB.

4.8. Caracterização molecular dos isolados de *Burkholderia*

Os isolados de *Burkholderia* que apresentaram amplificação positiva para o gene de xilose isomerase foram submetidas a extração de DNA genômico e amplificação por reação em cadeia da polimerase (PCR) do gene 16S rRNA. As sequências obtidas possuem alta qualidade, verificada através da utilização dos programas phredphrap e consed, e tamanho variando entre 400 e 600 pb tanto no sentido direto como no sentido reverso. Os isolados B7 e B20 não foram incluídos nesta análise porque não apresentaram sequências com qualidade satisfatória. O resumo dos relatórios obtidos no Blastn estão mostrados na tabela 8.

O resultado do sequenciamento parcial do gene 16S rRNA possibilitou a separação em nível de gênero, confirmando que todos os isolados pertencem ao gênero *Burkholderia*, mas não permitiu a classificação em nível de espécie, porque as sequências obtidas apresentam altos valores de similaridades com as sequências do gene 16S rRNA de diferentes espécies de *Burkholderias* disponíveis nos bancos de dados.

Tabela 8. Identificação dos isolados pela sequência parcial dos genes 16S rRNA obtida nos sentidos direto e reverso utilizando o algoritmo Blastn.

Linhagem	Verossimilhança ⁽¹⁾ 16S direto	S ⁽²⁾	ID ⁽³⁾	Frag. ⁽⁴⁾	GenBank ⁽⁵⁾	Verossimilhança ⁽¹⁾ 16S reverso	S ⁽²⁾	ID ⁽³⁾	Frag. ⁽⁴⁾	GenBank ⁽⁵⁾
B1	<i>Burkholderia gladioli</i>	800	99%	455	AB591835.1	<i>Burkholderia ambifaria</i>	874	99%	489	CP001027.1
B2	<i>Burkholderia ferrariae</i>	845	98%	481	DQ514537.1	<i>Burkholderia tropica</i>	677	98%	405	EF139183.1
B3	<i>Burkholderia tropica</i>	953	98%	553	AB252074.1	<i>Burkholderia tropica</i>	928	99%	524	FJ436053.1
B10	<i>Burkholderia gladioli</i>	798	100%	437	GU936679.1	<i>Burkholderia gladioli</i>	848	98%	484	AB024491.1

⁽¹⁾Verossimilhança, organismo que possui a sequência com a qual a sequência parcial do gene 16S rRNA da linhagem estudada apresentou maior homologia. ⁽²⁾S (Score), pontuação do alinhamento. ⁽³⁾Id (Identidade), percentagem de identidade entre a sequência das linhagens estudadas e o organismo relacionado. ⁽⁴⁾Frag. (Tamanho do fragmento), tamanho da sequência de nucleotídeos. ⁽⁵⁾GenBank, número de acesso da sequência do organismo relacionado.

Uma árvore filogenética foi construída, baseado nas sequências parciais do gene de 16S rRNA, no sentido direto, obtidas neste estudo e outras selecionadas de bancos de dados públicos. A árvore obtida está mostrada na Fig. 25. A árvore filogenética dividiu os isolados em alguns grupos: o isolado B1 formou um grupo monofilético com *B. ambifaria*, o isolado B2 formou um grupo monofilético com *B. glumae* e o isolado B3 formou um grupo monofilético com *B. viatnamiensis* e o isolado B10 formou um grupo monofilético com *B. gladioli*. Os resultados obtidos das análises do gene 16S rRNA foram consistentes com os disponíveis na literatura. COENYE & VANDAMME, em 2003, afirmaram que embora os genes ribossomais, particularmente o 16S rRNA, sejam hoje as moléculas mais utilizadas para estimar as relações filogenéticas entre as bactérias, foi constatado em *Burkholderia spp.* que a magnitude da divergência na sequência do gene 16S rDNA é limitada e insuficiente para distinguir, claramente, as espécies desse gênero.

