Universidade de Brasília
Instituto de Ciências Biológicas
Departamento de Biologia Celular
Programa de Pós-Graduação em Biologia Molecular

Atividades Amilolíticas Identificadas em Bibliotecas Metagenômicas da Microbiota do Solo do Cerrado

KAREN RAPP PY-DANIEL

Orientador: Prof. Fernando Araripe Gonçalves Torres

Brasília, 2010

Livros Grátis

http://www.livrosgratis.com.br

Milhares de livros grátis para download.

Universidade de Brasília
Instituto de Ciências Biológicas
Departamento de Biologia Celular
Programa de Pós-Graduação em Biologia Molecular

Atividades Amilolíticas Identificadas em Bibliotecas Metagenômicas da Microbiota do Solo do Cerrado

KAREN RAPP PY-DANIEL

Dissertação de mestrado apresentada ao Departamento de Biologia Celular do Instituto de Ciências Biológicas da Universidade de Brasília como requisito parcial à obtenção do grau de Mestre em Biologia Molecular.

Orientador: Prof. Fernando Araripe Gonçalves Torres

Brasília, 2010

KAREN RAPP PY-DANIEL

Atividades Amilolíticas Identificadas em Bibliotecas Metagenômicas da Microbiota do Solo do Cerrado

Dissertação apresentada à banca como exigência parcial para a obtenção do título de Mestre em Biologia Molecular da Universidade de Brasília

Brasília, 2010

Banca Examinadora

Prof. Dr. Fernando Araripe Torres
Universidade de Brasília
Orientador

Prof. Dr. Ricardo Henrique Krüger
Universidade de Brasília
Membro Efetivo

Prof. Dra. Natália Florêncio Martins
Embrapa Cenargen
Membro Efetivo

Prof. Dra. Lidia Maria Pepe de Moraes
Universidade de Brasília
Suplente

A todos que compõem este todo e

o tornam tudo que ele o é!

"Life is what happens to you while you're busy making other plans." (John Lennon)

"If you want to be a hero, well just follow me." (John Lennon)

Agradecimentos:

[“Beautiful, beautiful, beautiful, beautiful boy.” (John Lennon)

"I get by with a little help from my friends." (Paul McCartney; John Lennon)

"...sitting on a sofa with a sister or two." (Paul McCartney; John Lennon)

Que longa caminhada, muito bem escoltada, assessorada e tranquilizada que realizei.

As bênçãos que recebi e recebo, não sei o que fiz para merecer, mas sou um poço sem fim de gratidão.]

Ao professor Fernando, serei eternamente grata pelas oportunidades, auxílios e votos de confiança que você me proporcionou. Espero que ainda possamos trabalhar muito juntos nos intermináveis desafios da biologia molecular.

À professora Lidia, sua disposição para ajudar é inspiradora. Sem o seu constante suporte este trabalho não chegaria onde chegou. Muito Obrigada!!!

Ao grande professor Spartaco, sem você eu sequer estaria no Laboratório de Biologia Molecular. Eterna gratidão pela inspiração sem fim e portas que você me ajudou a abrir.

Colegas do Laboratório de Biotecnologia Molecular: vocês são o máximo! Agradeço muito a todos por terem conseguido me aceitar e tornar a minha rotina tão divertida. Nosso companheirismo e prestatividade são um exemplo, claro que não são perfeitos, mas pra que perfeição se ela é tão sem graça? (Viviane você é prestativa e fofa demais. Assim você nem consegue trabalhar!! Obrigada!!! Obrigada!! Obrigada!! Ju, o que seria da Bio Mol sem esse seu jeitinho recifense tão único de ser? Luanne, tutora e amiga, sem você eu provavelmente não teria durado muito no lab. Ninguém conseguiria pedir por pessoa melhor para começar a trabalhar na Bio Mol. Carol, Paula,Fernanda (já és do nosso Lab, viu?), Vivícius, Osmar, Theyssa, Bruno, Profª. Janice, Loise, Talita e Ana Lívia, valeu messssmo por tudo!!

Ao Marciano, Rafael, Alex, Daniel, Miguel e Túlio, valeu demais guris por tornarem o nosso subsolo um lugar mais alegre e divertido (e pelas muitas duvidas tiradas no caminho).

Aos Professores Ildinete, Sueli, Marcelo, Andrea e Élida, obrigada pelos muitos auxílios prestados no caminho. Ao professor Henrique Krüger por ter nos possibilitado trabalhar com as suas bibliotecas metagenômicas e seu constante auxílio. Ao professor Georgios Pappas pelo auxilio na utilização das ferramentas de bioinformática.

A todos os professores do programa de Pós-graduação em Biologia Molecular.

Ao Thompson, Dona Fátima e Dona Ivolnildes!!! (Nossos braços esquerdos, direitos, pernas e todo o resto no laboratório).

À Faculdade UnB Planaltina por ter me acolhido tão bem e ter me fornecido disponbilidade de tempo para terminar a minha pós-graduação.

Aos amigos: Fabiany, João-Maria, Lulukabio, Thalitaion, e Erlaine, peças chaves na minha existência brasiliense. A todos os amigos e colegas da UnB.

À Prof^{a}. Drª Lúcia Helena Rapp Py-Daniel, minha mãe fofa, maravilhosa e fantástica que conquistou o mundo com a sua simplicidade (com três molecas nas costas). Você é minha principal fonte de inspiração.

Ao Prof. Dr. Pyum, meu papito maravilhoso que, como fonte sem fim de conhecimento e sabedoria, me ajuda a me tornar uma pessoa melhor a cada dia que passa.

Às irmãs Anne, Tainã e Sarah por terem sobrevivido à nossa infância de constantes atritos e por serem hoje minhas melhores amigas.

À maravilhosa família Reinaldo e associados, sem vocês não teria tido metade da graça. Obrigada por tudo!!

Ao melhor marido do mundo, meu Amora. E ao melhor filho do mundo, meu Rei Arthur. Vocês são partes tão importantes de mim, que não sou mais eu sem vocês.

Ao CNPq pelo apoio financeiro.

E a todos que contribuíram neste trabalho.

[“All you need is love” (John Lennon)]

SUMÁRIO

ÍNDICE DE FIGURAS i
ÍNDICE DE TABELAS iii
ABREVIATURAS iv
RESUMO vi
ABSTRACT vii

1. INTRODUÇÃO 1
1.1. Amilases 1
1.1.1. Amido 1
1.1.2. Enzimas envolvidas na degradação do amido 2
1.1.3. Características de α-amilases 6
1.1.4. Aplicações industriais de enzimas amilolíticas 7
1.2. Estudos Metagenômicos 11
1.2.1. Microrganismos do solo 11
1.2.2. Metagenômica 13
1.2.2.1. Triagens funcionais de bibliotecas metagenômicas 14
1.2.2.2. Triagem de genes de bibliotecas metagenômicas por homologia de sequência 17
1.2.3. Utilização de metagenômica para prospecção de novas enzimas amilolíticas 18
1.2.4. Metagenômica da microbiota do solo do Cerrado 20
2. OBJETIVOS 24
2.1. Geral 24
2.2. Específico 24
3. MATERIAL 25
3.1. Microrganismos utilizados 25
3.2. Enzimas utilizadas 25
3.3. Marcadores de DNA utilizados 25
3.4. *Kits* utilizados 26
3.5. Ferramentas de bioinformática 26
3.6. Vetores 27
3.7. Meios de cultura 28

3.8. Soluções utilizadas nas preparações caseiras de DNA plasmidial 30
3.9. Análise de DNA em gel de agarose 31
3.10. Soluções de ensaio dextrinizante de amido 32
4. METODOS 33
4.1. Digestão de DNA com enzimas de restrição 33
4.2. Análise em gel de agarose 33
4.3. Purificação e eluição de fragmentos de DNA 33
4.4. Ligação de fragmentos de DNA 34
4.5. Preparação de células bacterianas competentes para choque térmico 34
4.6. Transformação de *E. coli* por choque térmico 34
4.7. Preparação de DNA plasmidial em pequena escala 35
4.8. Preparação de DNA plasmidial em média escala 35
4.9. PCR 36
4.10. Identificação de atividades amilolítica em placas. 36
4.11. Ensaio em meio líquido para detecção de atividade dextrinizante de amido 36
4.12. Sequenciamento de DNA e análise *in silico* 38
4.12.1. *Primer walking* 38
4.12.2. Análises filogenéticas 39
5. RESULTADOS 40
5.1. Biblioteca de pequenos insertos 40
5.1.1. Confirmação de atividade amilolítica de clones da biblioteca de pequenos insertos 40
5.1.2. Determinação de atividade dextrinizante de A1.1 em meio líquido 41
5.1.3. Determinação da influência de arabinose na produção de halos de hidrólise em amido 42
5.1.4. Identificação de perfil de restrição de pA1 43
5.1.5. Sequenciamento de fragmento clonado em pA1 44
5.1.6. Amplificação e clonagem do gene codificante da proteína predita 1 em pET21a 49
5.2. Biblioteca de grandes insertos 51
5.2.1. Confirmação de atividade amilolítica de clones da biblioteca de grandes insertos 51

5.2.2. Determinação de atividade dextrinizante de B1.1 em meio líquido 52
5.2.3. Determinação de influência de arabinose na produção de halos de hidrólise em amido 53
5.2.4. Identificação de perfil de restrição de pB1 54
5.2.5. Subclonagem de pB1 em pBluescript II KS 55
5.2.6. Sequenciamento de subclones pB1.2, pB1.3, pB1.4 e pB1.5 56
5.2.7. Análise filogenética 65
6. DISCUSSÃO 67
7. CONCLUSÕES E PERSPECTIVAS 77
8. REFERÊNCIAS BIBLIOGRÁFICAS 78
9. ANEXOS
Anexo 1. Marcadores de DNA utilizados 91
Anexo 2. Representação esquemática do mapa físico do vetor pCF430 e seus sítios múltiplo de clonagem 92
Anexo 3. Representação esquemática do mapa físico do vetor do fosmídio CopyControl™ pCC1FOS™ 93
Anexo 4. Representação esquemática do mapa físico do vetor pABC1 94
Anexo 5. Representação física dos vetores pBluescript II KS/SK + 95
Anexo 6. Sequência dos p*rimers* utilizados em reações PCR com suas respectivas 96
temperaturas de anelamento e sítios de restrição 97
Anexo 7. Sequência nucleotídica e de aminoácidos da proteína predita 1 do plasmídeo pA1 98
Anexo 8. Alinhamento múltiplo da sequência de aminoácidos da proteína predita 1 de pA1 com sequências similares 99
Anexo 9. Alinhamento local entre a proteína predita 1 de pA1 com a sequência do domínio catalítico de amilase *Acidobacterium* sp 101

ÍNDICE DE FIGURAS

Figura 1. Representação esquemática da organização dos resíduos de glicose nos polímeros amilose e amilopectina 1

Figura 2. Representação esquemática da ação das enzimas amilolíticas 3

Figura 3. Representação esquemática da relação entre as enzimas que compõem a família α-amilase de acordo com seus tipos de reações 4

Figura 4. Representação em fitas da estrutura do barril $(\beta/\alpha)_8$, característico da família α-amilase 6

Figura 5. Sequências de aminoácido conservados nas enzimas da família α-amilase 7

Figura 6. Representação esquemática de etapas contidas nas duas formas de triagens de bibliotecas metagenômicas 14

Figura 7. Abundância de clones de rDNA de diferentes grupos de *Eubacteria* de biblioteca de rDNA 16S da microbiota do solo do cerrado *sensu stricto* 21

Figura 8. Curva padrão de amido 37

Figura 9. Confirmação de halo de hidrólise em amido de clones da biblioteca de pequenos insertos 41

Figura 10. Representação da concentração de amido residual em solução no cultivo do clone A1.1 e do clone 41 no decorrer do ensaio 42

Figura 11. Verificação da influencia de arabinose na presença e diâmetro de halos de hidrólise de clones amilolíticos da biblioteca de pequenos insertos 43

Figura 12. Visualização em gel de agarose 1% da digestão de pA1 com enzimas de sítio único no vetor pCF430 43

Figura 13. Visualização do *contig* de fragmento metagenômico de pA1 45

Figura 14. Domínios conservados identificados pelo programa InterProScan na sequência da proteína predita 1 de pA1 47

Figura 15. Cladograma da proteína predita 1, sequências similares e amilases de microrganismos relacionados 49

Figura 16. Visualização em gel de agarose 1% de PCR de pA1 com *primers* A1-CLOF/A1-CLOR1 e A1-CLOF/A1-CLOR2 50

Figura 17. Ilustração das próximas etapas para clonagem dos *amplicons* resultantes da PCR para purificação e caracterização proteica. 51

Figura 18. Confirmação de halo de hidrólise em amido de clones da biblioteca de grandes insertos. 52

Figura 19. Representação da concentração de amido residual em solução no cultivo de B1.1 e P1 (controle negativo) no decorrer do ensaio 53
Figura 20. Verificação da influencia de arabinose na presença e diâmetro do halo de hidrólise de clone amilolítico da biblioteca de grandes insertos 54
Figura 21. Visualização em gel de agarose 0,8% da digestão de pB1 com enzimas de restrição 54
Figura 22. Estratégia utilizada para subclonar fragmentos resultantes da digestão de pB1 com *Not*I no vetor pBluescript II KS 55
Figura 23. Visualização em gel de agarose 1% da digestão dos plasmídeos pB1.2, pB1.3, pB1.4 e pB1.5 com enzima de restrição *Not*I 55
Figura 24. Visualização do *contig* de fragmento metagenômico de pB1.2 57
Figura 25. Visualização do *contig* de fragmento metagenômico de pB1.3 58
Figura 26. Visualização do *contig* de fragmento metagenômico de pB1.4 59
Figura 27. Visualização do *contig* de fragmento metagenômico de pB1.5 60
Figura 28. Cladograma da proteína predita 4 de pB1.3 e sequências similares 65
Figura 29. Cladograma da proteína predita 5 de pB1.4 e sequências similares 66
Figura 30. Cladograma da proteína predita 4 de pB1.5 e sequências similares 67

ÍNDICE DE TABELAS

Tabela 1. Trabalhos de caracterização de amilases metagenômicas 18

Tabela 2. Anotação de sequências protéicas mais similares às proteínas preditas identificadas no plasmídeo pA1 46

Tabela 3. Anotação de sequências protéicas mais similares às proteínas preditas identificadas no plasmídeo pB1.3 62

Tabela 4. Anotação de sequências protéicas mais similares às proteínas preditas identificadas no plasmídeo pB1.4 63

Tabela 5. Anotação de sequências protéicas mais similares às proteínas preditas identificadas no plasmídeo pB1.5 64

ABREVIATURAS

Asn	asparagina
Asp	aspartato
ATP	adenosina trifosfato
BAC	bacterial artificial cromossome
BrEt	brometo de etídeo
cm	centímetro
DNA	ácido desoxiribonucléico
dNTP	desoxirribonucleotídeos fosfatados
EDTA	ácido etilenodiaminotetracético
EC	Enzyme Comission
G	guanina
g	grama
g	força gravitacional
Gly	glicina
h	hora
His	histidina
IPTG	isopropil β-D-1-tiogalactopiranosídeo
kb	quilobase
KV	quiloVolts
L	litro
Lys	lisina
LB	Luria-Bertani
M	molar
min	minuto
Met	metionina
mg	miligrama

mL	mililitro
mM	milimolar
NCBI	National Center for biotechnology Information
ng	nanograma
nm	nanômetro
OD_{600}	densidade óptica a 600 nm
ORF	*open reading frame*
p	peso
pb	pares de bases
PCR	polymerase chain reaction
pH	potencial hidrogeniônico
Phe	fenilalanina
RNA	ácido ribonucléico
RNAse	ribonuclease
rpm	rotações por minuto
s	segundos
SDS	dodecilsulfato de sódio
Ser	serina
T	timina
Tris	tri (hidroximetil) aminometano
v	volume
V	Volts
v/v	volume/volume
X-Gal	5-bromo-4-cloro-3indoxil-β-D-galactopiranosídeo
ºC	graus Celsius
µg	micrograma
µL	microlitro

RESUMO

As amilases são enzimas que degradam o amido e encontram diversas aplicações nas indústrias de alimentos, de bebidas, de detergentes, têxteis e de papel. Esta enzima é produzida por diversos organismos incluindo a microbiota do solo. O solo tem sido uma excelente fonte para o isolamento de microrganismos produtores de novas biomoléculas, entretanto, estima-se que menos de 1% da microbiota do planeta é cultivável em laboratório. Técnicas recentes permitem a extração de DNA total de uma amostra ambiental, seguida da sua clonagem em microrganismos cultiváveis. O Cerrado é o segundo maior bioma da América do Sul sendo que sua biodiversidade é considerada a mais rica dentre regiões similares no mundo. Duas bibliotecas metagenômicas foram construídas a partir da microbiota do solo do Cerrado: uma de grandes insertos (~30 kb) e a segunda contendo pequenos insertos (~7 kb). O objetivo do presente trabalho foi realizar a análise de clones amilolíticos obtidos a partir destas bibliotecas e identificar os genes responsáveis por atividades detectadas em placas.

Foi realizada a extração de plasmídio de clones previamente identificados nas duas bibliotecas metagenômicas citadas e estes foram utilizados para realizar a retransformação da célula em que a biblioteca foi construída, para confirmação de fenótipo. Foi realizado ensaio para detecção de atividade dextrinizante em meio líquido com os clones que apresentaram o perfil amilolítico em placa. Os fragmentos de interesse foram diretamente sequenciados ou subclonados para posterior sequenciamento. Foi realizada a montagem e análise das sequências e estas foram comparadas com sequências depositadas no GenBank e submetidas a alinhamentos e análises filogenéticas.Identificou-se em um clone ORF relacionada com a superfamília das α-amilases, com domínios da família Glicosil hidrolase e com similaridade a sequências depositadas do gênero *Acidobacterium*. O sequenciamento parcial de um segundo clone identificou ORFs que indicam uma relação com o gênero *Burkholderia*.

A utilização de ferramentas metagenômicas permitiu acessar enzimas hidrolíticas presentes na microbiota do solo do Cerrado contornarndo a necessidade do cultivo dos microrganismos para explorar o potencial que este ambiente possui. Portanto, estas ferramentas podem e devem ser utilizadas como forma de identificarmos genes que dificilmente seriam descobertos de outra forma. O presente estudo identificou o gene de uma glicosil hidrolase. Para confirmar a aplicabilidade desta enzima em processos biotecnológicos, será realizada a clonagem do gene de interesse em vetor de expressão apropriado e realização de caracterização bioquímica.

ABSTRACT

Amylases are starch-degrading enzymes which encounter several applications in the food, beverage, detergent, textile and paper industries. This enzyme is produced by several organisms including the soil microbiote. The soil has been used as an excellent source of new microorganism that produce novel biomolecules, however, current estimates indicate that less the 1% of the planet's microorganisms are culturable in laboratory conditions. Recent techniques permit the extraction of total DNA from an environmental sample followed by cloning of its DNA in culturable microorganisms. Cerrado is South America's second biggest biome and is considered to be the most biodiverse amongst similar regions in the world. Two metagenomic libraries were built from Cerrado's soil microbiome: one of large inserts, (~30 kb) and another comprising small inserts (~7 kb). The goal of this work was to analyze amylolytic clones detected in these libraries and identify genes responsible for amylytic activity detected on plates.

Plasmid extraction from positively identified clones from both libraries was accomplished and plasmids were utilized for retransformation of cells used in library constructions to confirm phenotypes. Liquid dextrinizing assays were performed of clones with positive activity on plates for detection of amylolytic profile. Cloned fragments of interest were either directly sequenced or subcloned and then sequenced. *Contigs* were assembled, analyzed and compared with sequences in GenBank database as well as used in alignments and phylogenetic analysis. An ORF was identified with similarities to sequences of the amylase superfamily with domains of glycosyl hydrolase family similar to sequences of *Acidobacterium*. Parcial sequencing of second clone identified ORFs with similarities to sequences of *Burkholderia*.

The use of metagenomic techniques permitted access to hydrolytic enzymes present in Cerrado's microbiome soil dispensing culturing of microorganisms to explore the potential this environment possesses. Therefore, these techniques can and should be used as a form of identification of novel genes that would difficultly be accessed in another manner. This work identified a glycosyl hydrolase gene. To confirm the applicability of this enzyme in biotechnological processes, cloning of gene in appropriate expression vector and biochemical characterizations will be carried out.

1. INTRODUÇÃO

1.1. Amilases

1.1.1. Amido

O amido é o mais importante polissacarídeo de reserva do reino vegetal sendo encontrado em sementes de cereais (exemplos: milho, cevada, trigo e arroz), tubérculos (batata) e raízes (mandioca). O amido é composto por resíduos de glicose, ligados entre si por ligações *O*-glicosídicas. Esta ligação é estável em pH alcalino, mas é hidrolisada em pH ácido. No final da cadeia polimérica está presente um aldeído latente que define a extremidade redutora do polímero (Nelson & Cox, 2006).

No amido, os resíduos de glicose se organizam na forma de dois tipos de polímeros: amilose e amilopectina (Figura 1). A amilose é um polímero linear insolúvel em água fria, composto em média por 500-20.000 unidades de D-glicose unidas por meio de ligações α-1,4. A média de conteúdo de amilose nas fontes de amido encontradas na natureza é de aproximadamente 25%. O restante da molécula de amido é composto pelo polímero amilopectina que é solúvel em água e consiste de pequenas cadeias lineares de resíduos de glicose unidos por ligações α-1,4, unidos entre si por ligações α-1,6 (Buléon *et al*., 1998; [Jenkins *et al*., 1993 citado por Smith, 2001]).

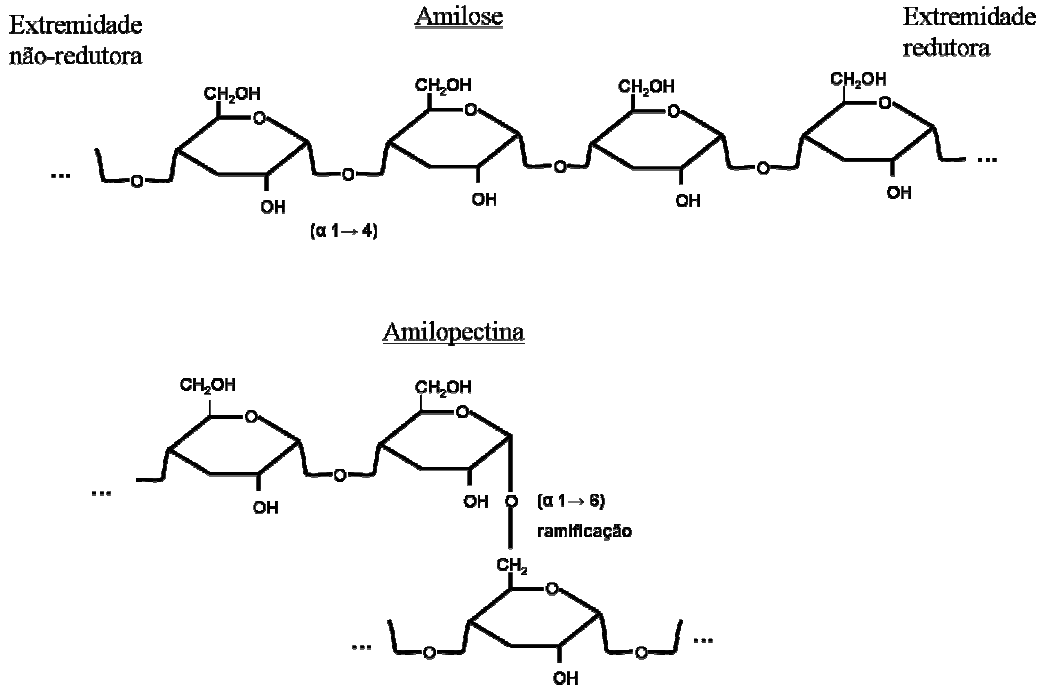


Figura 1. Representação esquemática da organização dos resíduos de glicose nos polímeros amilose e amilopectina.