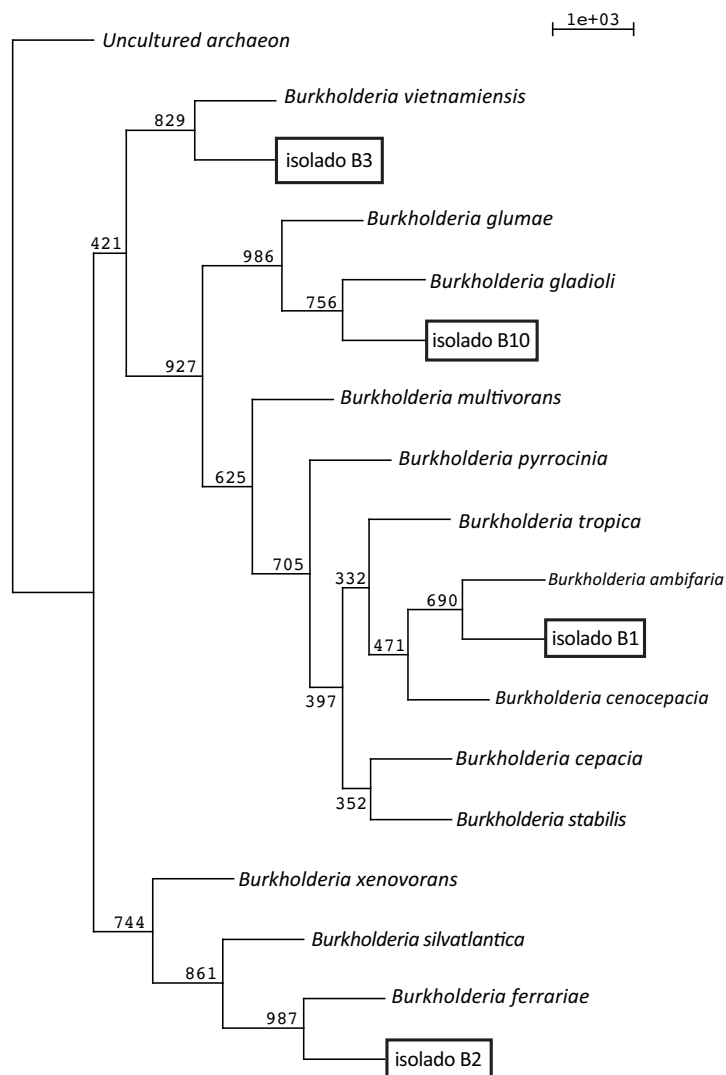


Figura 25. Análise filogenética do gene de 16S rRNA. A reconstrução filogenética foi computada pelo método de distância, utilizado o algoritmo Jukes-Cantor e o método neighbor-joining de construção de filogramas, com 1000 "bootstraps". O número de substituições de nucleotídeos é proporcional ao comprimento da escala. As sequências obtidas neste trabalho estão ilustradas em vermelho. Os números de acesso das sequências oriundas dos bancos de dados são: *B. ambifaria* CP001027.1, *B. cepacia* FJ652617.1, *B. cenocepacia* HM042678.1, *B. stabilis* AY616143.1, *B. pyrrocinia* GQ169784.1, *B. vietnamiensis* FJ436055.1, *B. ferrariae* DQ514537.1, *B. xenovorans* EU723243.1, *B. silvatlantica* AY965240.1, *B. tropica* FJ436053.1, *B. multivorans* AP009386.1, *B. gladioli* GU936679.1, *B. glumae* DQ355167.1 e Uncultured archaeon FN691750.1.

V. CONCLUSÕES

Nas bibliotecas metagenômicas, foi possível identificar genes que codificam álcool desidrogenase, oxirredutase, acil-CoA desidrogenase, luciferase, transportadores de membrana, thiolase, aminoglicosídeo fosfotransferase, desidrogenase de cadeia curta, proteínas repressoras TetR, enoil-CoA hidratase/isomerase, entre outras.