1.1.2. Enzimas envolvidas na degradação do amido

As enzimas que catalisam a hidrólise ou síntese das ligações glicosídicas presentes no amido são: endoamilases, exoamilases, enzimas desramificadoras e transferases.

Endoamilases quebram as ligações glicosídicas α-1,4 presentes na parte interna das cadeias de amilose e amilopectina levando à formação de oligossacarídeos com configuração α e α-dextrinas limite (carboidratos ramificados de baixo peso molecular). A α-amilase é uma endoamilase bem conhecida que pode ser encontrada em uma grande variedade de organismos representados em todos os domínios (Pandey *et al*., 2000). A Taka-amilase A de *Aspergillus oryzae* já foi intensamente estudada e caracterizada e serve de modelo para estudos de α-amilases (Matsuura *et al*., 1984; Swift *et al*., 1991; Nagashima *et al*., 1992).

Exoamilases agem sobre os resíduos externos de glicose da amilose e amilopectina produzindo somente glicose ou maltose e β-dextrinas limite. As exoamilases podem atuar exclusivamente nos resíduos ligados em α-1,4, como a enzima β-amilase, ou em ambas as ligações, α-1,4 e α-1,6, como as enzimas glicoamilase e α-glicosidase. As glicoamilases possuem maior afinidade por polissacarídeos de cadeia longa e as α-glicosidases possuem maior afinidade por pequenos maltooligossacarídeos.

As enzimas desramificadoras podem hidrolisar exclusivamente ligações α-1,6 (isoamilase e pululanase tipo I) ou tanto ligações α-1,6 como α-1,4 (amilopululanase). A amilopululanase, por possuir tanto a atividade de α-amilase como de pululanase é também denominada de α-amilase-pululanase. A maior diferença entre isoamilases e pululanases é a habilidade de hidrolisar pululana, um polissacarídeo com unidades repetidas de maltotriose ligadas em α-1,6.

As transferases quebram as ligações glicosídicas α-1,4 da molécula doadora e transferem parte do doador para um aceptor glicosídico com a formação de uma nova ligação glicosídica. As amilomaltases e ciclodextrina glicosiltransferases formam uma nova ligação α-1,4, enquanto que a enzima ramificadora forma uma nova ligação α-1,6 (Van der Maarel *et al.*, 2002). A Figura 2 apresenta os produtos da reação das amilases citadas.

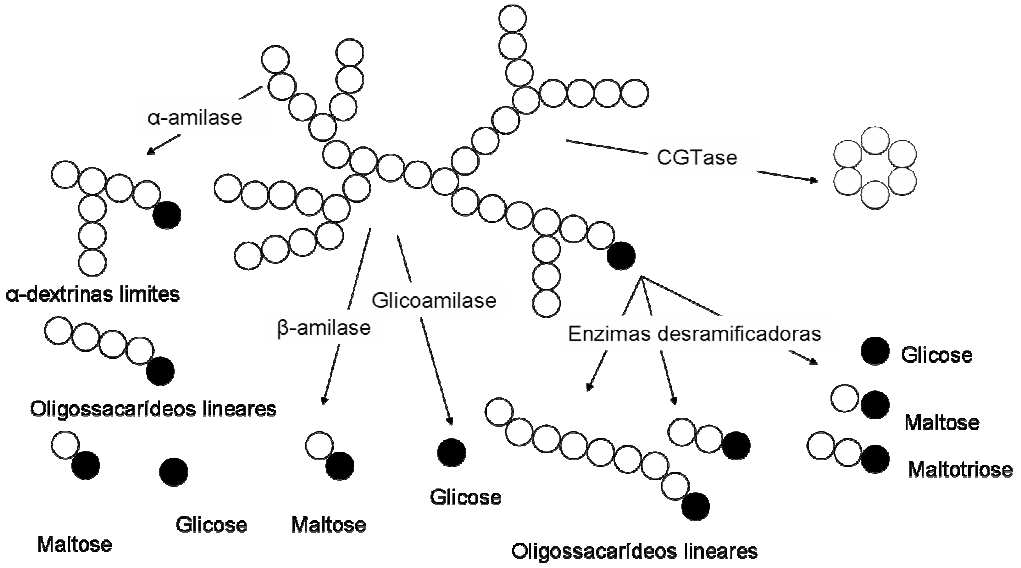


Figura 2. Representação esquemática da ação das enzimas amilolíticas. Círculos representam resíduos de glicose e círculos preenchidos destacam extremidade redutora.

Uma forma alternativa de organizar as enzimas envolvidas na degradação do amido é de acordo com o tipo de ligação em que atuam: hidrolisadoras de ligações α-1,4-glicosídicas (α-amylase, EC 3.2.1.1), hidrolisadoras de ligações α-1,6-glicosídicas (pululanase, EC 3.2.1.41 ou isoamilase 3.2.1.68), transglicosilação para formação de ligações α-1,4 (ciclodextrina glicosiltransferase - GCTase, EC 2.4.1.19) ou transglicosilação para formação de ligações α-1,6 (enzima ramificadora, EC 2.4.1.18) (Kuriki & Imanaka, 1999).

Todas as enzimas conhecidas possuem um número da Comissão de Enzimas (do inglês *Enzyme Comission* - EC) seguindo a recomendação da *Internacional Union of Biochemistry and Molecular Biology* (IUBMB, 1992). Glicosil hidrolases (endoamilases, exoamilases e enzimas desramificadoras) recebem o código EC 3.2.1.x e glicosiltransferases (transferases) recebem o código EC 2.4.1.x, onde x representa o substrato específico e em alguns casos também representa o mecanismo molecular ou o tipo de ligação. A classificação da Comissão de Enzimas se baseia em dois fatores: na reação catalisada e na especificidade do substrato, consistentes com o conceito de "chave-fechadura" proposto por Emil Fischer em 1894. Entretanto, o problema intrínseco de uma classificação baseada em especificidade de substrato ou produto é o enquadramento inapropriado de enzimas que atuam sobre vários substratos em determinada categoria (Henrissat & Davies, 1997).

Dessa forma, observa-se na literatura que as formas de agrupamento de amilases não são rígidas e excludentes uma vez que em muitos casos não há uma distinção clara entre a função de glicosil hidrolases (endo e exoamilases e enzimas desramificadoras) e de glicosil transferses (transferases). As α-amilases catalisam fracamente a transglicosilação α-1,4 em adição à sua reação principal de hidrólise de ligações α-1,4 (Hehre *et al.*, 1971, citado por Kuriki & Imanaka, 1999), assim como foram relatadas como sendo capazes de hidrolisar ligações α-1,6 (Okada & Mizokami, 1980; Sakano *et al.*, 1983; citado por Kuriki & Imanaka, 1999). CGTases hidrolisam fracamente ligações α-1,4 em adição à sua reação principal de transglicosilação α-1,4 (Kitahata & Okada, 1982, citado por Kuriki & Imanaka, 1999).

Imanaka & Kuriki (1989), identificaram uma nova pululanase-neopululanase (EC 3.2.1.135) capaz tanto de catalisar a hidrólise de ligações glicosídicas α-1,4 e α-1,6 como de transglicosilar, estabelecendo ligações glicosídicas α-1,4 e α-1,6. Considerando este histórico, Kuriki & Imanaka (1999) revisaram o conceito de família α-amilase. Na Figura 3 pode-se observar a representação esquemática utilizada pelos autores para demonstrar a relação das enzimas pertencentes a esta família.

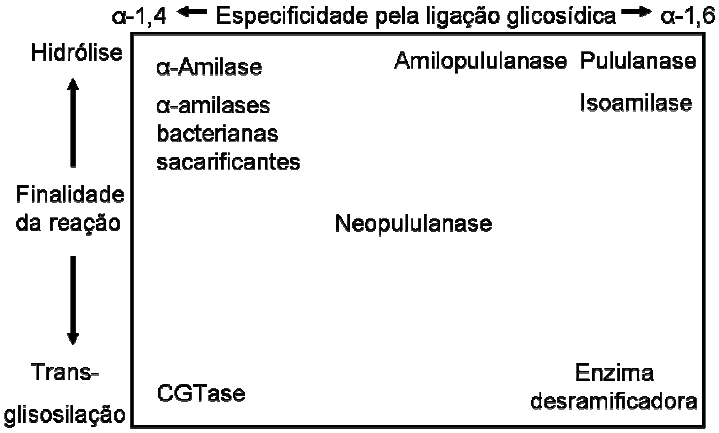


Figura 3. Representação esquemática da relação entre as enzimas que compõem a família α-amilase (adaptado de Kuriki & Imanaka, 1999)**.**

Henrissat (1991) classificou as glicosil hidrolases em famílias baseadas em homologia de sequência. De acordo com essa classificação, 301 sequências de aminoácidos de glicosil hidrolases, correspondendo a 39 números da Comissão de Enzimas, foram classificadas em 35 famílias. Dentre estas, a família 13 das glicosil hidrolases abrange pelo menos 30 especificidades enzimáticas diferentes, dentre elas a α-amilase. No entanto essa classificação apresenta alguns problemas de ambiguidade:

enzimas com diferentes especificidades de substratos são agrupadas na mesma família assim como enzimas com a mesma especificidade são agrupadas em famílias diferentes.

Coutinho & Henrissat (1999) observaram uma relação direta entre a sequência proteica e a similaridade de dobramento e propuseram uma nova classificação para as amilases baseada na homologia de sequências e em similaridades estruturais. Por serem baseadas em similaridades de sequências, estas classificações se correlacionam mais com mecanismos enzimáticos e com dobramentos protéicos do que com especificidade enzimática. Consequentemente, estas famílias são utilizadas para, conservadamente, comparar e classificar proteínas com funções não caracterizadas, cuja única característica conhecida é a similaridade de sequência, com uma enzima caracterizada experimentalmente, evitando super-predição de atividades enzimáticas (Cantarel *et al.*, 2008).

A classificação de Coutinho & Henrissat está disponível, desde 1998, no endereço eletrônico www.cazy.org, como "Enzimas Ativas em Carboidratos" (Carbohydrate-Active enZymes - CAZy). O CAZy abrange aproximadamente 300 famílias protéicas sendo 115 destas organizadas na classe de glicosil hidrolases (GHs) e 92 delas como glicosil transferases (GTs). Especificamente as α-amilases, β-amilases e glicoamilases se encontram nas famílias GH13, GH14 e GH15, respectivamente.

As amilases são proteínas com múltiplos domínios e aproximadamente 10% delas possuem um domínio distinto e não catalítico conhecido por facilitar a ligação e degradação do amido insolúvel (Lo *et al.*, 2002), apesar desta atividade já ter sido identificada em enzimas sem esse domínio (Rodriguez-Sanoja *et al.*, 2005). Esse domínio de ligação do amido (do inglês *starch binding domain* - SBD) parece se limitar às amilases de microrganismos (fungos e bactérias) e foi utilizado em 1999 por Janecek & Sevcik para demonstrar, na forma de arvore evolucionária, que os SBDs se comportam independentemente em relação aos domínios catalíticos de α-amilases, β-amilases e glicoamilases. Em outras palavras, a evolução do SBD reflete principalmente a evolução das espécies ao invés da evolução das amilases individualmente.

Atualmente, os SBDs são melhores caracterizados como parte da classificação baseada nas sequências dos módulos de ligação ao carboidrato (do inglês *carbohydrate binding module* - CBM) e têm sido classificados em sete diferentes famílias de CBMs: CBM20, CBM21, CBM25, CBM26, CBM34, CBM41, e CBM45 (Machovic & Janecek, 2006). Estas famílias também podem ser encontradas no endereço eletrônico www.cazy.org.

1.1.3. Características de α-amilases

Kuriki & Imanaka (1999) propuseram a definição da família α-amilase como enzimas que satisfizessem os seguintes requerimentos:

- Hidrolisam ligações α-glicosídicas produzindo mono ou oligossacarídeios α-anoméricos, formando ligações α-1,4 ou α-1,6 (transglicosilação), ou uma combinação de ambas as atividades;
- Os resíduos catalíticos estão contidos na estrutura de barril $(\beta/\alpha)_8$ ou Barril TIM (Fig. 4);
- Todas possuem quatro regiões altamente conservadas em sua sequência primária contendo o sítio catalítico e os mais importantes sítios de ligação ao substrato (Fig. 5);
- Todas possuem resíduos Asp, Glu, Asp no sítio catalítico correspondendo a Asp 206, Glu 230, e Asp 270 da Taka-amilase A.

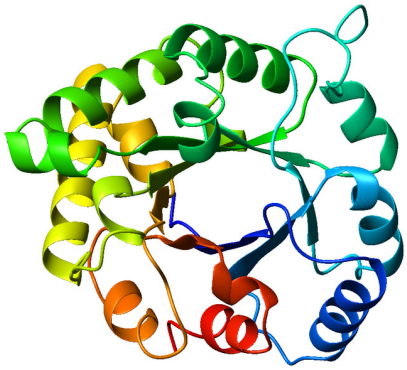


Figura 4. Representação em fitas da estrutura do barril $(\beta/\alpha)_8$, característico da família α-amilase (Fonte: RCSB Protein Data Bank, identificador: 2X1R).

1

I	116 VDVVANH 122	II	202 GLRIDTVKH 210
III	226 YCIGEVLD 233	IV	292 FVENHD 297

2

I	XDXXXNH	II	GXRXDXXZZ
III	XXX(G/A)EZZZ	IV	XXBBHD

Figura 5. Sequências de resíduos de aminoácido conservados nas enzimas da família α-amilase. 1 - sequência de aminoácidos conservados de Taka-amilase A; 2 - Sequência de aminoácidos generalizado para demais enzimas da família. X - Comumente um resíduo hidrofóbico. B - Comumente um resíduo hidrofílico. Z - Resíduo de importante especificidade.

1.1.4. Aplicações industriais das enzimas amilolíticas

As amilases encontram suas principais aplicações nas indústrias de alimentos, de bebidas, de detergentes, têxtil e de papel, chegando a representar 25% do mercado de enzimas (Saxena *et al.*, 2007).

a) Liquefação e sacarificação do amido

O maior mercado para a aplicação das α-amilases é na produção de amido hidrolisado na forma de xaropes de glicose e frutose. O xarope de frutose é um adoçante de baixa caloria e alto poder adoçante, obtido pela isomerização do xarope de glicose e utilizado principalmente na indústria de bebidas carbonatadas (refrigerantes). O processo de liquefação do amido requer o uso simultâneo de α-amilases e glicoamilases que necessariamente devem apresentar elevada termoestabilidade. (Pandey *et al.*, 2000; Van Der Maarel *et al.*, 2002).

b) Indústria de panificação

A complementação da farinha de trigo com α-amilase não somente aumenta a taxa de fermentação e reduz a viscosidade da massa, resultando em melhora no volume e textura do produto, como também gera o aumento da concentração de açúcares que melhora o gosto, coloração da crosta e qualidades do pão para assar (Gupta *et al.*, 2003)

Uma aplicação mais recente da α-amilase na indústria de panificação é na diminuição da taxa de envelhecimento do pão (Van Dam & Hille, 1992). A retrogradação do amido é um fenômeno natural decorrente de perda de água e reorganização das rígidas fibras de amilose. A retrogradação é um importante fator no envelhecimento do pão, pois seu resultado é o endurecimento do produto. A utilização de α-amilases estáveis em temperatura intermediária (do inglês *intermediate temperature stable* - ITS), ativas após a gelatinização do amido e inativas antes do término do processo de assar o pão, diminuem a retrogradação e são inativadas antes que atuem em excesso e tornem a massa pegajosa (Kulp, 1993).

Atualmente, as α-amilases utilizadas na fabricação de pães são produzidas por bactérias, principalmente da espécie *Bacillus subtilis*, e por fungos do gênero *Aspergillus*. A α-amilase proveniente de fungos possui pH ótimo ligeiramente ácido enquanto a α-amilase bacteriana possui pH ótimo de aproximadamente 7,0. Como o pH das massas de panificação tem valor ao redor de 5,0, tem-se observado que as amilases de origem fúngica são mais apropriadas para o uso em panificação (Quaglia, 1991 citado por Spier, 2008).

c) Indústria de bebidas alcoólicas

Tradicionalmente, produtores de bebidas alcoólicas permitem que os grãos de cevada germinem produzindo as enzimas necessárias para a degradação do amido em glicose para a fermentação. A cevada germinada, formando o malte, geralmente produz enzima suficiente para degradação natural do amido sem a necessidade de suplementação enzimática. No entanto, o processo natural esta sujeito à variação biológica, resultando em maior ou menor deficiência de enzimas do malte. Atualmente, a indústria de bebidas alcoólicas suplementa com amilases a produção do malte de forma a obter extratos economicamente viáveis e em menor tempo, reduzindo de 10 para quatro dias de germinação o processo de produção do malte ([Bickerstaff, 1987; Tubb, 1986] citados por Moraes, 2004).

A complementação da germinação com amilases também possui importante papel na redução de oligossacarídeos não fermentáveis, reduzindo o valor calórico da bebida alcoólica (Liu *et al*., 2004).

d) Alimentação infantil

Outro mercado para as amilases é o processo de fabricação dos cereais para alimentação infantil, tais como farinha láctea (à base de trigo), mingau de milho, arroz e misturas de cereais. As amilases são adicionadas para diminuir a viscosidade da mistura de cereais durante o processo de pré-gelatinização facilitando o escoamento pela tubulação. As amilases são inativadas com o aquecimento durante a etapa final do processo de secagem desses cereais, por isso são consideradas coadjuvantes tecnológicas de fabricação (Allbrands, 2004 citado por Spier, 2008).

e) Alimentação animal

As amilases também são empregadas para aumentar a digestibilidade das rações no sistema digestivo dos animais, diminuindo a quantidade de estrume e gases liberados ao meio ambiente. Amilases auxiliam também na alimentação de animais após o desmame por facilitar a digestão da ração por sistemas digestivos imaturos (Novozymes, 2004).

A empresa *Enzyme Development Corporation* (2010) comercializa combinações enzimáticas, dentre estas a amilases fúngicas, para auxiliar em alimentação animal com diferentes substratos.

f) Indústria Têxtil

O atual processo mecânico de produção têxtil introduz muita tensão nas fibras durante o processo de tecelagem. Os fios devem, portanto, ser protegidos para que não

ocorra sua ruptura. Uma forma de proteger os fios é pela adição de camada de amido, uma matéria prima barata, muito disponível e de fácil remoção. O acabamento dos tecidos ou roupas contendo a camada de amido é obtido pela adição de α-amilases que irão remover seletivamente a cobertura de amido e não danificar as fibras dos tecidos (Hendriksen *et al*., 1999).

g) Indústria de detergentes e produtos de limpeza

Atualmente, as enzimas constituem um dos ingredientes empregados na constituição dos detergentes. As vantagens da aplicação de enzimas em detergentes são: proporcionam condições mais suaves de uso do que em detergentes que não contêm enzimas, enzimas são 100% biodegradáveis e não são tóxicas ao meio ambiente. (Van Ee *et al*. 1992, citado por Gupta *et al*. 2003).

Alguns tipos de α-amilases nativas são sensíveis a oxidantes os quais geralmente são componentes da formulação de detergentes. A estabilidade contra os oxidantes em detergentes de uso doméstico foi alcançada pela utilização de estratégias bem sucedidas por parte das duas maiores fabricantes de enzimas, Novozymes e Genencore International, utilizando a engenharia proteica para melhorar a estabilidade das α-amilases em alvejantes (Svendsen & Bisgaard-Frantzen, 1994; Bisgaard-Frantzen *et al.*, 1995 citados por Gupta *et al,*. 2003). Trabalhando de forma independente, as duas empresas substituíram o aminoácido metionina, sensível à oxidação, na posição 197, pelo aminoácido leucina, da amilase de *Bacillus licheniformis*, o que levou a um aumento da resistência da enzima a compostos oxidativos. A melhoria na estabilidade oxidativa resultou em melhor estabilidade de armazenamento e no desempenho da enzima junto aos alvejantes contidos em detergentes. As empresas Genencore International e Novozymes introduziram estes novos produtos no mercado com os nomes de Purafect OxAm® e Duramyl®, respectivamente (Gupta *et al*., 2003).

h) Indústria de Papel e Celulose

De forma semelhante ao processamento de produtos têxteis, a cobertura do papel com amido é feita para proteger o papel contra danos mecânicos durante a sua produção. A cobertura com amido tem também função importante na melhoria da qualidade final do papel: aumenta a resistência do papel e melhora a sua apagabilidade (Bruinenberg *et al*., 1996). α-Amilases são utilizadas para diminuir a viscosidade natural da solução de amido utilizada, mantendo a viscosidade constante, de acordo com o tipo de papel desejado. A dextrinização parcial do amido nas indústrias papilíferas ocorre na faixa de 45-60ºC e atualmente existem no mercado as enzimas Amizyme®

(PMP Fermentation Products, Peoria, EUA), Termamyl®, Fungamyl, BAN® (Novozymes, Dinamarca) e α-amilase G9995® (Enzyme Biosystems, EUA) que são utilizadas para esta finalidade (Gupta *et al.*, 2003).

i) Produção de bioetanol a partir de amido

O amido é um substrato para a produção de etanol de alto rendimento. Entretanto, é necessária a sua hidrólise em açucares para que ocorra a fermentação. A hidrólise do amido era tradicionalmente obtida pelo tratamento com ácidos, mas a especificidade das enzimas, suas condições de reações inerentemente moderadas e a ausência de reações secundária, tornaram as amilases os catalisadores mais utilizados na geração de açúcares para a fermentação. As α-amilase mais utilizadas com esta finalidade são da bactéria termorresistente *Bacillus licheniformis* e de linhagens geneticamente modificadas de *Escherichia coli* e *B. subtillis* (Sánchez & Cardona, 2008).

Considerando o grande potencial do amido para a produção de bioetanol, o grupo de Biotecnologia Molecular da Universidade de Brasília vem desenvolvendo há 25 anos pesquisas com amilases para a hidrólise do amido da mandioca (Astolfi-Filho *et al.*, 1986; Ribeiro-dos-Santos, 1988; Moraes *et al.*, 1995; Moraes *et al.*, 1999; Borges, 2009).

1.2. Estudos Metagenômicos

1.2.1. Microrganismos do solo

O solo é um ambiente complexo e considerado o maior reservatório da diversidade genética microbiana. A complexidade da diversidade microbiana é resultante da interação de fatores como pH, conteúdo de água, estrutura do solo, variação climática e atividade biótica (Robe *et al*., 2003). Ward *et al*. (1990) realizaram uma análise cultura-independente da microbiota de uma fonte termal, identificando sequências do RNA ribossomal 16S, e confirmaram a especulação da grande diversidade de microrganismos não-cultiváveis presentes neste ambiente. Estimativas atuais sugerem que 4 – 6 x 10^{30} células microbianas estejam presentes nos solos do mundo (Sleator *et al*., 2008) e Aman *et al.* (1995) afirmam que menos de 1% desta microbiota é cultivável em laboratório.