Das 13 linhagens de *Burkholderias* testadas para a presença do gene de xilose isomerase, 6 apresentaram amplificação positiva. Foi possível recuperar a sequência completa dos 6 genes identificados. As sequências foram utilizadas em análises computacionais, que permitiram estabelecer a identidade entre as sequências e a dedução da função das proteínas baseado em similaridades. O trabalho foi bem sucedido em fornecer novas sequências de genes de xilose isomerase.

A análise do índice de adaptação de códons mostrou que os genes encontrados possuem boas possibilidades de expressão em *Escherichia coli*, mas que podem apresentar uma taxa de expressão ineficiente em *Saccharomyces cerevisiae*.

O gene do isolado B3 foi escolhido para realização de ensaios de complementação gênica, e foi capaz de restaurar a função do gene de uma linhagem de *E. coli* cujo gene de xilose isomerase estava nocauteado.

O resultado do seqüenciamento parcial do gene 16S rRNA possibilitou a separação em nível de gênero, confirmando que os isolados seqüenciados pertencem ao gênero *Burkholderia*.

VI. REFERÊNCIAS

AMANN, R.I.; LUDWIG, W.; SCHLEIFER, K.H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. **Microbiological Reviews**, Washington, v. 59, p. 143-169, 1995.

AMORE, R.; WILHEM, M.; HOLLENBERG CP. The fermentation of xylose - an analysis of the expression of *Bacillus* and *Actinoplanes* xylose isomerase genes in yeast. **Applied Microbiology and Biotechnology**, v. 30, p.351–357, 1989.

BHOSALE, S. H.; RAO, M. B.; DESHPANDE, V. Molecular and industrial aspects of glucose isomerase. **Microbiological reviews**, v. 60, p. 280-300, 1996.

BORNEMAN, J.; SKROCH, P.W.; OSULLIVAN, K.M.; PALUS, J.A.; RUMJANEK, N.G.; JANSEN, J.L.; NIENHUIS, J.; TRIPLETT, E.W. Molecular microbial diversity of an agricultural soil in Wisconsin. **Applied and Environmental Microbiology**, Washington, v. 62, n. 6, p. 935-943, 1996.

BOXMA, B.; VONCKEN, F.; JANNINK, S.; VAN ALLEN, T.; AKHMANOVA, A.; VAN WEELDEN, S.W.; VAN HELLEMOND, J.J.; RICARD, G.; HUYNEN, M.; TIELENS, A.G.; HACKSTEIN, J.H. The anaerobic chytridiomycete fungus *Piromyces* sp. E2 produces ethanol via pyruvate:formate lyase and an alcohol dehydrogenase. **Molecular Microbiology**, v. 51, p. 1389–1399, 2004.

BRAT, D.; BOLES, E.; WIEDEMANN, B. Functional expression of a bacterial xylose isomerase in *Saccharomyces cerevisiae*. **Applied and Environmental Microbiology**, v. 75, p. 2304-2311, 2009.

BURKHOLDER, W. H. Sour skin, a bacterial root of onion bulbs. **Phytopathology**, Lancaster, v. 40, p. 115-117, 1950.

COENYE, T.; VANDAMME, P. Diversity and significance of *Burkholderia* species occupying diverse ecological niches. **Environmental Microbiology**, v. 5, p. 719–729, (2003).

COURTOIS, S., et al. Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. **Applied and Environmental Microbiology**, Washington v.69, n.1, p. 49-55, 2003.

COWAN, D.; MEYER, Q.; STAFFORD, W.; MUYANGA, S.; CAMERON, R.; WITTER, P. Metagenomic gene discovery: past, present and future. **Trends in Biotechnology**, London, v. 23, n. 6, p. 321-329, 2005.

EWING, B.; GREEN, P. Base-calling of automated sequencer traces using Phred. II. Error probabilities. **Genome Research**, New York, v. 8, p. 186-194, 1998.

EWING, B.; HILIER, L.D.; WENDL, M.C.; GREEN, P. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. **Genome Research**, New York, v. 8, p. 175-185, 1998.

FELSENSTEIN J. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Departament of Genome Sciences, University of Washington, Seattle, 2005.