Para explicar essa incapacidade de cultivar em laboratório uma parte tão representativa da microbiota dos solos, existem duas principais hipóteses. A primeira hipótese sugere que as condições a que os microrganismos estão submetidos provocam mudanças fisiológicas específicas para aquele ambiente, ressaltando que não necessariamente os microrganismos não-cultiváveis sejam filogeneticamente distantes dos microrganismos cultiváveis (Borneman, 1999).

Dentro das condições ambientais específicas sugeridas nesta hipótese, podemos atribuir ao *quorum sensing* e à formação de biofilmes um importante papel na constituição de microambientes:

- Segundo Fuqua *et al.* (1994) determinados comportamentos bacterianos só podem ser realizados de forma eficiente quando a população atinge um número suficientemente de indivíduos, descrevendo essa unidade mínima comportamental como um quórum de bactérias. Essa identificação do número de indivíduos (*quorum sensing*) é realizada pela emissão e detecção de pequenos elementos difundidos no meio, cuja concentração é crucial para a coordenação do comportamento. Portanto, possivelmente a remoção de indivíduos de um meio onde a concentração desses elementos defina rotas metabólicas, pode ser um fator importante na impossibilidade de cultivá-los em laboratório.
- Os biofilmes bacterianos são comunidades sésseis de microrganismos que coexistem como associações altamente diferenciadas em uma matriz extracelular produzida pelas células constituintes. A formação desta complexa estrutura multicelular requer comunicação célula-a-célula através de mensageiros moleculares (*quorum sensing*)

ou por contato direto célula-a-célula. A matriz do biofilme é tipicamente constituída por polissacarídeos extracelulares e pode conter proteínas e DNA, mas a estrutura precisa e composição dos biofilmes varia com as espécies componentes e as condições ambientais. Os biofilmes provêem um ambiente mecanicamente estável e protetor para bactérias, resultando em maior tolerância a condições extremas de pH e temperatura (Dickschat, 2010).

A segunda hipótese sugere que os microorganismos não-cultiváveis são filogeneticamente distantes dos cultiváveis e seus requerimentos nutricionais desconhecidos e distintos impossibilitam a sua cultura em laboratório (Hugenholtz & Pace, 1996). Rondon *et al.* (1999) afirmam que a combinação de ambas as hipóteses podem ajudar a decifrar a discrepância que observamos ao contar o número de células viáveis de determinada amostra de solo e o quanto desta microbiota conseguimos cultivar em laboratório.

Os seres humanos têm recorrido ao solo para obtenção de novas biomoléculas há décadas. As etapas tradicionais envolvidas na obtenção destas biomoléculas incluem a triagem de uma grande variedade de microrganismos cultiváveis na busca da atividade desejada, purificação e caracterização da biomolécula de interesse. Porém, recentemente, o rendimento de novos produtos naturais do solo tem sido baixo, em parte porque os mesmos microrganismos têm sido identificados recorrentemente em culturas (Handelsman *et al.*, 1998).

Durante as últimas duas décadas o desenvolvimento de métodos para isolamento de ácidos nucleicos de amostras ambientais possibilitou o acesso a uma diversidade microbiana desconhecida. A análise dos ácidos nucleicos diretamente extraídos em amostras ambientais permite aos pesquisadores estudar comunidades microbianas naturais sem a necessidade do cultivo (Riesenfeld *et al.*, 2004).

A nova era da ecologia microbiana iniciou-se quando o sequenciamento de RNA ribossomal (rRNA) e os genes que os codificavam foram apresentados para descrever as bactérias não cultivadas no meio ambiente. O primeiro trabalho nessa área sequenciou o DNA de clones de biblioteca de rRNA 5S obtidos da comunidade simbiótica do verme *Riftia pachyptila* (Stahl *et al.*, 1984 citado por Riesenfeld *et al.*, 2004). Vários métodos similares geraram um conjunto de técnicas independentes de cultivo para: reconstruir filogenias, comparar a distribuição microbiana entre amostras utilizando sequências nucleotídicas ou polimorfismo de fragmentos de restrição (RFLPs, do inglês *restriction fragment length polymorphisms*) e também para quantificar a abundância relativa de

cada grupo taxonômico utilizando hibridização de membrana ou hibridização fluorescente *in situ* ([Head *et al*., 1998; Kent *et al*., 2002; Pace, 1997; Pace *et al*., 1985; Pace *et al*., 1986; Amann *et al*., 1995] citados por Riesenfeld *et al*., 2004).

Portanto, a taxonomia baseada na filogenia, incluído os dados de microrganismos não-cultivados, está mudando rapidamente e substituindo a taxonomia baseada exclusivamente em parâmetros morfológicos, fisiológicos e bioquímicos de microrganismos cultivados (Garrity, 2005).

Entretanto, descrever a diversidade filogenética dos microrganismos não-cultivados é somente o primeiro passo, o principal desafio será designar seus respectivos papeis ecológicos. A microbiota não cultivada deve possuir papel chave nos processos ambientais naturais e representa um grande recurso inexplorado para aplicações biotecnológicas.

1.2.2. Metagenômica

Em 1998, Handelsman *et al*., denominaram o genoma microbiano coletivo de uma amostra ambiental de "metagenoma" e a análise funcional ou baseada em sequência decorrente deste metagenoma tem sido principalmente denominada de "metagenômica" (Riesenfeld *et al*., 2004). O termo metagenoma faz referência ao conceito estatístico de *meta*-análise, processo de combinar estatisticamente análises separadas, correspondendo à combinação da análise de múltiplos genomas (Handelsman *et al*., 1998).

A metagenômica desvenda a informação contida no DNA metagenômico pelo isolamento do DNA diretamente da amostra ambiental, clonagem deste DNA em vetores e utilização destes vetores para transformação de células facilmente cultiváveis como a *E. coli*. Os transformantes decorrentes da transformação constituem uma biblioteca metagenômica, contendo o DNA presente em amostras ambientais.

Atualmente, as bibliotecas metagenômicas são submetidas a triagens na busca de sequências gênicas homólogas a sequência conhecida ou pela detecção de atividade/função de interesse (Schmeisser *et al*., 2007). A Figura 6 apresenta as duas abordagens de triagens esquematizadas.

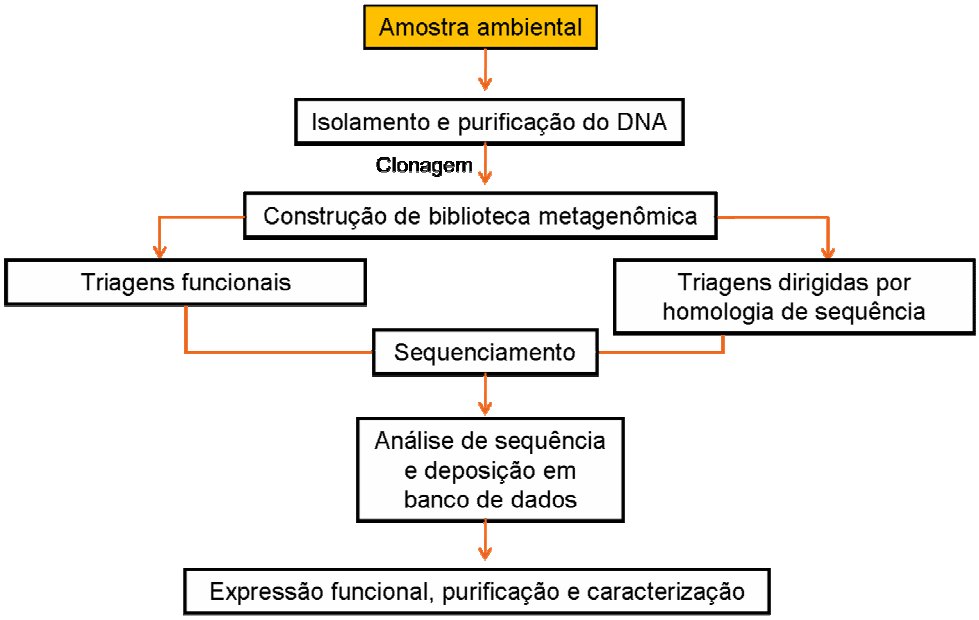


Figura 6. Representação esquemática de etapas contidas nas duas formas de triagens de bibliotecas metagenômicas. Etapas de triagens funcionais e triagens por homologia de sequência.

1.2.2.1. Triagens funcionais de bibliotecas metagenômicas

A triagem baseada em atividade é a forma mais direta de se obter genes com a atividade desejada, sendo aplicada principalmente a estudos biotecnológicos. A probabilidade de determinado gene ser identificado pela triagem funcional de uma biblioteca metagenômica depende dos seguintes fatores: a) sistema hospedeiro-vetor; b) tamanho do gene alvo; c) abundância do gene na biblioteca; d) método do ensaio sendo utilizado; e e) eficiência da expressão heteróloga (Uchiyama & Miyazaki, 2009).

a) Sistema hospedeiro-vetor: várias linhagens altamente competentes de *E. coli* estão disponíveis comercialmente e são as células hospedeiras mais utilizadas por pesquisadores para a construção de bibliotecas metagenômicas. Esta escolha se deve às deleções de genes para a recombinação homóloga (*rec*A e *rec*BC) e de sistemas de restrição (*mcr*A e *mcr*BC) presentes nessas linhagens que tornam a célula mais receptiva a DNA modificado e exógeno. A escolha do vetor depende do tamanho dos insertos a serem clonados. Existem quatro classes de vetores sendo utilizados na construção de bibliotecas (o tamanho do fragmento suportado pelo vetor está discriminado entre parêntesis): plasmídeos (até 10 kb), fosmídios (25-40 kb), cosmídios (25-35 kb) e BACs (100-200 kb).

b) O tamanho do gene alvo: quanto menor o gene, maior a probabilidade de estar contido na biblioteca de forma funcional.

c) Abundância do gene na biblioteca: espera-se que quanto maior o número de cópias do gene na biblioteca mais facilmente será sua detecção. Entretanto, Simon *et al*., (2009) compararam a frequência de detecção de um gene de interesse em duas bibliotecas, uma fosmídial e outra plasmidial, construídas a partir do mesmo metagenoma e observaram em ambas as bibliotecas uma quantidade aproximada de clones positivos por número de clones analisados. Os plasmídeos, por terem um maior número de cópias por célula e por possuírem promotores fortes, facilitam a detecção do gene de interesse, entretanto, uma biblioteca de plasmídeo precisa ter mais clones analisados para identificar os clones positivos. Ao mesmo tempo bibliotecas fosmidiais possuem clones mais representativos do material genético, consequentemente exigindo a análise de um número menor de clones, entretanto, estes são mais dificilmente detectados. Mayumi *et al*. (2008) observaram ainda que a contaminação da amostra com o substrato dos genes de interesse pode aumentar a abundancia de genes desejados na biblioteca.

d) Métodos do ensaio sendo utilizado: as atividades enzimáticas de interesse em bibliotecas metagenômicas são normalmente detectadas em placas, com ágar e o substrato da enzima de interesse, na forma da presença de halo ou coloração ao redor de clones positivos. Três formas de facilitar a detecção da atividade de interesse são: exposição do substrato aos clones da biblioteca após lise das células, facilitando a ação de enzimas contidas dentro das células hospedeiras (Suenaga *et al.*, 2007); associação da sobrevivência da célula hospedeira à identificação da atividade de interesse (Mori *et al.*, 2008; Kazimierczak *et al*., 2009; Simon *et al*., 2009); e utilização de genes repórteres (Guan *et al*., 2007; Uchiyama *et al.*, 2005).

O método desenvolvido por Uchiyama *et al.* (2005) de triar bibliotecas metagenômicas na busca de vias metabólicas utilizando expressão gênica induzida por substrato (do inglês *substrate-induced gene expression screening*- SIGEX) foi uma forma elegante e eficiente de otimizar a triagem funcional. Esta triagem se baseia no fato de genes envolvidos em vias metabólicas normalmente estarem organizados em *operons* que são induzidos pelos substratos da via. Os autores se aproveitam desta propriedade para clonar DNA metagenômico *upstream* do gene *gfp* (do inglês *green fluorescent protein*), colocando a expressão da proteína GFP sobre o controle de promotores contidos no DNA metagenômico. Em seguida, ocorre a identificação de clones que tiveram sua expressão induzida pelo substrato observando a fluorescência de clones somente na presença do substrato.

e) Eficiência da expressão heteróloga: a utilização da maquinara da célula hospedeira, normalmente *E. coli*, é a principal limitação da triagem funcional. Esta limitação se deve ao fato que dificilmente a célula hospedeira consegue reconhecer os sinais transcricionais e traducionais contidos no metagenoma. Gabor *et al.* (2004) realizaram análise computacional de 32 genomas de procariotos na busca por sinais de expressão funcionais em *E. coli* e sugeriram que aproximadamente 40% das atividades enzimáticas podem ser prontamente detectadas na clonagem aleatória in *E. coli*. Warren *et al.,* (2008), utilizaram RT-PCR e microarranjo para quantificar a proporção do genoma de *Haemophilus influenzae, Pseudomonas aeruginosa e* de DNA humano que era transcrito em *E. coli*: 50% do genoma de *H. influenzae*, menos de 30% do genoma de *P. aeruginosa* e transcrição pontual do DNA humano.

Algumas alternativas propostas para melhorar a expressão em *E. coli* são: realizar modificações na maquinaria transcricional e traducional da célula (Bernstein *et al.,* 2007) e realizar a introdução de chaperonas que ajudem do dobramento de proteínas heterólogas (Ferrer *et al*., 2004).

Observa-se também a tentativa de facilitar a identificação de genes de interesse em bibliotecas metagenômicas pela utilização de *E. coli* somente para a construção da biblioteca e de outros hospedeiros para a identificação dos clones com as atividades de interesse. Wang *et al*. (2007) usaram o vetor bifuncional *E. coli-Rhizobium* para construir biblioteca metagenômica e observaram diferentes perfis de expressão dependendo do hospedeiro utilizado: *E. coli* ou *R. leguminosarum*. Ono *et al*. (2007) utilizaram para construção de biblioteca metagenômica cosmídio com duas origens de replicação, ColE1 e RK2, capaz de se replicar tanto em bactérias Gram-positivas como Gram-negativas.

Já Martinez *et al*. (2004) aumentaram a abrangência de triagens funcionais pela utilização de linhagens de *Streptomyces lividans* e *Pseudomonas putida* como células hospedeiras para construção e triagem de biblioteca metagenômica.

1.2.2.2 Triagem de genes de bibliotecas metagenômicas por homologia de sequência

A triagem baseada em homologia de sequência consiste na procura de genes em bibliotecas metagenômicas pela utilização de *PCR* ou hibridização de sondas utilizando *primers*/sondas degeneradas baseados em sequências consenso de genes homólogos. Como esse método tem como alvo a sequência interna de determinado gene, é preciso clonar as regiões flanqueadoras para obter a sequência completa do gene (Uchiyama & Miyazaki, 2009).

Esta forma de triagem tem como principal limitação a impossibilidade de identificarmos genes efetivamente novos, uma vez que essa técnica se baseia em semelhanças de sequência com genes já conhecidos (Morimoto & Fujii, 2009; Tang *et al*., 2008; Marzorati *et al*., 2007).

Alguns métodos têm sido otimizados para realizar o *gene walking* de bibliotecas metagenômicas (Xu *et al*., 2008; Chae *et al*., 2008; Fieseler *et al*., 2007). Elsied *et al*., (2007) e Koenig *et al*. (2009) utilizaram as sequências conservadas de *integrons* para a elaboração de *primers* a fim de sequenciar genes flanqueadores.

O sequenciamento maciço de metagenomas já foi realizado em amostras do mar de Sargasso (Venter *et al*., 2004) e de biofilme de água residual ácida (Tyson *et al*., 2005) e avanços recentes em sequenciamento automático sugerem que esta técnica tende a ser cada vez mais aplicada a metagenômas (Margulies *et al*., 2005; Goldberg *et al*., 2006).

Uchiyama & Miyazaki (2009) atribui à biologia sintética o futuro papel de permitir abordagens mais radicais para a triagem baseada em homologia de sequência: será possível obter todos os possíveis genes candidatos para determinada atividade de bancos de dados, sintetizar esses genes e testar suas atividades.

Os recentes avanços mais promissores na triagem de bibliotecas metagenômicas por sequências homólogas foram advindos da bioinformática (Schmeisser *et al*., 2007). Um destes avanços, decorrente de projetos metagenômicos de sequenciamento em larga escala (Venter *et al*., 2004, Tyson *et al*., 2005) foi o desenvolvimento de programas computacionais como o DOTUR, desenvolvido para determinar quando uma biblioteca tem genes suficientes para ser considerada representativa da diversidade microbiana da amostra (Schloss & Handelsman, 2005).

1.2.3. Utilização de metagenômica para prospecção de novas enzimas amilolíticas

A metagenômica tem se desenvolvido na última década de forma a ser aplicada a duas grandes vertentes: na elucidação do genoma dos microrganismos não-cultiváveis, visando esclarecer a ecologia microbiótica, e na procura por novas enzimas e biomoléculas de interesse biotecnológico (Schmeisser *et al.,* 2007).

Já foram realizadas triagens de bibliotecas metagenômicas por uma grande variedade de enzimas bio-catalíticas (Streit *et al.*, 2004). Vários relatos da literatura reportam a detecção e caracterização de amilases a partir de bibliotecas metagenômicas: (Ferrer *et al.*, 2007; Richardson *et al.*, 2002; Rondon *et al.*, 2000; Voget *et al.*, 2003, Yun *et al.*, 2004; Sharma *et al*., 2010). Ao todo, foram identificadas 15 amilases, sendo seis caracterizadas (Tabela 1).

Tabela 1. Trabalhos de caracterização de amilases metagenômicas.

Fonte	Vetor	Número de clones triados	Tamanho de inserto	Tamanho total	Número total	Nível de caracterização	Referências
Solo	BAC	3.648	27 kb	100 Mb	8	Análise de restrição	Rondon *et al*., 2000
Várias	Plasmídio	50.000	Não especificado	Não especificado	3	Purificação protéica, pH,°C	Richardson *et al*., 2002
Solo	Cosmídio	1.532	32,5 kb	49,8 Mb	1	Sequenciamento	Voget *et al*., 2003
Solo	Plasmídio	30.000	3,5 kb	105 Mb	1	Purificação protéica, pH,°C	Yun *et al*., 2004
Intestino bovino	Vetor fago λ	14.000	5,5 kb	77 Mb	1	Atividade do extrato bruto	Ferrer *et al*., 2005, 2007
Solo	Cosmídio	90.700	30-45 kb	122 Mb	1	Purificação protéica parcial, pH,°C	Sharma *et al*., 2010

Rondon *et al*. (2000) identificaram oito clones com propriedades amilolíticas de bibliotecas metagenômicas da microbiota de solo e análise de restrição dos clones confirmou que os clones são decorrentes de eventos de clonagem independentes e não o resultado de clones repetidos.

Yun *et al.* (2004) construíram biblioteca metagenômica da microbiota do solo da província de Kuwon na Coreia do Sul e identificaram amilase com característica de ser estável e ativa em condições alcalinas com um pH ótimo de 9.0.

Richardson *et al*. (2002) construíram biblioteca metagenômica a partir de diversas amostras ambientais e submeteram essa biblioteca a triagens por atividade

amilolítica em condições ótimas típicas do processo de trituração úmida de milho: pH 4,5 e 95°C. Após triagem inicial, 15 clones demonstraram atividade em uma das duas condições pré-estabelecidas. Subsequentes caracterizações indicaram três clones com propriedades excepcionais que foram seqüenciados e suas sequências utilizadas para realização de rearranjo gênico. Das quimeras obtidas, uma amilase demonstrou excepcional compatibilidade com o processo de trituração úmida de milho.

Ferrer *et al*. (2007) construíram biblioteca metagenômica da microbiota de rúmen bovino e identificaram enzima pertencente à família das α-amilases com atividade sobre ligações α-D-(1,4) e α-D-(1,6) e atividade de transglicosilação. A enzima possui amplo espectro de atividade, de pH de 5.5 a 9.0 e manteve 80% da sua atividade a pHs 5.0 e 9.5 a 75°C.

Voget *et al.* (2003) construíram biblioteca metagenômica de cultura enriquecida em laboratório e identificaram por sequenciamento vários genes de enzimas biocatalíticas, dentre elas, gene muito similar ao gene que codifica a α-amilase de *Dictyoglomus thermophilum* e a atividade amilolítica do clone foi confirmada. A utilização de cultura enriquecida deve ter levado a perda de diversidade, entretanto o rendimento de enzimas biocatalíticas foi maior do que o normalmente obtido em triagens funcionais.

Sharma *et al*. (2010) construíram biblioteca metagenômica da microbiota de solo e identificaram gene com propriedades amilolíticas muito similar a gene de bactéria não-cultivada. A enzima possui atividade ótima a 40°C e pH 7.0 e mantém mais de 90% da sua atividade a baixas temperaturas, característica incomum a amilases de microrganismos cultivados e não-cultivados.

Entretanto, Steele *et al.* (2009) afirmam ainda não haver enzimas derivadas de estudos metagenômicos em aplicações biotecnológicas. Estipulam que esta aplicação deve ocorrer somente quando forem encontradas enzimas metagenômicas que tenha acesso a novos substratos ou que atuem em condições sob as quais as enzimas clássicas não atuam.

1.2.4 Utilização de metagenômica na microbiota do solo do Cerrado

O Cerrado é o segundo maior bioma da América do Sul, ocupando mais de 200.000.000 hectares. O clima é sazonal, com duas estações bem definidas: seca e chuvosa (Pinto, 1990). A combinação de estações sazonais com solos pobres em nutrientes, ricos em alumínio e queimadas frequentes, infligem ao Cerrado pressões ambientais que resultaram em um amplo espectro de adaptações das suas regiões fitogeográficas (Ratter *et al.*, 1997). A biodiversidade do Cerrado, tanto no âmbito de sua flora como da sua fauna, é considerada a mais rica dentre regiões similares no mundo (Pinto, 1990), sendo classificada pela *Conservation International* como um *hotspot* de diversidade biológica (Silva & Bates, 2002). Entretanto a diversidade biológica do cerrado está cada vez mais ameaçada, tendo em vista que nas últimas décadas, este bioma tem sofrido um rápido processo de urbanização e ocupação agrícola.

Estudos de diversidade microbiana do solo do Cerrado têm sido realizados utilizando ferramentas de isolamento de material genético dos microrganismos do solo. Quirino *et al.* (2009) utilizaram abordagens baseadas em DNA ribossomal (rDNA) 16S para estudar a comunidade bacteriana associada ao solo do cerrado nativo (*sensu stricto*) e de área de cerrado convertida a pastagem. O grupo bacteriano mais abundante no solo do cerrado *sensu stricto* foi de α-Proteobacteria e de Actinobacteria na área de pastagem. A Figura 7 demonstra a abundância dos clones das duas áreas amostradas de acordo com a organização de grupos bacterianos.

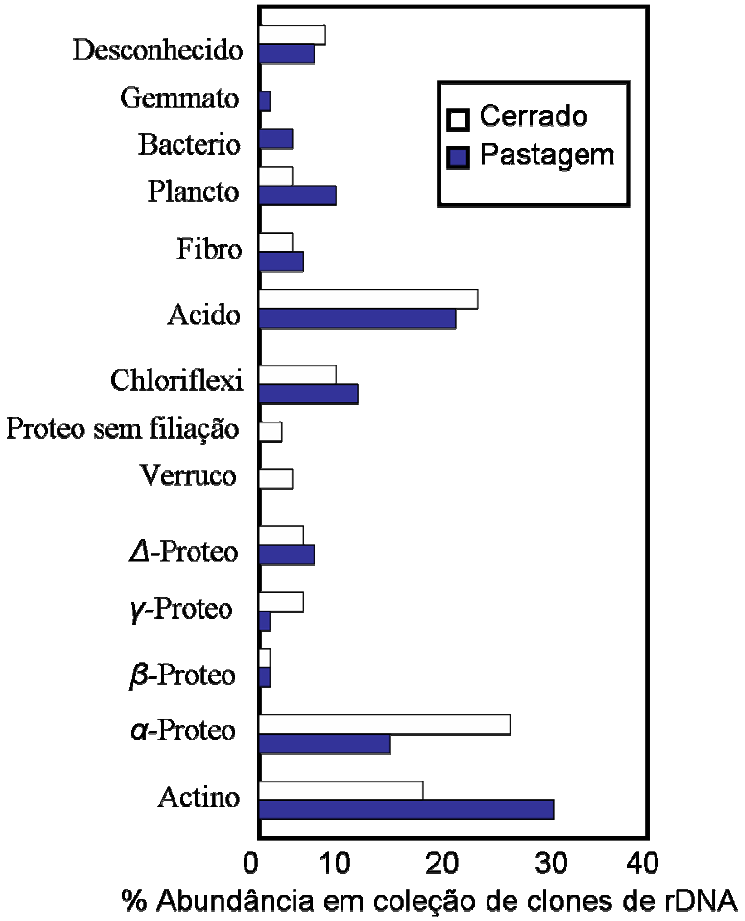


Figura 7. Abundância de clones de rDNA de diferentes grupos de *Eubacteria* em biblioteca de rDNA 16S da microbiota do solo do cerrado *sensu stricto*, em barras preenchidas e de área de pastagem (barras vazias). As barras indicam a percentagem de clones de rDNA dentro de cada grupo filogenético: bactérias do solo desconhecidas (desconhecido), gemmatobacteria (gemmato), bacteriodetes (bacterio), planctomycetes (plancto), fibrobacteria (fibro), acidobacteria (acido), chloriflexi, verrucobacteria (verruco), proteobacterias sem filiação (proteo sem filiação), *Δ*-proteobacteria (*Δ*-proteo), *γ*-proteobacteria (*γ*-proteo), *β*-proteobacteria (*β*-proteo), *α*-proteobacteria (*α*-proteo), actinobacteria (actino). Grupos com a ausência de barras indicam que clone daquele grupo filogenético foi identificado (adaptado de Quirino *et al*., 2009).

De Castro *et al*. (2008) utilizaram várias técnicas moleculares para caracterizar a comunidade de fungos presentes no solo em quatro diferentes áreas do bioma Cerrado: cerrado *sensu stricto*, pastagem, mata de galeria e plantação de soja. A técnica *Ribossomal RNA Intergenic Spacer Analysis* (RISA) revelou diferenças consideráveis entre a comunidade fúngica nas diversas áreas estudadas, onde padrões únicos de bandas de DNA para cada amostra foram identificados em cada área. Posteriormente, foram utilizados *primers* específicos para rDNA de fungos para amplificar pela reação em cadeia da polimerase (do inglês *polymerase chain reaction* - PCR) a região

conservada de rDNA 18S. Os produtos de PCR obtidos foram clonados e usados para construir quatro bibliotecas de rDNA 18S e identificaram principalmente o decréscimo da diversidade da biblioteca referente à plantação de soja.

1.2.5 Biblioteca metagenômica da microbiota do solo do Cerrado

De Castro (2008) construiu duas bibliotecas metagenômicas com a clonagem de grandes e pequenos insertos de DNA diretamente isolados do bioma Cerrado em *E. coli* usando o fosmídeo pCC1FOS (Epicentre) e o plasmídeo pCF430 como vetores de clonagem.

Para determinar a variabilidade dos genomas presentes nas bibliotecas metagenômicas e validar suas construções, duas abordagens foram utilizadas:

- Sequenciamento direto dos clones através da utilização de *primers* desenhados a partir das sequências localizadas na porção final dos vetores utilizados, sequenciando as pontas finais dos insertos de DNA.
- Construção de biblioteca 16S rDNA a partir das bibliotecas metagenômicas

Ambas as análises confirmaram a informação genômica presente nos clones das bibliotecas metagenômicas. A biblioteca metagenômica clonada em fosmídeo possui cerca de 65,000 clones com uma mediana de inserto de DNA de 35 Kb. A biblioteca metagenômica clonada em pCF430 tem cerca de 150,000 clones com uma mediana de inserto de DNA de aproximadamente 8 Kb. Os grupos bacterianos mais representados nas bibliotecas metagenômicas foram: Actinobactérias, Proteobactérias, Acidobactérias e Verrucomicrobias. Estes quatro filos são predominantes e onipresentes nos solos, indicando que ambas as bibliotecas metagenômicas são representativas de comunidades bacterianas contidas no solo.

Em 2007 e 2008 o grupo de Biotecnologia Molecular do Laboratório de Biologia Molecular da Universidade de Brasília realizou triagens funcionais por atividades amilolítica, celulolítica e xilanolítica das bibliotecas de grandes insertos (metagenoma da microbiota do solo do Cerrado clonado em fosmídios) e de pequenos insertos (metagenoma da microbiota do solo do Cerrado clonado em plasmídeos) construídas por De Castro (2008).

Levas de identificação foram realizadas e podem ser simplificadas nas seguintes etapas: a) Realização de diluições seriada das células da biblioteca para identificação de número de células viáveis; b) Realização de novas diluições de forma a se obter o número de colônias isoladas que se deseja triar em cada leva; c) Inoculação de células

das colônias isoladas em poços de placas de Elisa contendo meio e antibiótico apropriado; d) Utilização de carimbo metálico, específicos para placas de Elisa, para "carimbar" a microcultura das placas de Elisa em placas contendo meio sólido e os substratos das enzimas de interesse; e) Identificação de halos de hidrólise dos clones nos substratos específicos.

Triagens funcionais de 12672 clones da biblioteca de pequenos insertos da (BPI) identificaram cinco clones capazes de produzir halos de hidrólise em placas com amido. Clones celulolíticos e xalanolíticos inicialmente identificados não foram confirmados ao serem transferidos para novas placas com seus respectivos substratos.

Triagens funcionais de 3072 clones da Biblioteca de grande insertos da microbiota do Cerrado (BGI) identificaram 67 clones capazes de produzir halos de hidrolise em placas com amido. Novamente, clones celulolíticos e xilanolíticos inicialmente identificados não foram confirmados ao serem transferidos para novas placas com seus respectivos substratos.

Neste contexto, o presente trabalho se propôs a sequenciar, expressar e caracterizar novas amilases a partir de clones com perfis amilolíticos de bibliotecas metagenômicas da microbiota de solo do Cerrado.

2. OBJETIVOS

2.1. Objetivo geral

Identificar genes associados a atividades amilolíticas encontradas em bibliotecas metagenômicas da microbiota do solo do Cerrado.

2.2. Objetivos específicos

1) Realizar identificação confirmatória de clones amilolíticos.
2) Selecionar clones com atividade amilolítica promissoras.
3) Sequenciar diretamente fragmentos de interesse clonados em plasmídeos e subclonar fragmentos de interesse clonados em fosmídio para posterior sequenciamento.
4) Amplificar genes de interesse e clonar no vetor pBluescript II KS.

3. MATERIAL

3.1 Microrganismos utilizados

Escherichia coli linhagem DH5α - *End*A1, *rec*A1, *hsd*R17, *sup*E44, *gyr*A96, *thi*-1, *rel*A1, Δ*lac*U169 (ϕ 80*lac*ZΔM15) (Sambrook e Russel, 2001)

Escherichia coli linhagem EPI300 - F^- *mcr*A D(*mrr-hsd*RMS-*mcr*BC) f80d *lac*ZDM15 D*lac*X74 *rec*A1 *end*A1 *ara*D139 D(*ara*, *leu*)7697 *gal*U *gal*K l– *rps*L *nup*G *trf*A *ton*A *dhfr*. (EPICENTRE)

Escherichia coli linhagem XL10 Gold - Tet^r, Δ(*mcr*A)183, Δ(*mcr*CB-*hsd*SMR-*mrr*)173, *end*A1, *sup*E44, *thi*-1, *rec*A1, *gyr*A96, *rel*A1, *lac* Hte, [F', *proAB*, $lacI^q$ZΔM15 ,Tn10(tet^r) Amy Cam^r]^a (Invitrogen)

Estoques das linhagens utilizadas foram mantidos a -80ºC, em meio LB (ver seção meios de cultura) contendo 25% de glicerol.

3.2 Enzimas utilizadas

Enzimas de restrição

As enzimas *Xho*I, *Not*I, *Sac*I, *Kpn*I, *Pst*I, *Not*I, *Xba*I, *Sal*I, *Nde*I e *Hind*III foram fornecidas pela New England Biolabs, Promega, QBiogene e Fermentas Life Sciences e utilizadas como indicado pelo fabricante.

Enzima T4 DNA ligase (New England Biolabs) – utilizada na concentração de 40 U/µL nas reações de ligação.

RNAse A (Sigma-Aldrich) -Concentração de uso 10 mg/ml.

3.3 Marcadores de DNA utilizados (Anexo 1)

1 kb Ladder (Gibco BRL);

1 kb Plus DNA Ladder (O'GeneRuler)

3.4 *Kits* utilizados

PureYield^{TM} Plasmid Miniprep System **(Promega)** - utilizado para purificação de plasmídeos

QIAquick Gel Extraction Kit **(QIAGEN)** - utilizado para extração e purificação de DNA de gel de agarose.

3.5 Ferramentas de bioinformática

Busca de sequências: GenBank (NCBI) - (http://www.ncbi.nlm.nih.gov/)

Tradução, manipulação e identificação de janelas de leitura: "Molecular Toolkit" - (http://arbl.cvmbs.colostate.edu/molkit/index.html)

Análises de restrição: NEBcutter V2.0 - (http://tools.neb.com/NEBcutter2/index.php)

Alinhamento: ClustalW - (http://www.ebi.ac.uk/Tools/clustalw2/index.html)
Sequencher 4.10.1 - Gene Codes Corporation
MEGA 4 (Tamura *et al.*, 2007)
BioEdit Sequence Alignment Editor (Hall, 1999)

Identificação de proteínas preditas e *operons*: FGENESB (www.softberry.com)

Identificação de domínios conservados:
Pfam (http://pfam.sanger.ac.uk/)
InterProScan (http://www.ebi.ac.uk/Tools/InterProScan/)

3.6 Vetores

As representações esquemáticas dos mapas físicos dos vetores podem ser observadas nos Anexos 2-5.

pCF430 (Newman & Fuqua, 1999) - A biblioteca metagenômica de pequenos insertos utilizada neste trabalho foi resultante da clonagem de DNA isolado da microbiota do solo Cerrado no vetor pCF430. De Castro (2008) extraiu o DNA do solo utilizando o *PowerSoil DNA isolation kit* (MOBIO). O DNA foi digerido com a enzima *Pst*I e fragmentos com tamanho entre 2 a 8 kB foram clonados no vetor pCF430. Este vetor possui 10 kb e um sítio múltiplo de clonagem. O plasmídio pCF430 possui o promotor P_{BAD} dos genes catabólicos de L-arabinose de *E. coli*. Esse promotor já foi intensamente estudado e possibilita um controle fino da expressão gênica. A regulação é feita pela proteína AraC, que atua tanto como regulador positivo como negativo. Na ausência de L-arabinose, AraC se liga à montante de P_{BAD} reprimindo a expressão e na sua presença induz a expressão em um espectro de concentração de L-arabinose de até 1000 vezes.

pCC1FOS (EPICENTRE) - A biblioteca metagenômica de grandes insertos utilizada neste trabalho foi resultante da clonagem de DNA isolado da microbiota do solo Cerrado no vetor pCC1FOS. De Castro (2008) realizou a extração direta do solo segundo Williamson *et al*. (2005) com adaptações. Em seguida, com o auxilio do kit comercial *CopyControl™ Fosmid Library Production Kit*, fragmentou o DNA, tornou as extremidades das moléculas de DNA coesivas, purificou de gel de agarose o DNA de tamanho desejado (aproximadamente 40 kB) e clonou os fragmentos no fosmídio pCC1FOS™ linearizado em *Eco*72I.

pBluescript II KS + (Invitrogen) - A construção dos plasmídios contendo os fragmentos resultantes da digestão do plasmídio pB1 foram feitas utilizando o vetor comercial pBluescript II KS+ assim como a clonagem dos produtos de PCR. O vetor tem aproximadamente 3,0 kb e possui o gene de resistência a ampicilina, a origem de replicação do fago f1, parte do gene *lac*Z da enzima β-galactosidase e sítios múltiplos de clonagem

pABC1 (Souza, 1986) - Como controle positivo para os ensaios de atividade foi utilizado o plasmídio pABC1. O plasmídio possui o gene truncado da *α*-amilase de *Bacillus* sp. e é um plasmídio bifuncional, capaz de se replicar tanto em *E. coli* como em *B. subtilis*.

3.7 Meios de cultura

Os meios e soluções foram preparados com água destilada e esterilizados em autoclave por 20 min a 120 ºC. Soluções sensíveis ao calor foram esterilizadas por filtração usando-se membranas de 0,2 µm. Quando necessário foram adicionados arabinose e os devidos antibióticos nas seguintes concentrações finais: arabinose 200 µg/mL, ampicilina 100 µg/mL, tetraciclina 20 µg/mL., cloranfenicol 12 µg/mL., imediatamente antes da inoculação com os microrganismos.

Meio Luria-Bertani (LB)

Extrato de levedura	0,5% (p/v)
Peptona de caseína	1,0% (p/v)
NaCl	1,0% (p/v)

pH ajustado para 7,2 com NaOH.

Meio LB Ágar

Meio LB	
Ágar	1,5% (p/v)

Meio LB-Ágar-Amido:

Meio LB-Ágar	
Amido solúvel	0,5% (p/v)

Meio LB-Amido

Meio LB	
Amido solúvel	0,5% (p/v)

Meio SOB

Triptona	2,0% (p/v)
Extrato de levedura	0,5% (p/v)
NaCl	10 mM
KCl	2,5mM

Meio SOC

Meio SOB acrescido de:

$MgCl_2$	10 mM
$MgSO_4$	10 mM
Glicose	20 mM

Solução de Sais M9 5X

$Na_2HPO_4 \cdot 7H_2O$	6,4% (p/v)
KH_2PO_4	1,5% (p/v)
NaCl	0,25% (p/v)
NH_4Cl	0,5% (p/v)

Meio M9-Amido pH 7,0

Solução de sais M9 5X	200 ml
Amido	0,5% (p/v)
Ágar bacteriológico	1,5% p/v

H2O q.s.p. 1000ml

Solução de IPTG

Isopropil-β-D-tiogalactosídeo 1M

Diluído em água destilada.

Estocar a -20°C.

X-Gal 2%

X-Gal 20 mg/mL

Diluído em N, N dimetilformamida

Estocar a -20°C.

Ampicilina

Ampicilina 50mg/mL

Dissolver 0,5 g de ampicilina em 7 mL de H2O, (poderá ser necessária a adição de algumas gotas de NaOH, para auxiliar na diluição). Completar para 10 mL com H_2O destilada e homogeneizar, esterilizar por filtração em membrana *millipore* 0,22 µm. Estocar a -20°C.

Solução de Arabinose 25%

Dissolver 25 g de arabinose em 100 mL de água bidestilada e esterilizar por filtração em membrana 0,22 µm. Usar concentração final de 200µg/mL. Estocar a -20°C.

Solução de Tetraciclina 5 mg/mL

Dissolver 5 mg de tetraciclina em 1 mL de etanol. Concentração final utilizada 20 µg/mL. Estocar a -20°C.

Solução de cloranfenicol 20 mg/mL

Dissolver 20 mg de cloranfenicol em 1 mL de etanol. Solução de uso 20 mg/mL. Concentração final utilizada de 12 µg/mL. Estocar a -20°C.

3.8 Soluções utilizadas nas preparações caseiras de DNA plasmidial

Solução I (Tampão Tris – EDTA)

Tris-HCl pH 8,0	25 mM
EDTA	10 mM
Glicose	50 mM

Solução II (Solução de Lise alcalina)

NaOH	0,2 M
SDS	1 % (p/v)

Preparada no momento do uso.

Solução III (Solução de Neutralização)

Acetato de sódio	3 M
Ácido acético	2 M

pH ajustado para 5,0

Solução de RNAse

RNAse A	10 mg/mL
Acetato de sódio pH 4,8	50 mM

Solução incubada em água fervente, por 20 minutos.

Tampão TE

Tampão Tris-HCl (pH 8,0)	25mM
EDTA	10mM

Clorofil

Clorofórmio	24% (v/v)
Álcool Isoamílico	1% (v/v)

Acetato de Amônia

Acetato de Amônia	7,5M

3.9 Análise de DNA em gel de agarose

Tampão Tris-acetato EDTA (TAE)	10X
Tris-Acetato	40 mM
EDTA	1 mM

pH ajustado para 8,5

Brometo de Etídeo

EtBr	10mg/mL

Dissolvido em H_2O destilada.

Azul de Bromofenol

Azul de Bromofenol	2mg
Xileno Cianol	2mg
H_2O destilada q.s.p.	2mL

Conservar em geladeira.

Tampão de amostra não desnaturante 6X

Azul de Bromofenol	0,25% (p/v)
Xileno Cianol	0,25% (p/v)
Glicerol	30,00% (p/v)

3.10 Soluções de ensaio dextrinizante de amido

Tampão acetato de sódio 0,5 M, pH 5,5

Acetato de sódio	41 g
Água destilada (q.s.p.)	1 L

Corrigir pH para 5,5 com ácido acétigo glacial.

Ácido acético 1M

Ácido acético glacial (PA)	2,9 mL
Água destilada (q.s.p.)	50 mL

Reagente Fuwa

Iodo (I_2)	1% (p/v) diluído em etanol P.A.
Iodeto de potássio (KI)	10% (p/v)
Preparar solução de uso I_2/KI/H_2O	1:1:3

4. MÉTODOS

4.1 Digestão de DNA com enzimas de restrição

As digestões foram realizadas conforme instruções dos fabricantes das enzimas. O tempo de incubação, os tampões, as temperaturas de reação, as concentrações de DNA e enzima variavam de acordo com o material a ser digerido e a eficácia da enzima. Geralmente foram utilizadas 10 U de enzima para cada 5 μg de DNA, seguindo-se incubação a 37 ˚C por 2 h. Teve-se o cuidado de se observar se o volume da enzima não tinha ultrapassado 10% do volume final da reação, com a finalidade de evitar a ação inespecífica da enzima sobre o DNA.

4.2 Análise em gel de agarose (Sambrook *et al*., 2001)

A eletroforese em gel de agarose foi utilizada para análise dos fragmentos de DNA. A agarose foi preparada na concentração de 0,8 % a 1,0% (p/v) em tampão de corrida TAE 1X e foi adicionado brometo de etídeo (análogo de base que se intercala na molécula de DNA e emite fluorescência violeta quando excitado com luz ultravioleta) para concentração final de 0,5 μg/mL. O tampão de amostra foi diluído para a concentração final de 1X nas amostras de DNA a serem analisadas. As amostras e os marcadores moleculares foram aplicados no gel e submetidos à eletroforese com a corrente e voltagem adequadas (1-5 V/cm). Para visualização da amostra de DNA o gel foi exposto à luz ultravioleta

4.3 Purificação e eluição de fragmentos de DNA

Após digestão dos plasmídios com as enzimas de restrição apropriadas, os fragmentos de DNA contendo as seqüências de interesse foram aplicados em gel de agarose e submetidos à eletroforese. Os fragmentos de DNA de interesse foram cortados do gel e purificados utilizando-se o *kit QIAquick Gel Extraction*, conforme as especificações do fabricante.

4.4 Ligação de fragmentos de DNA (Vetor-Inserto)

Os sistemas de ligação foram feitos de modo que a razão molar entre vetor e inserto ficasse entre 1:3 e 1:10. A enzima T4 DNA Ligase foi utilizada com os tampões de reação fornecidos pelos fabricantes. Os sistemas foram incubados a 4°C por pelo menos 14 h antes de serem utilizados para transformação de bactérias.

4.5 Preparação de células bacterianas competentes para choque térmico

O preparo de células de *E. coli* competentes para choque térmico foi realizado como descrito por Cohen (1972) com algumas modificações. As células da linhagem de *E. coli* desejada foram crescidas em 5 mL de meio LB e incubadas a 37 ºC durante a noite sob agitação a 250 rpm. Em seguida 1 mL do pré-inóculo foi adicionado a 30 mL de meio LB e essa cultura incubada a 37ºC sob agitação (250 rpm) até atingir uma OD600 de 0,2 a 0,3. As células foram coletadas por centrifugação a 3.000 x *g* por 10 minutos a 4ºC e ressuspensas em 10 mL de $CaCl_2$ 100 mM estéril e gelado. Em seguida as células foram submetidas a uma nova centrifugação nas mesmas condições. Finalmente, as células foram aliquotadas em tubos Eppendorff (50 ou 100 µL/tubo) e estocadas a -80 ºC.

4.6 Transformação de *E. coli* por choque térmico

Uma alíquota de célula competente previamente preparada foi utilizada para cada sistema de ligação. As células foram retiradas do freezer a -80 ºC e deixadas no gelo até que descongelassem. Em seguida, foram adicionados 5 µL do sistema de ligação às células, e as mesmas foram novamente incubadas no gelo por 30 minutos. Após este período, as células foram submetidas a um choque térmico a 37 ºC, por 5 minutos ou 42°C por 90 segundos e, o sistema foi incubado no gelo por 2 minutos. Foram adicionados 900 µL de meio LB ao sistema que foi incubado a 37 ºC por 1 hora. O plaqueamento foi feito em duas placas contendo LB-ágar e o antibiótico adequado, em diferentes concentrações. As placas foram incubadas a 37 ºC durante a noite.

4.7 Preparação de DNA plasmidial em pequena escala (mini-prep) (adaptado de Sambrook *et al*., 2001)

Uma colônia de bactéria foi inoculada em 3 mL de meio LB contendo o antibiótico adequado seguindo-se cultivo a 37ºC por 16 horas sob agitação de 200 rpm. Estas células foram coletadas através de centrifugação a 12000 x *g* por 2 minutos a 4 ºC, descartando-se o sobrenadante. O sedimento foi ressuspenso em 200 µL de solução I, e em seguida adicionou-se 360 µL de solução II recém-preparada. A amostra foi homogeneizada, com cuidado, por inversão rápida do tubo. Após incubação por 5 minutos à temperatura ambiente, foram adicionados 300 µL de solução III gelada ao lisado, procedendo-se nova homogeneização por inversão do tubo, gentilmente. A amostra foi então incubada no gelo por 5 minutos e centrifugada a 12000 x *g* por 5 minutos a 4°C. O sobrenadante foi transferido para outro tubo contendo 750 µL de isopropanol. A solução foi homogeneizada por inversão e submetida a uma nova centrifugação a 12000 x *g* por 5 minutos. O sobrenadante foi descartado e o precipitado ressuspenso em 200 µL de solução TE. Foi adicionado 110 µL de acetato de amônio 7,5 M. Após forte agitação, o sistema foi centrifugado a 12000 x *g* por 10 minutos. O sobrenadante foi recuperado em um tubo limpo, onde foi adicionado 750 µL de etanol 100% gelado. A mistura foi centrifugada a 12000 x *g* por 5 minutos a temperatura ambiente. O precipitado foi lavado (sem ressuspender) com 1 mL de etanol 70 % gelado. Após secagem por exposição ao ar, o precipitado foi ressuspenso em 50 µL de TE com RNAse A (1 µL de RNAse 10 mg/mL para 100 µL de TE). O DNA plasmidial foi estocado a -20 °C até o uso.