GELSOMINO, A; KEIJZER-WOLTERS, A.C.; CACCO, G.;VAN ELSAS, J.D. Assessment of bacterial community structure in soil by polymerase chain reaction and denaturing gradient gel electrophoresis. **Journal of Microbiological Methods**, Amsterdam, v. 38, p 1-15, 1999.

GINOLHAC, A.; JARRIN, C.; GILLET, B.; ROBE, P.; PUJIC, P.; TUPHILE, K.; BERTRAND, H.; VOGEL, T.M.; PERRIERE, G.; SIMONET, P.; NALIN, R. Phylogenetic analysis of polyketide synthase I domains from soil metagenomic libraries allows selection of promising clones. **Applied and Environmental Microbiology**, Washington, v. 70, n. 9, p. 5522-5527, 2004.

GORDON, D.; ABAJIAN, C.; GREEN, P. Consed: A Graphical Tool For Sequence Finishing. **Genome Research**, New York, v. 8, p. 195-202, 1998.

GOVINDASWAMY, S.; VANE, L.M. Kinetics of growth and ethanol production on different carbon substrates using genetically engineered xylose-fermenting yeast. **Bioresource Technology**, v. 98, p. 677–685, 2007.

HAHN-HÄGERDAL, B.; KARHUMAA, K.; FONSECA, C.; SPENCER-MARTINS, I.; GORWA-GRAUSLUND, M.F. Towards industrial pentose-fermenting yeast strains. **Applied Microbiology and Biotechnology**, v. 74, p. 937–953, 2007.

HANDELSMAN, J. Sorting out metagenomes. **Nature Biotechnology**, London, v. 23, n. 1, p. 38-39, 2005.

HOLMES, A., GOVAN, J. AND GOLDSTEIN, R. Agricultural use of *Burkholderia (Pseudomonas) cepacia*: a threat to human health? **Emerging Infectious Diseases**, Atlanta, v. 4, p. 221–227, 1998.

INGRAM, L.O.; ALDRICH, H.C.; BORGES, A.C.; CAUSEY, T.B.; MARTINEZ, A.; MORALES, F.; SALEH, A.; UNDERWOOD, S.A.; YOMANO, L.P.; YORK, S.W.; ZALDIVAR, J.; ZHOU, S. Enteric bacterial catalysts for fuel ethanol production. **Biotechnology Progress**, v. 15, p. 855–866, 1999.

JEFFRIES, T.W. Engineering yeasts for xylose metabolism. **Current Opinion in**

Biotechnology, v. 17p. 320–326, 2006.

KARHUMAA, K.; GARCIA-SANCHEZ, R.; ÄGERDAL, B.; GORWA-GRAUSLUND, M.F. Comparison of the xylose reductase-xylitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant *Saccharomyces cerevisiae*. **Microbial Cell Factories**, v. 6, p. 5, 2007.

KUYPER, M.; HARHANGI, H.R.; STAVE, A.K.; WINKLER, A. A.; JETTEN, M.S.M.; DE LAAT, W.T.A.M.; DEN RIDDER, J.J.J.; OP DEN CAMP, H.J.M.; VAN DIJKEN, J.P.; PRONK, J.T. High-level functional expression of a fungal xylose isomerase: the key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae*? **FEMS Yeast Res**, v. 4, p. 69–78, 2003.

LARKIN, M.A.; BLACKSHIELDS, G.; BROWN, N.P.; CHENNA, R.; MCGETTIGAN, P.A.; MCWILLIAM, H.; , VALENTIN, F.; WALLACE, I. M.; WILM, A.; LOPEZ, R.; THOMPSON, J. D.; GIBSON, T. J.; HIGGINS, D. G. Clustal W and Clustal X version 2.0. **Bioinformatics** v. 23, p.2947-2948, 2007.

LOPES, M. S. G.; GOMES, J. G. C.; SILVA, L. F. Cloning and overexpression of the xylose isomerase gene from *Burkholderia sacchari* and production of polyhydroxybutyrate from xylose. **Canadian Journal of Microbiology**, v. 55, p. 1012-1014, 2009.