4.8 Preparação de DNA plasmidial em média escala (midi-prep) (adaptado de Sambrook *et al*., 2001)

O procedimento foi o mesmo do descrito para a preparação de plasmídios em pequena escala. Apenas os volumes das soluções utilizadas foram ajustados ao volume de cultura de 50 mL. A análise e a estocagem do material, também seguiram os procedimentos da mini-preparação.

4.9 PCR

Para a amplificação de DNA por meio de PCR os sistemas foram montados com um volume final de 50µL, como descrito: cerca de 1 ng de DNA plasmidial como molde, 0,2 µM de cada um dos *primers*, 0,2 mM de uma solução de dNTPs; tampão de Taq DNA polimerase para uma concentração final de 1X, cloreto de magnésio (na concentração ótima para cada par de *primers*), 2 U de Taq DNA polimerase(Cenbiot) ou *Phusion High Fidellity Polymerase* (Finnzymes). Uma vez homogeneizada, a amplificação foi executada por 35 ciclos onde o DNA foi desnaturado a 94 °C por 1 min, anelado a temperatura ideal para cada par de *primers* por 1 min e elongado a 72 °C por 2 min. Após a amplificação, cerca de 1/5 do sistema de reação foi analisado por eletroforese em gel de agarose.

4.10 Identificação de atividades amilolítica em placas.

Placas com o meio LB-Ágar-Ámido ou meio M9-Ámido, contendo arabinose e o antibiótico adequado foram estricadas com os clones a serem analisados. As placas foram incubadas a 37°C por 14 horas e em seguida coradas com vapor de iodo. A exposição das placas ao vapor de iodo foi realizada pela sobreposição das placas contendo os clones sendo analisados (posicionada em cima) e placas de Petri contendo os grãos de iodo (posicionada embaixo). Os clones positivos para atividade de amilases foram detectados pela formação de um halo claro, no qual não fica retido o vapor de iodo, ao redor da colônia, indicando a degradação do substrato.

4.11 Ensaio em meio líquido para detecção de atividade dextrinizante de amido (Fuwa, 1954).

Tubos de rosca de 100 mL contendo 10 mL de meio LB-Ámido ou LB, com os antibióticos adequados e arabinose, foram inoculados com colônias isoladas de placas feitas a no máximo uma semana e incubados a 37°C sob agitação (250rpm). A cultura foi acompanhada por até 90 horas com alíquotas sendo retiradas em intervalos de aproximadamente 12 horas

Nos tubos contendo meio LB sem amido, alíquotas de 200 µL da cultura foram retiradas e centrifugada (8000 x *g* por 3 minutos) e 60 µL do sobrenadante foram

utilizados para a realização do ensaio. O ensaio foi realizado a 40°C por 30 µL de sobrenadante, 100 µL de amido 0,5% e 40 µL de tampão acetato 0,5M. Decorrido este tempo, a reação era parada com 200 µL de ácido acético 1M e adicionava-se 200 µL de reagente FUWA e 4,4 mL de água destilada. As amostras foram homogeneizadas e suas absorbâncias foram lida a 660 nm. Como branco de leitura do aparelho a 660 nm, uma reação onde o sobrenadante e o amido foram substituídos com água destilada foi utilizada. Como quantidade total de amido, foi utilizada reação onde o sobrenadante foi substituído por água.

Nos tubos contendo amido, alíquotas de 200 µL da cultura foram retiradas e centrifugadas (8000 *g* por 3 minutos) e utilizadas para dosagem de amido residual. A cada 100 µL de sobrenadante foi adicionado: 60 µL de água destilada, 40 µL de tampão acetato 0,5 M, 200 µL de ácido acético 1 M, 200 µL de reagente FUWA (KI : I : H_2O, 1:1:3) e 4,4 mL de água destilada. A solução foi homogeneizada e sua absorbância foi lida a 660 nm. Como branco de leitura do aparelho a 660 nm uma reação onde a amostra e o amido foram substituídos com água destilada foi utilizada. Os valores de absorbância foram convertidos em concentração de amido utilizando a curva padrão de amido construída com concentrações que variavam de 0 a 0,5 g de amido por 100 mL de meio LB (Fig. 8).

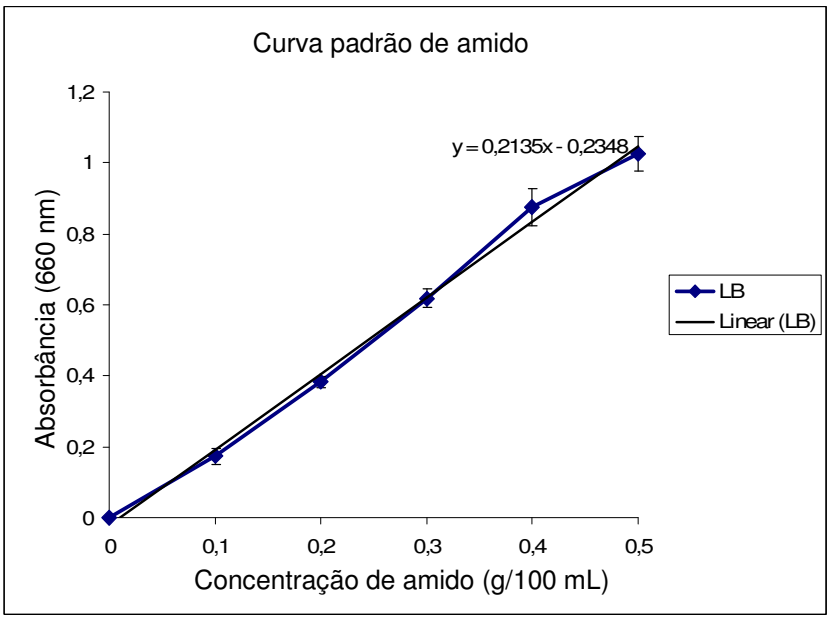


Figura 8. Curva padrão de amido.

4.12 Sequenciamento de DNA e análise *in silico*

As reações de sequenciamento realizadas no laboratório de Biologia Molecular da Universidade de Brasília foram realizadas com 200 a 300 ng de plasmídio quantificados por espectrofotometria utilizando o GeneQuant RNA/DNA *Calculator Pro* (Biochrom UK) ou pelo NanoDrop 2000 (Thermo Scientific) e analisados no sequenciador automático MegaBACE 1000(Amersham Biosciences), utilizando-se o *kit MegaBACE Dye Terminator*. A qualidade do sequenciamento foi analisada pelos programas PHRED (http://helix.biomol.unb.br/phph/) e a análise e montagem das sequências foi realizada com os programa BioEdit Sequence Alignment Editor e Sequencher 4.10.1.

Parte do sequenciamento foi realizado na Epoch Biolabs (EUA) e no Centro de Estudos do Genoma Humano (USP). Uma vez recebidas as sequências, as mesmas análises de qualidade e montagem foram realizadas.

ClustalW2 é um programa de alinhamento de múltiplas sequências. O programa calcula o melhor par para as sequências selecionadas e as alinha de forma que a identidade, similaridade e diferenças sejam visíveis.

As ferramentas InterProScan e Pfam foram utilizadas na identificação de domínios conservados nas sequências protéicas estudadas. InterProScan é uma ferramenta que combina diferentes métodos de reconhecimento de assinaturas protéicas nativas do banco de dados InterPro Member em uma única fonte.O banco de dados Pfam é uma ampla coleção de famílias proteicas, cada uma representada por alinhamento múltiplos de sequências e *Hidden Markov Models* (HMMs).

4.12.1 *Primer walking*

Esta técnica possibilitou o sequenciamento de fragmentos longos de DNA clonados em vetores cuja sequência seja conhecida e consiste no desenho de *primers* a partir da região 3' de uma sequência obtida anteriormente por *primers* que anelam no vetor. Novos *primers* são desenhados a cada vez que se avança na obtenção da seqüência até se montar um contínuo da sequência, também conhecido como *contig*. Os novos *primers* desenhados precisam estar a uma distância mínima da última base com qualidade da sequência de forma a possibilitar a montagem do *contig*. Dependendo do sequenciador utilizado, obtêm-se diferentes distâncias da região de anelamento dos

primers até as bases com qualidade identificadas. Para a montagem dos *contigs* somente bases de qualidade são utilizadas. Os nomes e sequências dos *primers* utilizados nos sequenciamentos realizados no laboratório de Biologia Molecular da UnB podem ser observados no Anexo 6.

4.12.2 Construção de cladogramas de proteínas ortólogas

Os cladogramas foram realizados com o programa MEGA 4.1 da seguinte forma: 20 sequências foram escolhidas para compor o alinhamento, sejam por grande similaridade atribuída pelo programa BLASTp ou para representarem grupos externos com conhecidas semelhança funcionais com as proteínas sendo estudadas. Uma vez alinhadas, regiões extremas em que não haviam bases representadas em alguma das sequências foram eliminadas para não introduzir erros nas análises. Os cladogramas foram construídos por *Neighbor Joining* com teste deduzido de filogenia do tipo *Bootstrap* com 1000 repetições.

5. RESULTADOS

5.1 Biblioteca de pequenos insertos

5.1.1 Confirmação de atividade amilolítica de clones da biblioteca de pequenos insertos

A triagem da biblioteca metagenômica de pequenos insertos (PI) da microbiota do solo do Cerrado realizada pelo grupo de Biotecnologia molecular da Universidade de Brasília, identificou cinco clones com halo de atividade em placas com meio LB-Ágar com amido 1%, coradas com vapor de iodo: clones A1, A2, A3, A4 e A5. Os plasmídeos dos clones amilolíticos foram extraídos por lise alcalina com o kit PureYield^{TM} Plasmid Miniprep System (Promega) e utilizados para transformação de *E. coli* linhagem EPI300. Os plasmídeos foram denominados, respectivamente, de pA1, pA2, pA3, pA3, pA4 e pA5 e os clones resultantes da transformação confirmatória de A1.1, A2.1, A3.1, A4.1 e A5.1.

O sistema de transformação foi plaqueado em meio LB-Ágar-Amido com tetraciclina e arabinose. Foi possível observar halos em todas as colônias dos transformantes que receberam o plasmídeo pA1 (dados não mostrados) e em nenhum outro transformante. Como controle positivo, *E. coli* linhagem EPI300 foi transformada com o plasmídeo pABC1 que possui fragmento clonado de *Bacillus* sp. com atividade amilolítica; e como controle negativo a mesma linhagem foi transformada com o vetor pCF430 intacto, denominada de 41.

Um clone de cada placa dos transformantes que receberam os plasmídeos pA1 a pA5 e os controles positivo e negativo foram transferidos para uma placa de Petri com dois compartimentos e as placas foram coradas com vapor de iodo (Fig. 9). Somente o halo de hidrólise do clone A1.1 e do controle positivo estão evidentes.

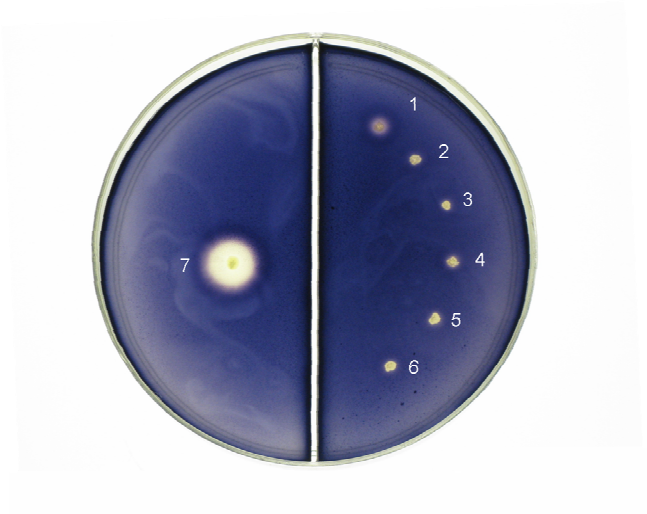


Figura 9. Confirmação de halo de hidrólise em amido de clones da biblioteca de pequenos insertos. Placa com amido 1% corada com vapor de iodo. 1 - A1.1; 2 - A2.1; 3 - A3.1; 4 - A4.1; 5 - A5.1; 6 - 41 (controle negativo); 7 - EPI300 transformada com plasmídeo pABC1 (controle positivo). Compartimento à esquerda: LB-Ágar-Amido com tetraciclina e arabinose. Compartimento à direita: LB-Ágar com tetraciclina.

5.1.2 Determinação de atividade dextrinizante de A1.1 em meio líquido

Com a finalidade de detectar a secreção de amilase foi realizado ensaio em meio líquido para detecção de atividade dextrinizante de amido do clone A1.1 e do controle negativo 41. Os clones foram cultivados em meio com amido 0,5% e o amido residual foi medido nas horas identificadas na Figura 10. No gráfico está demonstrada a média da concentração de amido residual presente no meio de cultura após 56 horas de ensaio em quatro experimentos. Os clones A1.1 e 41 foram crescidos em meio LB e alíquotas foram retiradas para realização do ensaio de FUWA em reações contendo 0,5% de amido. Não foi possível detectar atividade dextrinizante do sobrenadante dos cultivos sem amido mesmo após 56 horas de ensaio (dados não mostrados). Para verificar se a enzima estava no interior das células, as células foram sonicadas e foi realizado o ensaio de FUWA com o sobrenadante resultante da sonição com reações contendo 0,5% de amido. Não foi possível detectar atividade dextrinizante.

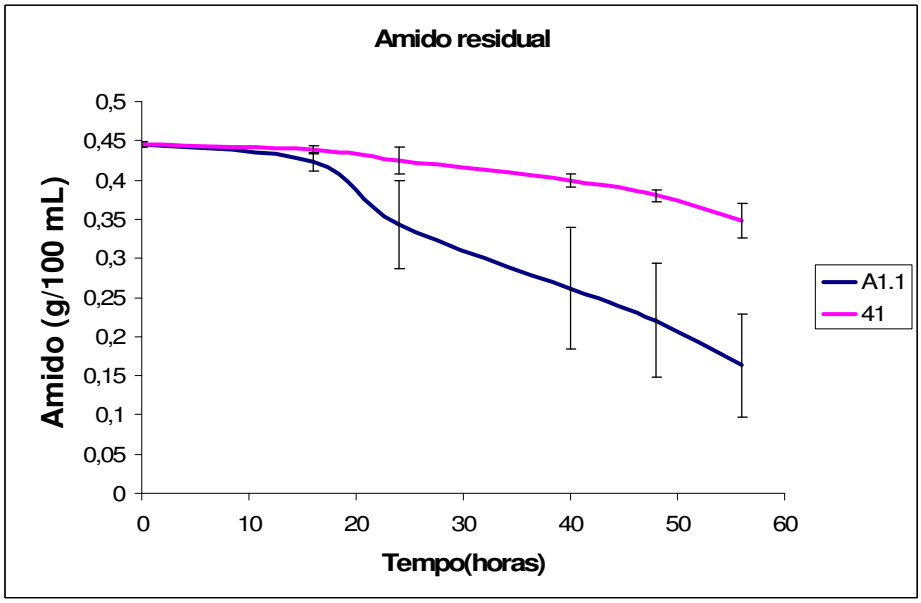


Figura 10. Representação da concentração de amido residual (g/100 mL) em solução no cultivo do clone A1.1 e do clone 41 (controle negativo) no decorrer do ensaio.

Um dos ensaios foi realizado por 90 horas e observou-se que o cultivo do clone A1.1 esgotou todo o amido do meio e o cultivo do clone 41 resultou na concentração de amido residual de 0,282 g/100mL.

5.1.3 Determinação da influência de arabinose na produção de halos de hidrólise em amido

Um clone de cada placa dos transformantes A1.1, A2.1, A3.1, A4.1, A5.1 e o controle negativo foram transferidos para uma placa de Petri com dois compartimentos na presença e ausência de arabinose e as placas foram coradas com vapor de iodo (Figura 11). Não foi possível distinguir aumento no diâmetro do halo de hidrólise do clone A1.1 na presença de arabinose.

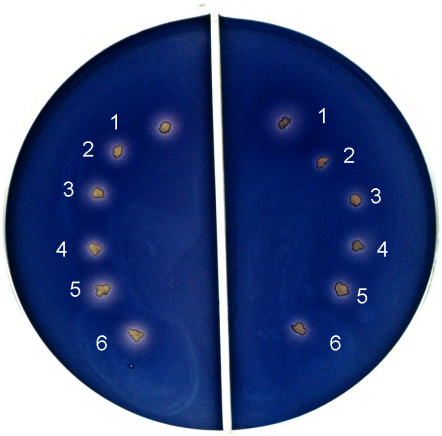


Figura 11. Verificação da influencia de arabinose na presença e diâmetro de halos de hidrólise de clones amilolíticos da biblioteca de pequenos insertos. Placa com amido 1% corada com vapor de iodo. 1 - A1.1; 2 - A2.1; 3 - A3.1; 4 - A4.1; 5 - A5.1; 6 - 41 (controle negativo); Compartimento à esquerda: LB-Ágar-Amido com tetraciclina. Compartimento à direita: LB-Ágar-Amido com tetraciclina e arabinose.

5.1.4 Identificação de perfil de restrição de pA1

Foi realizada análise de restrição de pA1 com as enzimas *Pst*I, *Hind*III, *Xba*I e *Kpn*I, enzimas de restrição com sítio único no vetor pCF430 (Figura 12).

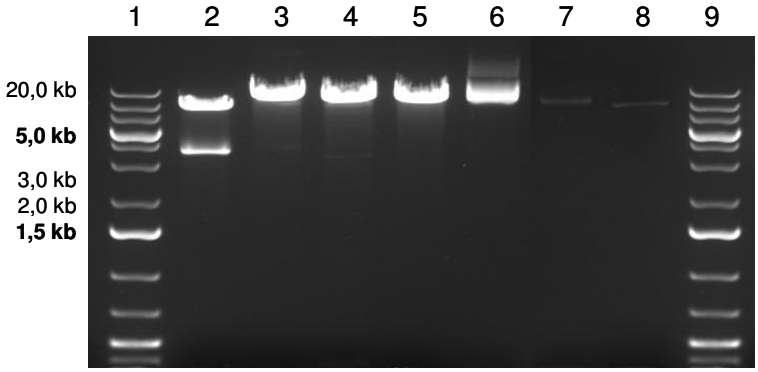


Figura 12. Visualização em gel de agarose 1% da digestão de pA1 com enzimas de sítio único no vetor pCF430. 1 - Marcador 1 kb Plus Ladder (O'GeneRuler); Os fragmentos de interesse do marcador estão discriminados à esquerda. 2 - pA1 digerido com *Pst*I. 3 - pA1 digerido com *Hind*III. 4 - pA1 digerido com *Xba*I. 5 - pA1 digerido com *Kpn*I. 6 - pA1 intacto.7 - Vetor pCF430 intacto. 8 - Vetor pCF430 digerido com *Pst*I; 9 - Marcador 1 kb Plus Ladder (O'GeneRuler).

A digestão de pCF430 com *Pst*I resultou em fragmento de aproximadamente 10 Kb. A digestão de pA1 com *Pst*I resultou em dois fragmentos: um com aproximadamente 10 kb e outro com aproximadamente 4 kb. A digestão de pA1 com *Hind*III, *Kpn*I e *Xba*I resultou em um fragmento de tamanho superior a 10 kb.

5.1.5 Sequenciamento de fragmento clonado em pA1

O fragmento clonado no plasmídeo pA1 foi sequenciado no laboratório de Biologia Molecular da Universidade de Brasília por *primer walking*. A montagem do *contig* foi realizada com auxílio do programa Sequencher 4.10.1 e pode ser observado na Figura 13. Foram identificadas 3625 bases nucleotídicas entre os sítios de *Pst*I, utilizados para a clonagem do fragmento. Um *primer* adicional foi sintetizado de forma a obtermos a sequência parcial do vetor de clonagem pCF430, que não possui sua sequência publicada.

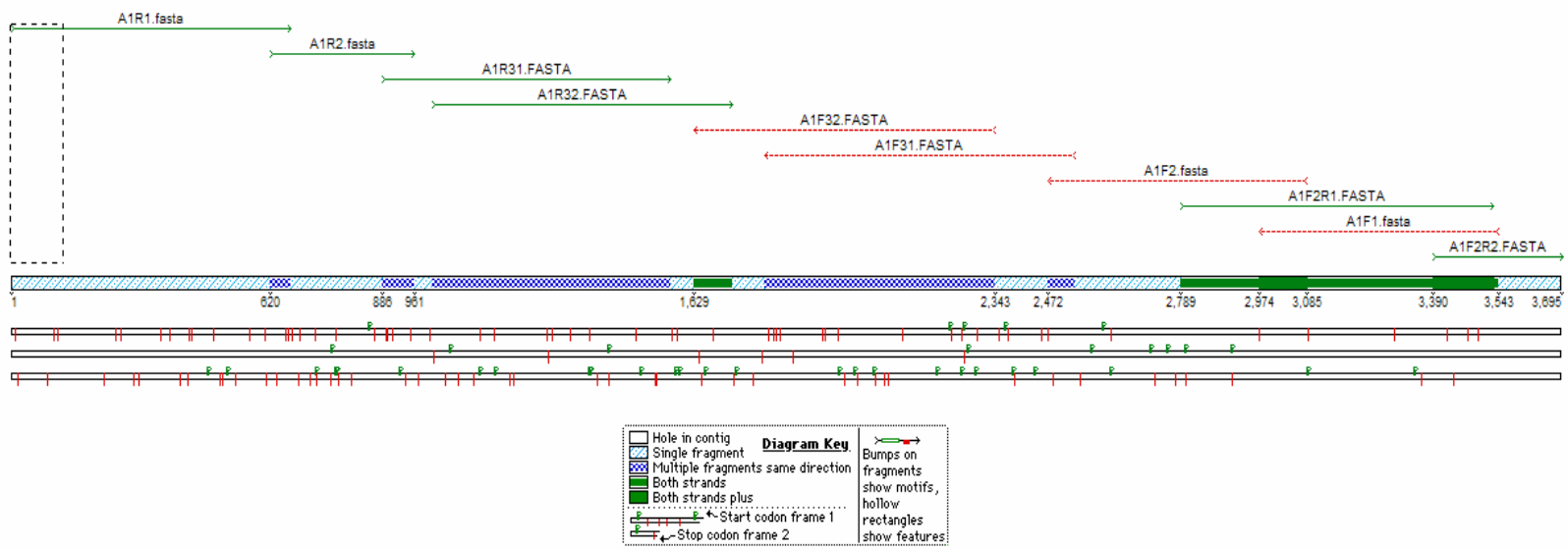


Figura 13. Visualização do *contig* de fragmento metagenômico de pA1. Sequências resultantes do sequenciamento foram alinhadas utilizando o programa Sequencher 4.10.1**.** Setas verdes representam sequências usadas diretamente para compor o *contig* e em setas vermelhas, sequências que foram invertidas para compor o *contig*.

O programa BLASTn (último acesso no dia 7/8/2010) não identificou sequências nucleotídicas com semelhança significativa à sequência do fragmento clonado em pA1 no banco de dados de coleção nucleotídica (nr/nt) do GenBank. O programa FGENESB: Bacterial Operon and Gene Prediction (www.softberry.com) foi utilizado para identificar proteínas preditas no fragmento metagenômico do plasmídeo pA1. Na Tabela 2 estão apresentadas as quatro proteínas preditas identificadas pelo programa FGENESB e a identificação das respectivas sequências proteicas contidas no banco de dados de sequências proteicas não redundantes (nr) do GenBank à qual foi atribuída o menor e-value pelo programa BLASTp (acessado no dia 7/8/2010) e domínios e famílias identificadas pelos programas BLAST e Pfam (último acesso no dia 7/8/2010).

Tabela 2. Anotação de sequências protéicas mais similares às proteínas preditas identificadas no plasmídeo pA1. Proteínas preditas pelo programa FGENESB da sequência de pA1 e identificação das sequências proteicas a que o programa BLASTp atribuiu menor E-value.

Proteínas preditas identificadas por FGENESB	Anotação	Domínio conservado putativo (GenBank)	Família Pfam	E-value / Coverage (GeneBank)
Proteína predita 1 - 281 aminoácidos	Região catalítica de alfa amilase [*Acidobacterium* sp. MP5ACTX8]	Superfamília Alfa-amilase	Alfa-amilase - Alfa amilase, domínio catalítico	2e-21 / 84%
Proteína predita 2 - 48 aminoácidos	-	-	-	-
Proteína predita 3 - 761 aminoácidos	Plug receptor TonB-dependente [*Acidobacterium* sp. MP5ACTX8]	Superfamília Canais_OM	Cna_B - proteína Cna domínio tipo B; DUF2846 - proteína de função desconhecida; e Plug - Domínio Plug receptor TonB-dependente	2e-164 / 97%
Proteína predita 4 - 41 aminoácidos	Proteína hipotética conservada [*E. coli* 83972]	-	-	1e-15 / 97%

A sequência nucleotídica e de aminoácidos da proteína predita 1 pode ser observada no Anexo 7. O programa InterProScan foi utilizado para realizar uma análise mais refinada de domínios na sequência predita. Todas as aplicações disponíveis no

programa foram utilizadas (último acesso no dia 08/07/2010) e o resultado pode ser observado na Figura 14.

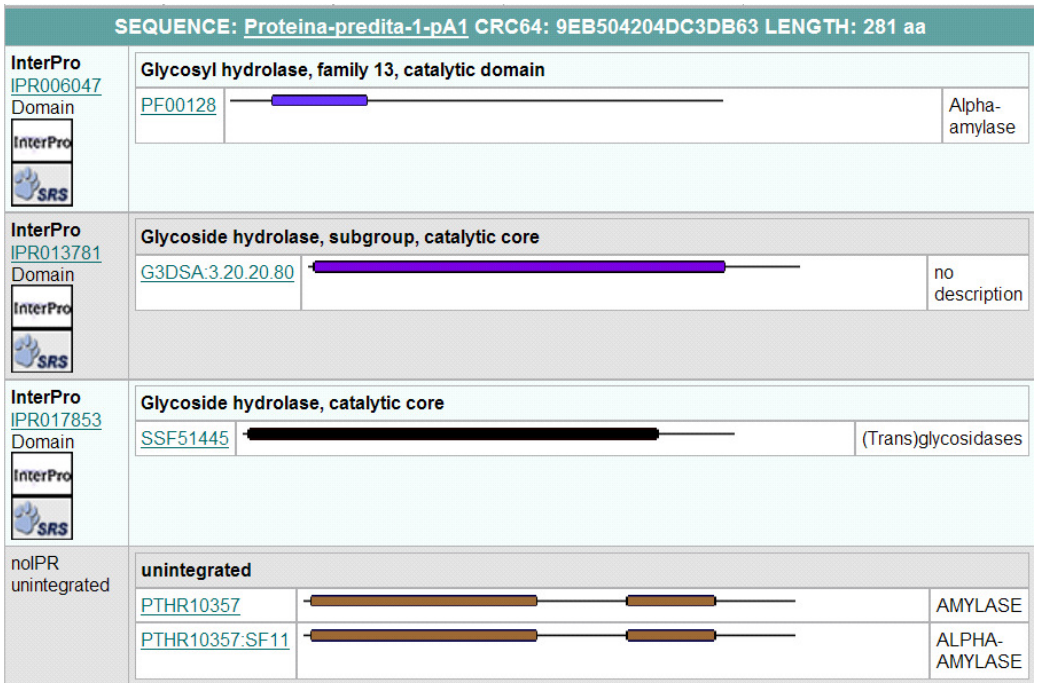


Figura 14. Domínios conservados identificados pelo programa InterProScan na sequência da proteína predita 1.

Foi realizado o alinhamento múltiplo da proteína predita 1 com proteínas com os menores e-values atribuídos pelo programa BLASTp de diferentes microrganismos e com a Taka-amilase de *Aspergillus oryzae* utilizando o programa ClustalW2 (Anexo 8). No alinhamento foi possível identificar que a proteína predita 1 conserva três das quatro assinaturas encontradas na superfamília das α-amilases destacadas na sequência da Taka Amilase de *Aspergillus orizae*, assim como possui os domínios secundários A, B e C de amilases. Foi também realizado o alinhamento local entre as sequências da proteína predita 1 com a sequência do domínio catalítico de amilase de *Acidobacterium* sp com o programa EMBOSS Pairwise Alignment (Anexo 9), o programa identificou 33,7% de identidade, 47,9% de similaridade e 20,3% de *gaps* entre as sequências.

As 15 sequências identificadas pelo programa BLASTp como similares à sequência de pA1 e cinco sequências de amilases de gêneros próximos foram utilizadas para construir um cladograma de proteínas ortólogas (Figura 15).

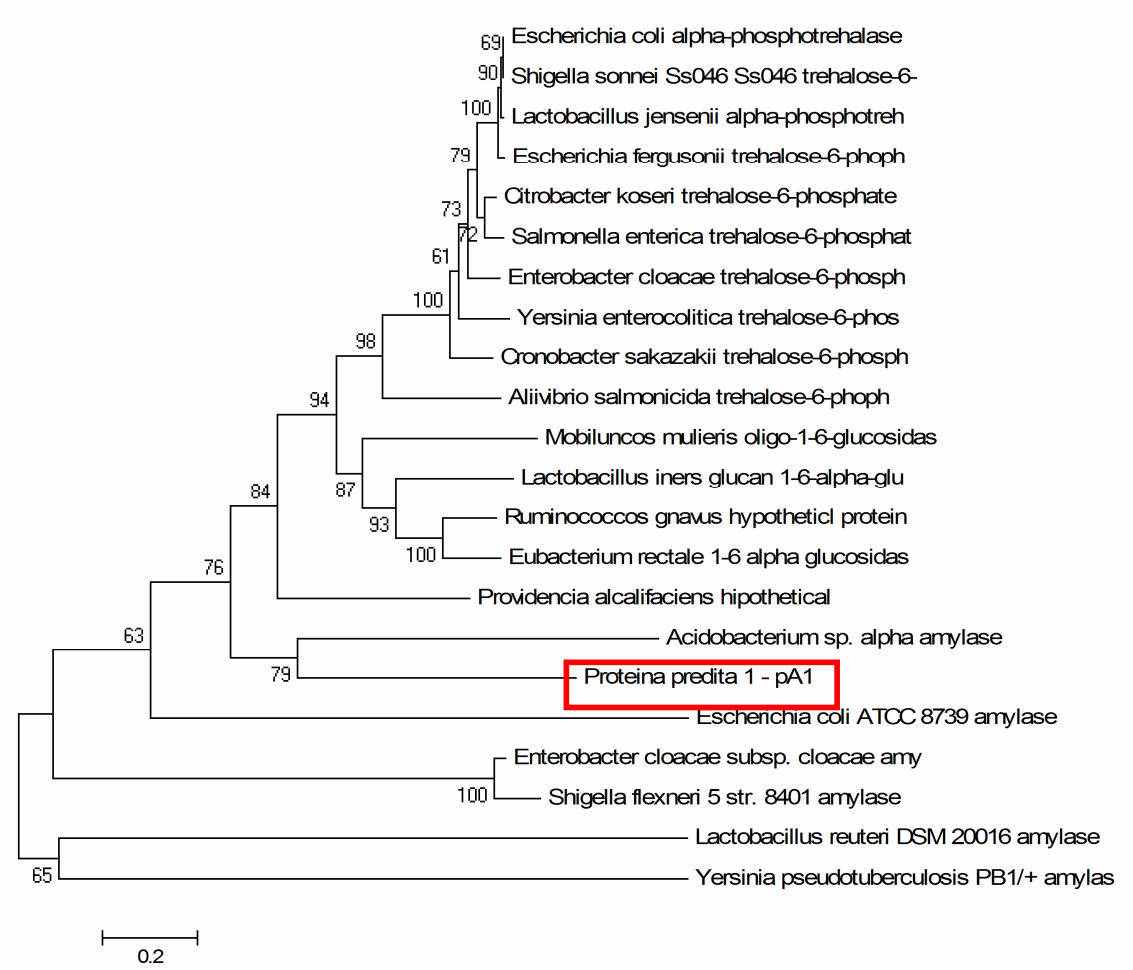


Figura 15. Cladograma da proteína predita 1 de pA1, sequências similares e amilases de microrganismos relacionados.

5.1.6 Amplificação e clonagem da proteína predita 1 em pET21a

Foram desenhados *primers* para amplificar a proteína predita 1 e adicionar sítios de restrição enzimática que permitissem a sua clonagem no vetor de clonagem pET21a. Um *primer forward* (A1CLOF) foi desenhado para adicionar o sítio de *Nde*I a jusante do códon de início e dois *primers reverse,* foram desenhados: o primero *primer reverse* (A1CLOR1) foi desenhado de forma a introduzir um códon de parada no final do *amplicon* contendo a sequência nucleotídica estritamente metagenômica de pA1 adicionando um sítio de *Sal*I após o códon de parada e o segundo *primer reverse* (A1CLOR2) resultou na amplificação de todo a sequência nucleotídica da proteína predita, contendo as bases referentes ao vetor de clonagem, seguidas de um sítios de *Sal*I. O fragmento contendo estritamente a sequência metagenômica (*amplicon* M) possui tamanho predito de 817 pb e o fragmento contendo a sequência nucleotídica

completa de proteína predita (*amplicon* MV) possui tamanho predito de 852 pb. A visualização em gel de agarose 1% demonstrou que o fragmento M possui ~0,7 kb e o fragmento MV ~0,8 kb (Figura 16).

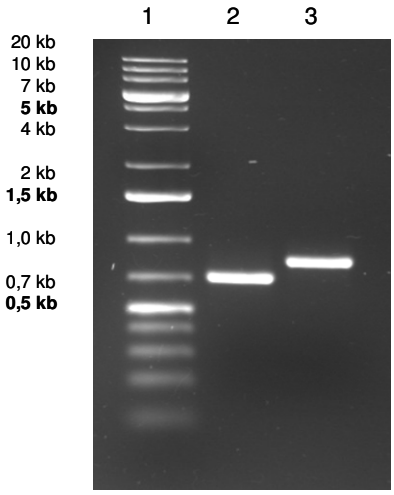


Figura 16. Visualização em gel de agarose 1% de PCR de pA1 com *primers* A1-CLOF/A1-CLOR1 e A1-CLOF/A1-CLOR2. 1 - Marcador O'Gene Ruler 1 kb Plus DNA Ladder. 2 - *Amplicon* de fragmento estritamente metagenômico (M). 3- *Amplicon* de fragmento metagenômico contendo bases do vetor (MV).

Os *amplicons* foram purificados em gel de agarose 1% e foram clonados no vetor pBlueScript II KS +, resultando nos plasmídeos pBLU-M e pBLU-MV. O sistema de ligação foi utilizado para transformar *E. coli* XL10 Gold. Na Figura 17 é possível observar as próximas etapas que serão realizadas com os plasmídeos pBLU-M e pBLU-MV.

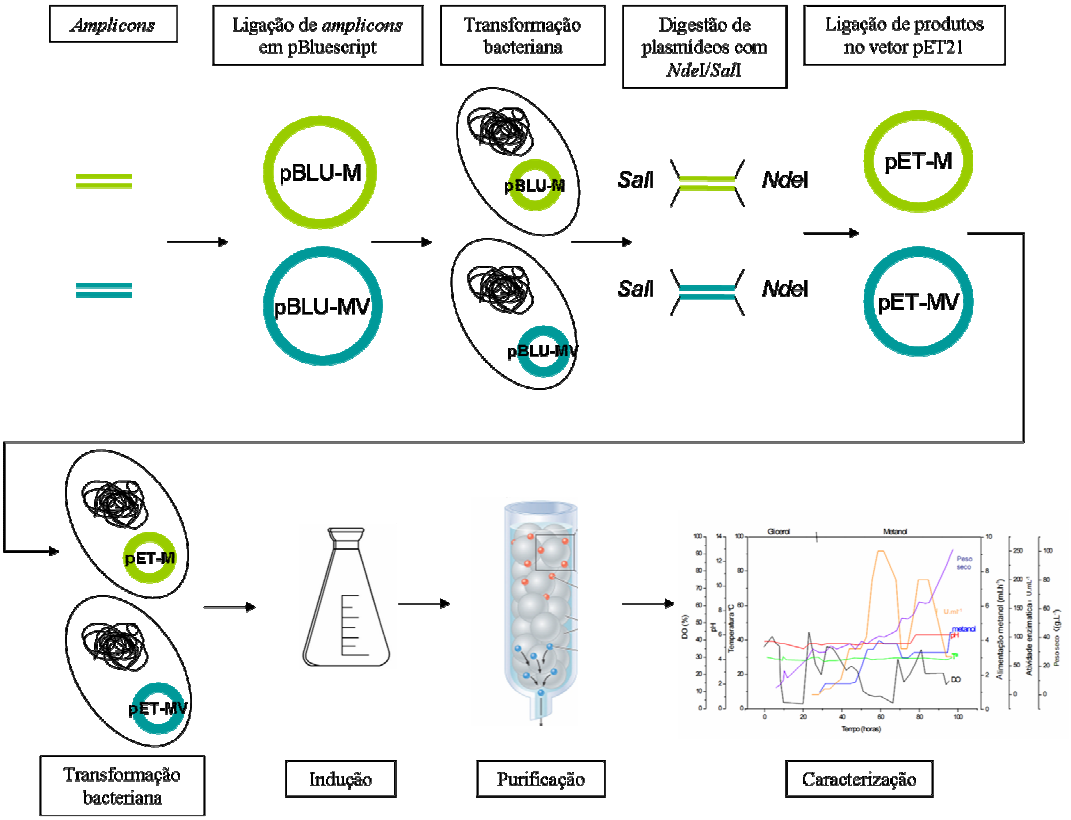


Figura 17. Ilustração das próximas etapas para clonagem dos *amplicons* resultantes da PCR, purificação e caracterização proteica.

5.2 Biblioteca de grandes insertos

5.2.1 Confirmação de atividade amilolítica de clones da biblioteca de grandes insertos

Na triagem da biblioteca metagenômica de grandes insertos (GI) da microbiota do solo do Cerrado, 67 clones apresentaram halo de atividade em placas com meio LB-Ágar com amido 1%, coradas com vapor de iodo. Dois clones foram escolhidos para transformação de EPI300 por apresentarem os maiores halos de hidrólise em placas: B1 e B2. Os plasmídeos destes dois clones (pB1 e pB2) foram extraídos por lise alcalina com o kit PureYield^{TM} Plasmid Miniprep System (Promega) e utilizados para transformação de *E. coli* linhagem EPI300, resultando nos clones B1.1 e B2.1. A transformação foi plaqueada em meio LB-Ágar-Amido com cloranfenicol e arabinose. Como controle positivo, *E. coli* linhagem EPI300 foi transformada com o plasmídeo pABC1 que possui fragmento clonado de *Bacillus* sp. com atividade amilolítica e como controle negativo a mesma linhagem foi transformada com o vetor pCC1FOS intacto, denominada de P1.

Todos os transformantes B1.1 e B2.1 apresentaram halo em placa com amido (dados não mostrados). Uma colônia de B1.1, B2.1 e o controle negativo foram transferidos para placa com amido e corados com vapor de iodo (Figura 18). Os clones transformados com pB1 e pB2 apresentam halo de hidrólise de amido, sendo o halo de B1.1 aparentemente maior que o halo de B2.1.

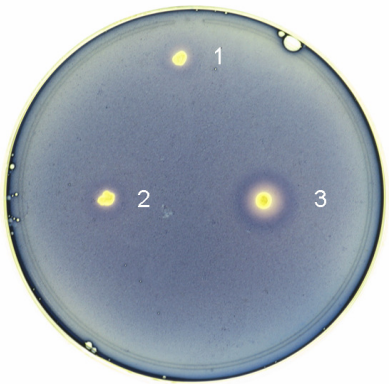


Figura 18. Confirmação de halo de hidrólise em amido de clones da biblioteca de grandes insertos. Placa com amido 1% corada com vapor de iodo. 1 - P1 (controle negativo); 2 - B1.1; 3 - B2.1.

5.2.2 Determinação de atividade dextrinizante de B1.1 em meio líquido

Foi realizado ensaio em meio líquido para detecção de atividade dextrinizante de amido com o clone B1.1 e com o clone P1. Na Figura 19 podemos observar a representação da concentração do amido residual presente no meio de cultura após 56 horas. Os clones B1.1 e P1 foram crescidos em meio LB e alíquotas foram retiradas para realização do ensaio de FUWA em reações contendo 0,5% de amido. Não foi possível detectar atividade dextrinizante do sobrenadante dos cultivos sem amido mesmo após 56 horas de ensaio (dados não mostrados). Para verificar se a enzima estava no interior das células, as células foram sonicadas e foi realizado o ensaio de FUWA com o sobrenadante resultante da sonição com reações contendo 0,5% de amido. Não foi possível detectar atividade dextrinizante.

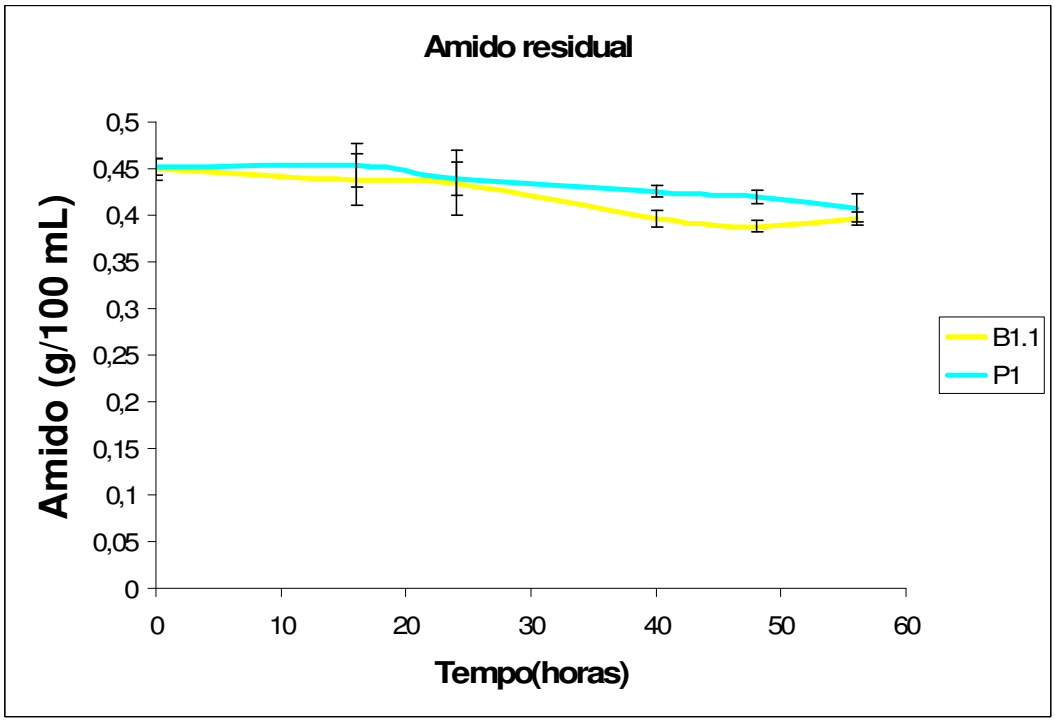


Figura 19. Representação da concentração de amido residual em solução no cultivo de B1.1 e P1 (controle negativo) no decorrer do ensaio.

Um dos ensaios foi realizado por 90 horas e observou-se que o cultivo do clone B1.1 resultou na concentração de amido residual de 0,283 e o cultivo do clone P1 resultou na concentração de amido residual de 0,269 g/100mL.

5.2.3 Determinação de influência de arabinose na produção de halos de hidrólise em amido

Um clone de B1.1 e de P1 foram repicados para uma placa de Petri com dois compartimentos na presença e ausência de arabinose e as placas foram coradas com vapor de iodo (Fig. 20). Houve aparente aumento no diâmetro do halo de hidrólise B1.1 na presença de arabinose.

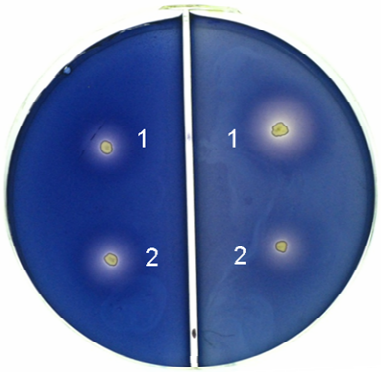


Figura 20 Verificação da influencia de arabinose na presença e diâmetro do halo de hidrólise de clone amilolítico da biblioteca de grandes insertos. Placa com amido 1% corada com vapor de iodo. 1 - B1.1; 2 - P1 (controle negativo). Compartimento à esquerda: LB-Àgar-Amido com cloranfenicol. Compartimento à direita: LB-Àgar-Amido com cloranfenicol e arabinose.

5.2.4 Identificação de perfil de restrição de pB1

Foi realizada análise de restrição de pB1 com as enzimas *Eco*RI, *Sca*I e *Hind*III enzimas de restrição com sítio único no vetor de clonagem PCC1FOS e com a enzima *Not*I (Figura 21) que possui dois sítios de clivagem no vetor, nas posição 2 e 643 no mapa do vetor mostrado no Anexo 3.

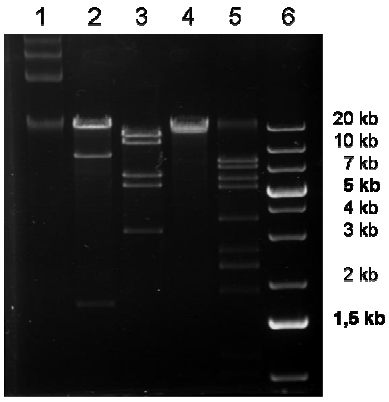


Figura 21. Visualização em gel de agarose 0,8% da digestão de pB1 com enzimas de restrição . 1 - pB1 intacto; 2 - pB1 digerido com *Hind*III; 3 - pB1 digerido com *Eco*RI; 4 - pB1 digerido com *Sca*I; 5 - pB1 digerido com *Not*I; 6 - Marcador 1 kb Plus Ladder (O'GeneRuler).