MARMUR, L.J. A producer for the isolation of deoxyribonucleic acid from microorganismis. **Journal of Molecular Biology**, v. 3, p. 208-218, 1961.

MARSHALL, R.O.; KOOL, E.R. The enzymatic conversion of d-glucose to d-fructose. **Science**, v. 125, p. 648-649, 1957.

MOES, C.J.; PRETORIUS, I.S.; VAN ZYL, W.H. Cloning and expressing of the

Clostridium thermosulfurogenes D-xylose isomerase gene (xylA) in *Saccharomyces cerevisiae*. **Biotechnology Letters**, v. 18, p. 269–274, 1996.

ORELLANA, C.; NETO, R.B. Brazil and Japan give fuel to ethanol market. **Nature Biotechnology**, v. 24, p. 232-232, 2006.

PACE, N. R. A molecular view of microbial diversity and the biosphere. **Science**, New York, v. 276, p. 734-740, 1997.

PALLERONI, N. J. Prokaryotic diversity and the importance of culturing. **Antonie Van Leeuwenhoek**, Amsterdam, v. 72, p. 3-19, 1997.

PEREIRA, R.M.; SILVEIRA, E.L.; SCAQUITTO, D.C.; PEDRINHO, E.A.N.; VAL-MORAES, S.P.; WICKERT, E.; CARARETO-ALVES, L.M.; LEMOS, E.G.M. Molecular characterization of bacterial populations of different soils. **Brazilian Journal of Microbiology**, São Paulo, v. 37, n. 4, p. 439-447, 2006.

PERRIÈRE, G.; GOUY, M. WWW-Query: An on-line retrieval system for biological sequence banks. **Biochimie**, v. 78, p. 364-369, 1996.

RICE, P.; LONGDEN, I.; BLEASBY, A. EMBOSS: The European Molecular Biology Open Software Suite. **Trends in Genetics**, v. 16, p. 276--277, 2000.

RIESENFELD, C.S.; SCHLOSS, P.D.; HANDELSMAN, J. Metagenomics: genomic analysis of microbial communities. **Annual Review of Genetics**, v. 38, p.525-552, 2004.

RONDON, M.R., AUGUST, P.R., BETTERMANN, A.D., BRADY, S.F., GROSSMAN, T.H., LILES, M.R., LOIACONO, K.A., LYNCH, B.A., MACNEIL, I.A., MINOR, C., TIONG, C.L., GILMAN, M., OSBURNE, M.S., CLARDY, J., HANDELSMAN, J., GOODMAN, R.M. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. **Applied and Environmental Microbiology**, Washington, v. 66, n. 6, p. 2541-2547, 2000.

ROSE, T.M.; SCHULTZ, E.R.; HENIKOFF, J.G.; PIETROKOVSKI, S.; MCCALLUM, C.M.; HENIKOFF, S. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly-related sequences. **Nucleic Acids Research**, v. 26, p. 1628-1635, 1998.

ROVERE, E. L.; PEREIRA, A. S.; SIMÕES, A. F., Biofuels and Sustainable Energy Development in Brazil. **World Development**, v. 39, p. 1026–1036, 2011.

ROZEN, S.; SKALETSKY, H.J. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, p. 365-386, 2000.

SALLES, J.F. *Burkholderia* community structure in soils under different agricultural management. 2005. 144 p. (Ph.D. thesis) - Leiden University, Institute of Biology Leiden, 2005.

SAMBROOK, J.; MANIATS, T.; FRITSCH, E.F. **Molecular cloning: a laboratory manual**, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989.

SARTHY, A.V.; MCCONAUGHY, B.L.; LOBO, Z.; SUNDSTROM, J.A.; FURLONG, C.E., HALL, B.D. Expression of the *Escherichia coli* xylose isomerase gene in *Saccharomyces cerevisiae*. **Applied and Environmental Microbiology**, v. 53, p. 1996–

2000, 1987.

SCHLOSS, P.D., AND HANDELSMAN, J. Biotechnological prospects from metagenomics. **Current Opinion in Biotechnology**. v. 14 p. 303-310, 2003.

SCHUCH, V. Construção de uma biblioteca metagenômica para prospecção de genes envolvidos na biossíntese de antibióticos. Dissertação. Programa de pós-graduação em microbiologia agropecuária, UNESP, Campus de Jaboticabal, 65 p., 2007.

SHARP, P.M.; LI, W.H. The codon Adaptation Index - a measure of directional synonymous codon usage bias, and its potential applications. **Nucleic Acids Research**, v. 15, p. 1281-95, 1987.

SILVEIRA, E.L.; PEREIRA, R.M.; SCAQUITTO, D.C.; PEDRINHO, E.A.N.; VAL-MORAES, S.P.; WICKERT, E.; CARARETO-ALVES, L.M.; LEMOS, E.G.M. Bacterial diversity of soil under eucalyptus assessed by 16S rDNA sequencing analysis. **Pesquisa Agropecuária Brasileira**, Brasília, v. 41, n. 10, p. 1507-1516, 2006.

SJÖSTEDT, A.; FRANCISELLA, B. In: Brenner DJ, Krieg NR, Staley JT, Garrity GM, editors. The proteobacteria, part B. Bergey's manual of systematic bacteriology. 2nd ed. New York: Springer-Verlag; 2005. p. 200-10.

SONG, S.; PARK, C. Utilization of D-ribose through D-xylose transporter. **FEMS Microbiology Letters**, v.163, p. 255–261, 1998.

STEIN, J.L., MARSH, T.L., WU, K.Y., SHIZUYA, H., DELONG, E.F. Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. **Journal of Bacteriology**, Washington, v. 178, p. 591–599, 1996.

TYSON, G.W., CHAPMAN, J., HUGENHOLTZ, P., ALLEN, E.E., RAM, R.J., RICHARDSON, P.M., SOLOVYEV, V.V., RUBIN, E.M., ROKHSAR, D.S., BANFIELD, J.F. Community structure and metabolism through reconstruction of microbial genomes from the environment. **Nature**, London, v. 428, p. 37-43, 2004.

VAN MARIS, A.J.A.; ABBOTT, D.A.; BELLISSIMI, E.; VAN DEN BRINK, J.; KUYPER, M.; LUTTIK, M.A.H.; WISSELINK, H.W.; SCHEFFERS, W.A.; VAN DIJKEN, J.P.; PRONK, J.T. (2006) Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status. **Antonie van Leeuwenhoek**, v. 90, p. 391–418, 2006.

VORONOVSKY, A.Y.; RYABOVA, O.B.; VERBA, O.V.; ISHCHUK, O.P.; DMYTRUK, K.V.; SIBIRNY, A.A. Expression of xylA genes encoding xylose isomerases from *Escherichia coli* and *Streptomyces coelicolor* in the methylotrophic yeast *Hansenula polymorpha*. **FEMS Yeast Research**, v.5, p.1055-1062, 2005.

WALFRIDSSON, M.; BAO, X.; ANDERLUND, M.; LILIUS, G.; BULOW, L.; HAHN-HÄGERDAL, B.; Ethanolic fermentation of xylose with *Saccharomyces cerevisiae* harboring the *Thermus thermophilus* xylA gene, which expresses an active xylose (glucose) isomerase. **Applied and Environmental Microbiology**, v. 62, p. 4648–4651, 1996.

WIEDEMANN, B.; BOLES, E. Codon-optimized bacterial genes improve L-arabinose fermentation in recombinant *Saccharomyces cerevisiae*. **Applied and Environmental Microbiology**, v. 74, p. 2043–2050, 2008.

YABUUCHI, E.; KOSAKO, Y.; OYAIZU, H.; YANO, I., HOTTA, H., HASHIMOTO, Y., EZAKI, T.; ARAKAWA, M. Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type

species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov. **Microbiology and Immunology**, Tokyo, v. 36, p. 1251-1275, 1992.

ZHOU, J. et al. DNA recovery from soils of diverse composition. **Applied and Environmental Microbiology**, v.62, p.316-322, 1996.