5.2.5 Subclonagem de pB1 em pBluescript II KS

Dez microgramas do plasmídeo pB1 foram digeridas com a enzima de restrição *Not*I e aplicadas em gel de agarose 0,8%. Os fragmentos de aproximadamente 9 kb, 7 kb, 6 kb e 3,7 kb foram purificados com o Illustra GFX PCR DNA and Gel Band Purification kit (GE Healthcare) e ligados a pBluescript II KS com a enzima T4 DNA ligase. O vetor foi previamente digerido com a enzima *Not*I e desfosforilado com a enzima SAP. A estratégia utilizada para obtenção dos subclones de pB1 está representada na Figura 22. Os clones resultantes da transformação de EPI300 com o sistema de ligação dos fragmentos subclonados não resultaram em clones que mantiveram o fenótipo (dados não mostrados).

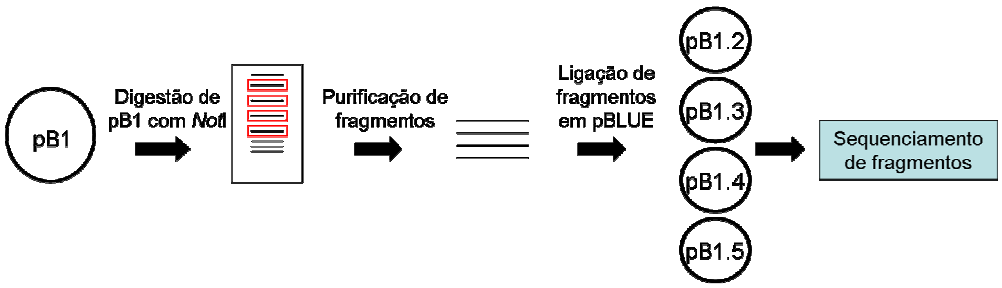


Figura 22. Estratégia utilizada para subclonar fragmentos resultantes da digestão de pB1 com *Not*I no vetor pBluescript II KS.

O sistema de ligação foi utilizado para transformação de *E. coli* DH5α. Os transformantes brancos tiveram seu plasmídeo extraído por mini preparação por lise alcalina caseira e digerido com a enzima *Not*I para confirmação de insertos. Os plasmídeos obtidos foram denominados pB1.2, pB1.3, pB1.4 e pB1.5 em ordem decrescente de tamanho e sua digestão com *Not*I pode ser observada na Figura 23.

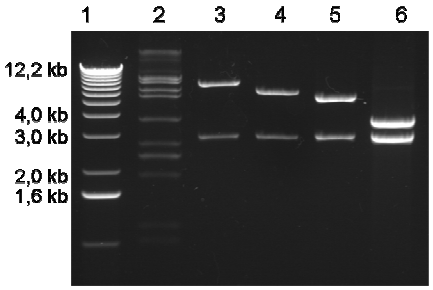


Figura 23. Visualização em gel de agarose 1% da digestão dos plasmídeos pB1.2, pB1.3, pB1.4 e pB1.5 com enzima de restrição *Not*I. 1 - Marcador 1 Kb DNA Lader (Gibco BRL); 2 - pB1 digerido com *Not*I; 3 - pB1.2 digerido com *Not*I; 4 - pB1.3 digerido com *Not*I; 5 - pB1.4 digerido com *Not*I; pB1.5 digerido com *Not*I.

Nas digestões dos plasmídeos pB1.2, pB1.3, pB1.4 e pB1.5 com *Not*I foi possível observar os fragmentos de tamanhos esperados e o fragmento de 3,0 kb referente ao vetor de clonagem pBluescript II KS.

5.2.6 Sequenciamento de subclones pB1.2, pB1.3, pB1.4 e pB1.5

Os transformantes contendo os plasmídeos pB1.2, pB1.3, pB1.4 e pB1.5 foram cultivados em meio LB para realização de Mini-prep com o Illustra plasmidPrep Mini Spin Kit (GE Healthcare). Os plasmídeos foram então enviados para Epoch Biolabs (EUA) para seqüenciamento por *Primer Walking*. Utilizou-se o programa Sequencher 4.10.1 para alinhar as sequências resultantes de cada leva de sequenciamento (Figuras 24, 25, 26 e 27). O número aproximado de bases que compõem cada fragmento clonado (retirando sequências referentes ao vetor) é de: 7,5kb, 6,2 kb, 5,4 kb e 3,3 kb.

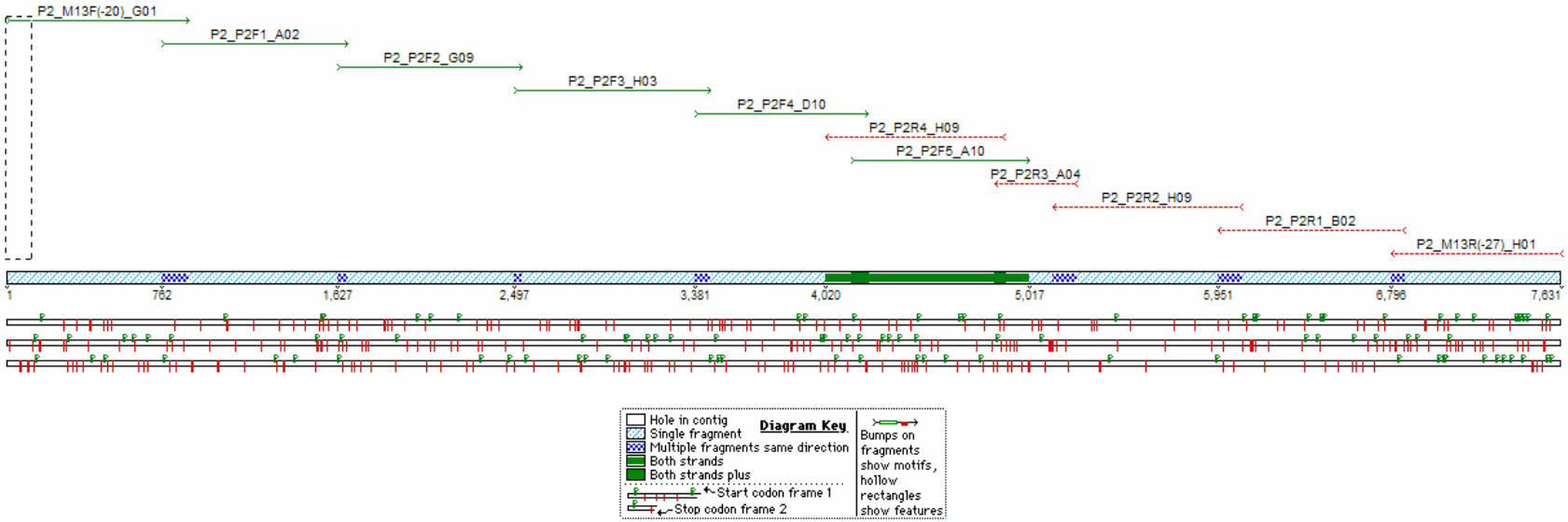


Figura 24. Visualização do *contig* de fragmento metagenômico de pB1.2. Sequências resultantes do sequenciamento foram alinhadas utilizando o programa Sequencher 4.10.1**.** Setas verdes representam sequências usadas diretamente para compor o *contig* e em setas vermelhas, sequências que foram invertidas para compor o *contig*.

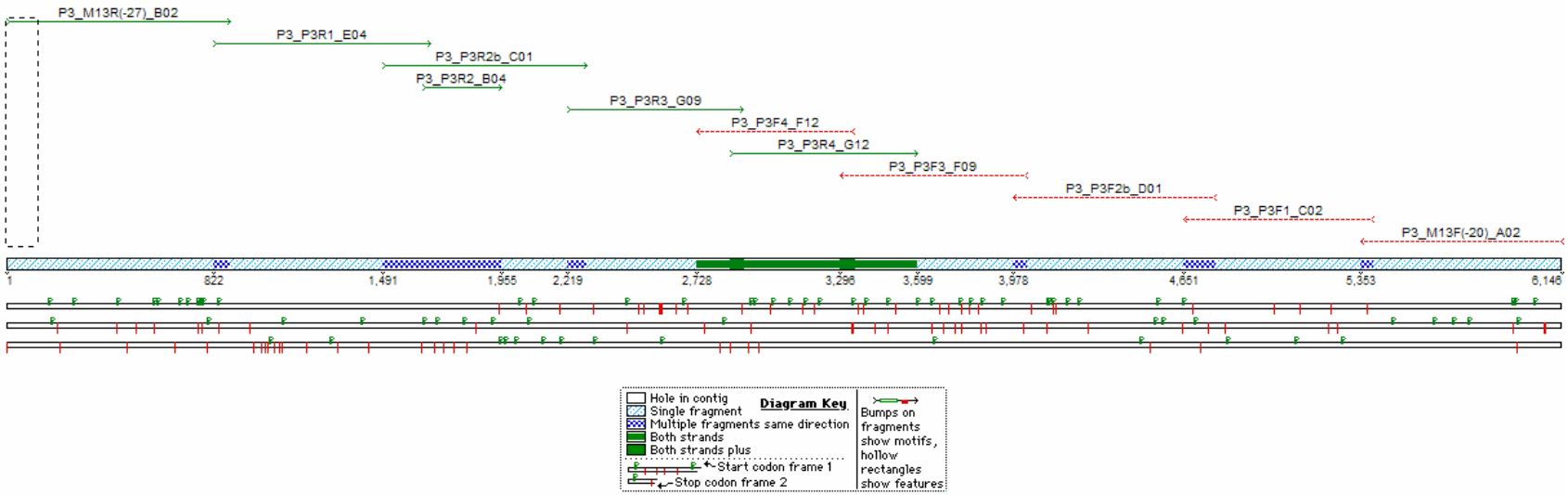


Figura 25. Visualização do *contig* de fragmento metagenômico de pB1.3. Sequências resultantes do sequenciamento foram alinhadas utilizando o programa Sequencher 4.10.1**.** Setas verdes representam sequências usadas diretamente para compor o *contig* e em setas vermelhas, sequências que foram invertidas para compor o *contig*.

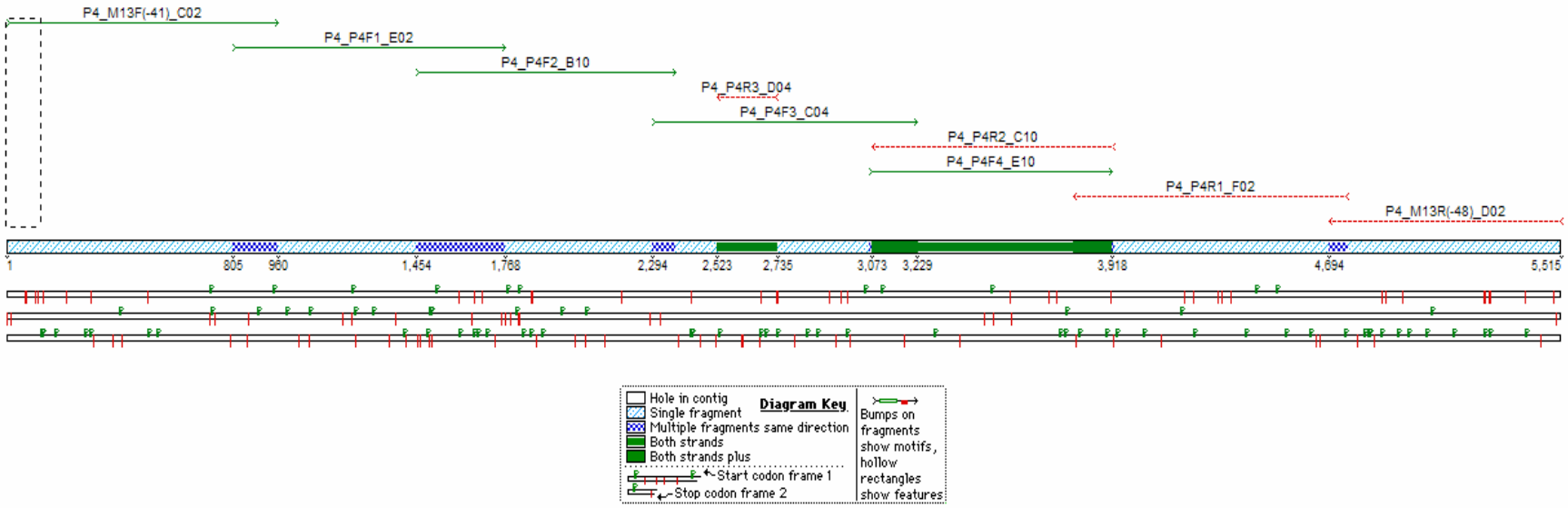


Figura 26. Visualização do *contig* de fragmento metagenômico de pB1.4. Sequências resultantes do sequenciamento foram alinhadas utilizando o programa Sequencher 4.10.1**.** Setas verdes representam sequências usadas diretamente para compor o *contig* e em setas vermelhas, sequências que foram invertidas para compor o *contig*.

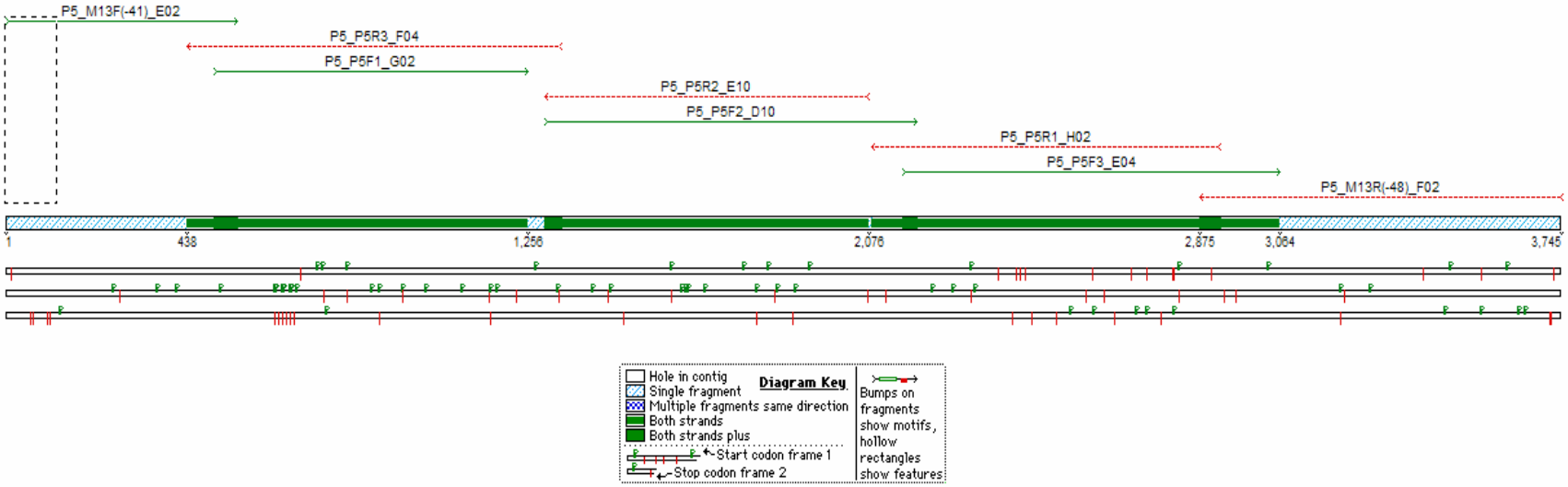


Figura 27. Visualização do *contig* de fragmento metagenômico de pB1.5. Sequências resultantes do sequenciamento foram alinhadas utilizando o programa Sequencher 4.10.1**.** Setas verdes representam sequências usadas diretamente para compor o *contig* e em setas vermelhas, sequências que foram invertidas para compor o *contig*.

O programa BLASTn (último acesso no dia 7/8/2010) foi utilizado para verificar semelhanças entre a sequência nucleotídica dos *contigs* e sequências depositadas no banco de dados de coleção nucleotídica (nr/nt) do GenBank. A sequência do fragmento clonado em pB1.2 foi identificada como sendo idêntica à sequência do vetor pCC1FOS depositada no banco de dados. O BLASTn não identificou sequências com semelhança significativa aos fragmentos clonados em pB1.3, pB1.4 e pB1.5. O programa FGENESB: Bacterial Operon and Gene Prediction (www.softberry.com) foi utilizado para identificar proteínas preditas nas sequências dos plasmídeos pB1.3, pB1.4 e pB1.5. As proteínas preditas foram utilizadas pelo programa BLASTp (acessado no dia 7/8/2010) para identificar sequências similares no banco de dados de sequências proteicas não redundantes (nr) do GenBank. Nas tabelas 3, 4 e 5 estão listadas as sequências protéicas identificadas pelo programa FGENESB, a identificação das respectivas sequências proteicas contidas no banco de dados GenBank às quais foi atribuído o menor e-value pelo programa BLASTp e domínios e famílias identificadas pelos programas BLAST e Pfam (ultimo acesso no dia 7/8/2010).

No plasmídeo pB1.4, aproximadamente dois terços do sua sequência em pares de bases são relacionadas ao metabolismo de ácidos graxos (proteínas preditas 5, 6 e 7). Foram identificadas como organizadas em um *operon* as proteínas preditas 5 e 6 e como uma unidade de transcrição separada a proteína predita 7 com o programa FGENESB (dados não mostrados). As proteínas preditas 5 e 6, por serem semelhantes a proteínas com a mesma anotação, foram alinhadas localmente com o programa EMBOSS Pairwise Alignment. Foram identificados 21.3% de identidade, 32,3% de similaridade e 29,5% de *gaps.*

Tabela 3. Anotação de sequências protéicas mais similares às proteínas preditas identificadas no plasmídeo pB1.3. Sequências proteicas a que o programa BLASTp atribuiu menor E-value de alinhamento com as proteínas preditas pelo programa FGENESB do plasmídeo pB1.3.

Proteínas preditas identificadas por FGENESB	Anotação	Domínio conservado putativo (GenBank)	Família Pfam	E-value / Coverage (GeneBank)
Proteína predita 1 - 288 aminoácidos	Produto protéico sem nome [*Tetraodon nigroviridis*]	-	-	3.0 / 30%
Proteína predita 2 - 294 aminoácidos	Proteína inter membrana transportadora de cadeia ramificada de aminoácido ABC [*Burkholderia* sp. CCGE1002]	Superfamília tipo TM_PBP1 cadeia ramificada de AA	BPD_trans 2 - Sistema de transporte de aminoácidos de cadeia ramificada	1e-104 / 99%
Proteína predita 3 - 319 aminoácidos	Translocador inter membrana [*Polaromonas naphthalenivorans* CJ2]	Superfamília tipo TM_PBP1 cadeia ramificada de AA	BPD_trans 2 - Sistema de transporte de aminoácidos de cadeia ramificada	1e-115 / 93%
Proteína predita 4 - 503 aminoácidos	Malonato decarboxilase, subunidade alfa [*Burkholderia* sp. CCGE1003]	Superfamília CoA_Trans	-	0.0 / 98%
Proteína predita 5 - 55 aminoácidos	-	-	TraI_2 - helicase putativa	-
Proteína predita 6 - 98 aminoácidos	Proteína predita [*Postia placenta* Mad-698-R]	-	-	6.3 / 47%
Proteína predita 7 - 215 aminoácidos	Regulador transcricional Dois components da família LuxR [*Anaeromyxobacter sp*. Fw109-5]	Superfamília REC	Response_reg - Dominio receptor de regulador	4e-27 / 53%
Proteína predita 8 - 203 aminoácidos	Histidina quinase híbrida sensora de GAF [*Acidovorax avenae subsp. avenae* ATCC 19860]	Superfamília GAF	HisKA_3 - Histidina quinase	2e-19 / 75%

Tabela 4. Anotação de sequências protéicas mais similares às proteínas preditas identificadas no plasmídeo pB1.4. Sequências protéicas a que o programa BLASTp atribuiu menor E-value de alinhamento com as proteínas preditas pelo programa FGENESB do plasmídeo pB1.4.

Proteínas preditas identificadas por FGENESB	Anotação	Domínio conservado putativo (GenBank)	Família Pfam	E-value / Coverage (GeneBank)
Proteína predita 1 - 157 aminoácidos	Proteína hipotética [*Curvibacter putative* symbiont of *Hydra magnipapillata*]	Superfamília arginase	Arginase - família arginase	2e-65 / 98%
Proteína predita 2 - 360 aminoácidos	ornithina ciclodeaminase [*Polaromonas* sp. JS666]	Superfamília NADB_Rossman	OCD_Mu_crystal 1 - Ornitina ciclodeaminase/ Família mucristalin	3e-135 / 95%
Proteína predita 3 - 67 aminoácidos	RNA-dependente RNA polimerase proteína de fusão P1-P2 [Rose spring dwarf-associated vírus]	-	-	1.7 / 95%
Proteína predita 4 - 140 aminoácidos	Regulador transcricional putative, AsnC/Lrp family [*Cupriavidus taiwanensis*]	HTH DNA-dominio de ligação de GntR	HTH_11 - domínio HTH e AsnC_trans_reg - família AsnC	2e-47 / 99%
Proteína predita 5 - 395 aminoácidos	Proteína contendo domínio de acil-CoA desidrogenase [*Sphingomonas wittichii* RW1]	Superfamília ACAD	Acyl-CoA_dh_N, M e 1 - Acyl-CoA desidrogenase, domínios N, mediano e C-terminal	8e-153 / 99%
Proteína predita 6 - 340 aminoácidos	Proteína acil-CoA desidrogenase [*Pseudoalteromonas haloplanktis* TAC125]	Superfamília ACAD	Acyl-CoA_dh_N e 1 - Acyl-CoA desidrogenase, domínios N, mediano e C-terminal	9e-74 / 95%
Proteína predita 7 - 188 aminoácidos	acetil-CoA acetiltransferase ou tiolase [*Pseudoalteromonas haloplanktis* TAC125]	Superfamília de enzimas condensadoras	Thiolase_N e C - Thiolase, domínios N e C terminal	6e-79 / 99%

Tabela 5. Anotação de sequências protéicas mais similares às proteínas preditas identificadas no plasmídeo pB1.5. Sequências protéicas a que o programa BLASTp atribuiu menor E-value de alinhamento com as proteínas preditas pelo programa FGENESB do plasmídeo pB1.5.

Proteínas preditas identificadas por FGENESB	Anotação	Domínio conservado putativo (GenBank)	Família Pfam	E-value / Coverage (GeneBank)
Proteína predita 1 - 40 aminoácidos	-	-	-	-
Proteína predita 2 - 125 aminoácidos	Proteína hipotetica BpseD_23784 [*Burkholderia pseudomallei* DM98]	Superfamília tipo NTF2	-	2e-30 / 97%
Proteína predita 3 - 97 aminoácidos	Proteína Y osmoticamente induzivel [*Escherichia albertii* TW07627]	Superfamília BON	BON - domínio putativo de ligação a fosfolipídio	7e-06 / 80%
Proteína predita 4 - 546 aminoácidos	Pirrolo quinolina-quinona [*Burkholderia* sp. H160]	Multi-domínio	PQQ - PQQ repetição de enzima	5e-162 / 91%
Proteína predita 5 - 197 aminoácidos	Proteína hipotética CC_2037 [*Caulobacter crescentus* CB15]	Superfamília Polyketide_cyc 2	Polyketide_ cyc - ciclase policetídica / Dehidrase e transporte lipídico	2e-19 / 73%

5.2.7 Cladograma de proteínas ortólogas

Foram escolhidas três proteínas preditas de cada um dos plasmídeos decorrentes da subclonagem de pB1 para realização da construção de cladograma: proteína predita 4 de pB1.3 (anotação: malonato decarboxilase, subunidade alfa), proteína predita 5 de pB1.4 (anotação: proteína contendo domínio de acil-CoA desidrogenase) e proteína predita 4 de pB1.5 (anotação: pirrolo quinolina-quinona). A escolha foi decorrente da proteína predita com menor E-value das encontradas nos respectivos plasmídeos (Figuras 28, 29 e 30). As sequências mais similares às proteínas preditas identificadas pelo programa BLASTp foram utilizadas para compor a filogenia.

Figura 28. Cladograma da proteína predita 4 de pB1.3 e sequências similares.

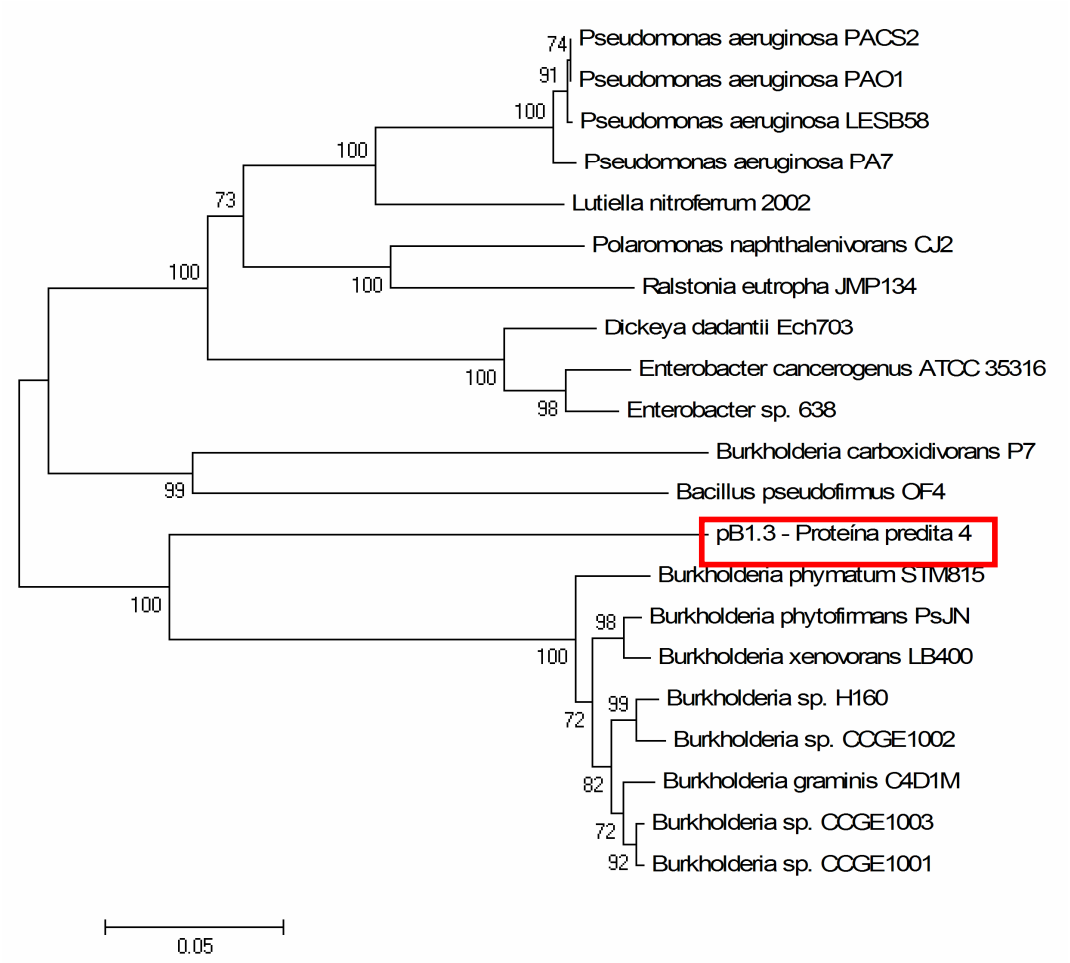


Figura 29. Cladograma da proteína predita 5 de pB1.4 e sequências similares.

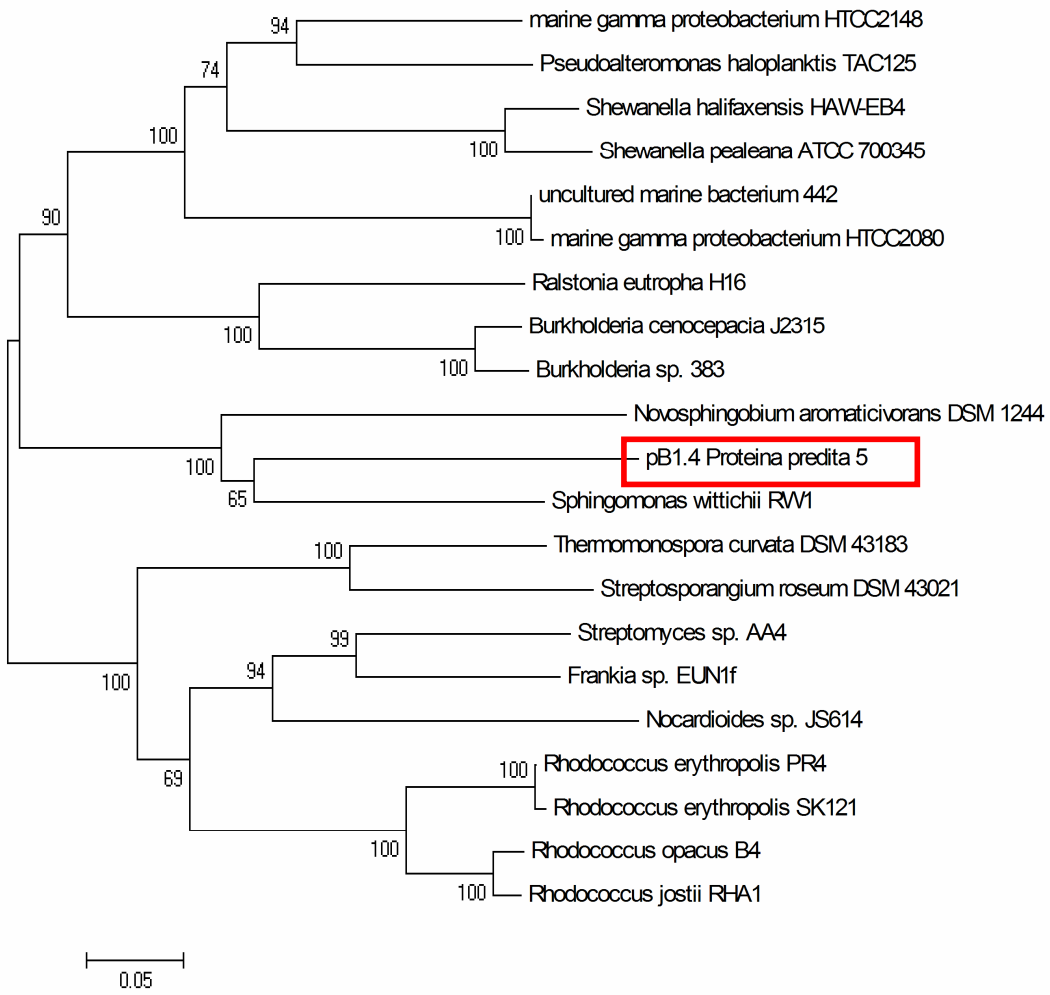
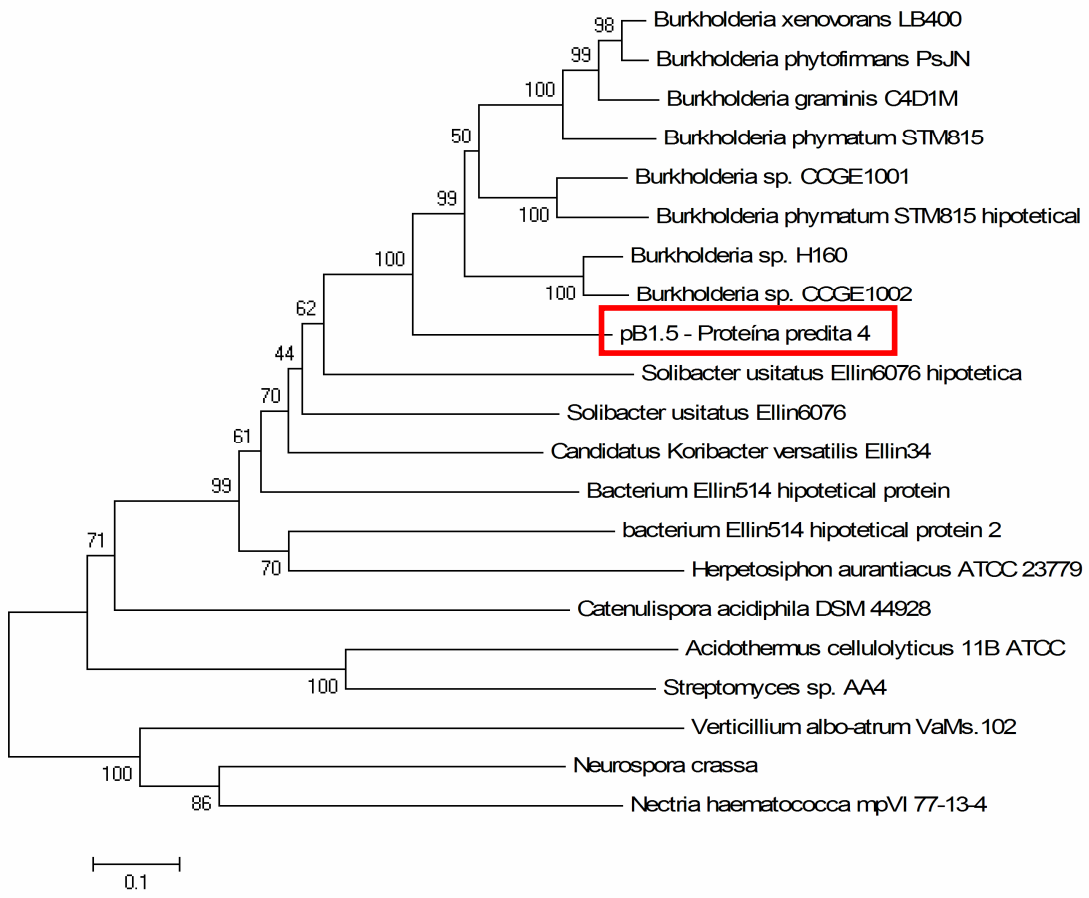


Figura 30. Cladograma da proteína predita 4 de pB1.5 e sequências similares.



6. DISCUSSÃO

A diversidade dos microrganismos do solo ainda incultiváveis e o potencial que representam para a descoberta de novas enzimas têm gerado um interesse cada vez maior neste ecossistema. A utilização das ferramentas metagenômicas tem possibilitado um conhecimento crescente destes microrganismos pela construção de bibliotecas de DNA extraído de amostras ambientais, dispensando a necessidade do cultivo. Devido ao interesse de aplicações biotecnológicas de novas enzimas hidrolíticas, bibliotecas metagenômicas da microbiota do solo do Cerrado foram construídas e triadas pelas atividades amilolítica, celulolítica e xilanolítica. Neste trabalho clones amilolíticos estáveis foram escolhidos para caracterização.

Dentre os clones amilolíticos identificados nas duas bibliotecas analisadas, dois foram escolhidos para caracterização: A1 e B1. A retransformação da mesma linhagem com plasmídeo extraído de clone positivo identificado é uma etapa importante para confirmar o fenótipo dos clones previamente identificados. A ausência de halos de atividade foi identificada nos três grupos enzimáticos identificados em 2007 (clones amilolíticos, xilanolíticos e celulolíticos). A razão que leva determinado clone a deixar de apresentar halo de atividade após a retransformação não é plenamente conhecida. Possivelmente, alguma modificação não-intencional nas condições de ensaio não permitiu a reprodutibilidade do fenótipo, apesar de pequenas modificações nas condições terem sido realizadas (dados não mostrados).

No caso de halos de clones amilolíticos, podemos relacioná-lo à possível presença de um fragmento metagenômico que codifique para uma lactamase. Métodos iodométricos são utilizados como forma de identificar microrganismos com resistência a antibióticos como penicilina e cefalosporina (Sawai *et al.*, 1978). O produto da atividade de β-lactamase (por exemplo, o ácido penicilóico no caso da ação da β-lactamase sobre a penicilina) é estequiometricamente oxidado por uma solução de iodo, resultando na descoloração do complexo azul amido-iodo (Novick, 1962). Desta forma, a descoloração observada pode ter sido decorrente da interação de algum composto contendo um anel β-lactâmico com uma β-lactamase sendo expressa pelo clone. Se a retranformação for plaqueada em meio onde não haja substrato para a β-lactamase, possivelmente o mesmo halo não seja observado.

Vizando atribuir o halo observado especificamente à ação hidrolítoca sobre o amido, foram realizadas tentativas de crescer os clones em meio mínimo M9-Amido-Ágar, entretanto não houve crescimento nas placas (dados não mostrados).

Em ensaio em meio líquido foi possível confirmar a atividade somente do clone A1.1 (Figura 10). Enquanto o controle positivo (linhagem EPI300 transformada com o plasmídeo pABC1) contendo a α-amilase de *B. subtillis*, esgotou o amido do meio em menos de 16 horas (dados não mostrados), ensaios com o clone A1.1 demonstraram que este clone leva 56 horas para degradar aproximadamente 63% do amido em solução e ~90 horas para esgotar o amido do meio. Entretanto, poucas comparações podem ser feitas entre estas atividades devido às diferenças nos vetores utilizados. A construção de um *amplicon* da α-amilase de *B. subtillis* contendo sítios para sua clonagem no vetor pCF430 não foram construídos em tempo hábil.

A atividade identificada, apesar de aparentemente baixa, era esperada e possui vários fatores que a justifiquem, dentre eles: a) a limitação da utilização da maquinaria celular da hospedeira na expressão heteróloga; b) a utilização do plasmídeo pCF430 para a clonagem dos fragmentos metagenômicos da biblioteca de PI; c) as condições ótimas para esta enzima ainda não serem conhecidas; d) a proteína sendo expressa está truncada; e e) α-amilases já foram descritas como tóxicas para determinados microrganismos.

a) Foi possível confirmar, tanto pela sequência de pA1, como pela realização de ensaio de atividade dextrinizante do clone A1.1, cultivado na ausência de amido, que o gene não está sob o controle do promotor P_{BAD} e sim sobre o controle de um promotor nativo. Uma sequência predita de promotor não foi identificável na sequência de pA1, mas, aparentemente, este é induzido na presença de amido, justificando a não detecção da atividade amilolítica na reação enzimática com o sobrenadante do cultivo de A1.1 que não possuía amido. A segunda confirmação da atuação de um promotor nativo é a distância de 772 bases do códon de iniciação da proteína predita 1 até o início do promotor P_{BAD} e a sua presença na fita inversa complementar à sequência do promotor.

Desta forma, a maquinaria celular de *E. coli* teve que reconhecer um promotor nativo heterólogo e ser capaz de transformar o produto da tradução em uma proteína funcional, etapa identificada na literatura como mais impactante na detecção da atividade desejada (Uchiyama & Miyazaki, 2009).

b) O plasmídeo pCF430 utilizado para a clonagem dos fragmentos metagenômicos da biblioteca de pequenos insertos realiza um controle fino da expressão de genes sobre o controle do promotor P_{BAD}. Possui ainda em sua sequência uma origem de replicação vegetativa (oriV) típica de plasmídeos conjugativos, sendo responsável pela manutenção deste plasmídeo em fita dupla, além de torná-lo um plasmídeo de amplo espectro de hospedeiros. Entretanto, esta origem confere tipicamente, um baixo número de cópias do plasmídeo por célula. No caso do fosmídio pCC1FOS utilizado na biblioteca de grandes insertos, há duas origens de replicação, uma de replicação do fator F e uma origem de replicação vegetativa dependente da proteína TrfA. TrfA tem sua expressão induzida na presença da proteína AraC produzida pela célula EPI300 na presença de arabinose. Nesse caso a presença de arabinose leva a um aumento do número de cópias do gene de interesse na célula independente da presença de um protoror P_{BAD} em fase com o gene.

c) As condições ótimas da enzima clonada em pA1 não foram determinadas nos ensaios em meio líquido devido à baixa atividade identificada e ao condicionamento da cultura do clone A1.1 ser realizada em amido para que ocorra a produção enzimática. Desta forma, as condições em que foram realizados os ensaios, pH inicial de 7,2 e temperatura de 37°C, podem estar distantes das condições idéias para a atividade da enzima sendo analisada. As propriedades ótimas da enzima serão identificadas com o gene de interesse clonado no vetor de expressão pET21a sobre o controle do promotor T7.

d) Ao analisarmos o sequenciamento do fragmento metagenômico clonado em pA1 observamos que não foi possível identificar um códon de parada entre o códon de inicio da proteína predita 1 e o sítio de *Pst*I utilizado para clonar o fragmento. A ausência da região C terminal da enzima pode ter tido impacto sobre a atividade enzimática observada.

Considerando que mesmo a enzima estando truncada, é possível observar a atividade em placa e em meio líquido, especulamos que um códon de terminação presente no vetor e na mesma fase de leitura da proteína predita identificada, está sendo utilizado para terminar a tradução da proteína. Comparativamente, enquanto em pA1 o sítio de *Pst*I utilizado para clonar o fragmento estava a 239 aminoácidos do códon de inicio da transcrição, a sequência da região catalítica de α-amilase de *Acidobacterium* sp. depositada no banco e dados do GenBank (sequência com maior similaridade com a proteína predita 1 de pA1) possui 480 resíduos de aminoácidos, 50% a mais de resíduos.

Sendo assim, considerando a composição aproximada de resíduos de aminoácidos de amilases como sendo em torno de 550 resíduos de aminoácidos, é possível que quase metade da sequência original responsável pela atividade amilolítica não tenha sido clonada em pA1.

De forma semelhante, a α-amilase de *B. subtillis*, intensamente caracterizada e utilizada nos laboratórios de Biologia Molecular e de Enzimologia da Universidade de Brasília, também está truncada. Durante a sua clonagem na construção do plasmídeo pABC1, um fragmento que codificava para 171 resíduos de aminoácidos foi substituído por um fragmento que codificava para 33 resíduos de aminoácidos, não presente na sequência original (De Marco *et al.*, 1996). Foi caracterizada uma diminuição na termoestabilidade devido à sequência estar truncada. Ohdan *et al.* (1999) caracterizaram duas amilases de *B. subtillis*, uma forma nativa e uma forma truncada em 186 resíduos de aminoácidos na região C-terminal (28% menor que a nativa) e descreveram as implicações estruturais do truncamento. Foi observado que as características enzimáticas das duas enzimas eram idênticas, a não ser pela termoestabilidade que aumentou na forma truncada. As análises de estrutura secundária e as predições de estrutura terciária identificaram na forma truncada os três domínios (A, B e C) necessários para que houvesse atividade amilolítica. Nesse contexto, é possível que a proteína predita 1 em pA1 tenha uma termoestabilidade baixa.

e) Alguns trabalhos já sugeriram uma inibição potente e específica de amilases sobre a composição da parede celular de *Neisseria gonorrhoeae* (Mellersh *et al.*, 1979) assim como a identificação do efeito de amilases no desmascaramento de componentes tóxicos do amido no crecimento de *Legionella pneumophila* (Bortner *et al.*, 1983). Desta forma, possivelmente a presença de uma amilase em solução ou no citoplasma, seja mais um estresse para as células em cultura. Apesar disso, algumas linhagens celulares de *E. coli,* como XL10 GOLD, são comercializadas com uma amilase constitutiva que auxilia na transformação bacteriana (Greener, 1996). A importância da expressão de uma possível amilase metagenômica na viabilidade celular será verificada na clonagem e expressão no vetor pET21a, onde há um controle fino da expressão devido à presença do indutor IPTG.

O motivo da ausência de atividade em meio líquido do clone B1.1 (Figura 19) não foi identificado, mas pode estar relacionado aos motivos citados acima que contribuíram para a baixa atividade do clone A1.1.

Principalmente, no caso de B1.1, cogitou-se que a não detecção de sua atividade em meio líquido fosse devido à enzima estar retida no interior na célula. Entretanto, mesmo realizando a sonicação das células, não foi possível detectar a atividade desejada. A sonicação foi realizada até que o meio tornou-se menos turvo, indicando uma diminuição na sua viscosidade. Como não tínhamos como confirmar a lise celular, possivelmente não foi realizado de forma a efetivamente sonicar bem todas as células.

Na verificação da ação de arabinose sobre os halos de atividade dos clones estudados, a distinção clara do efeito da arabinose sobre o tamanho do halo só foi possível com o clone B1.1 (Figuras 11 e 20). A presença da arabinose induz a expressão da proteína AraC em todas as células testadas devido a presença do gene *araC* na célula EPI300, AraC induz a expressão de TrfA requerida para a iniciação da replicação de pCC1FOS. Nos clones da biblioteca de pequenos insertos somente a presença de genes amilolíticos em fase com o promotor P_{BAD} resultariam no aumento do halo de hidrólise em placa. Já nos clones da biblioteca de grandes insertos, na presença de arabinose há um aumento no número de cópias dos plasmídeos de todos os clones. No caso do plasmídeo conter um gene amilolítico funcional, há um maior número de cópia dos genes para serem transcritos e consequentemente há um aumento no halo de hidrólise, como foi observado no clone B1.1.

O perfil de restrição de pA1 (Figura 12) demonstrou que não há sítios de restrição para as enzimas testadas no interior do fragmento. Já o perfil de restrição de pB1 demonstrou não haver sítio de restrição para *Sca*I no interior do fragmento e haver sítios para *Hind*III, *Eco*RI e *Not*I.

O sequenciamento de pA1 e montagem de seu contínuo de leitura permitiram a identificação de dois principais genes (Tabela 2) que tiveram maior semelhança aminoacídica com as proteínas preditas do genoma de *Acidobacterium* sp. MP5ACTX8 (depositado no banco de dados do GenBank no dia 23/06/2010). Este microrganismo foi isolado pela Dra. Minna Mannisto a partir do solo da região Kilpisjarvi, no norte da Finlândia.

O filo *Acidobacterium* é definido por uma ampla coleção de sequências gênicas de rRNA 16S (>1500 no Ribosome Database Project) obtidos de diversas fontes ambientais incluído solos e sedimentos, crosta de solo e dunas de areia, água residual, sistemas de distribuição de água, drenos de minas ácidas, entre outras (Eichorst *et al*., 2007). O filo está ubiquamente distribuído em solos porem possui poucos representantes cultiváveis. *Acidobacterium capsulatum* foi o primeiro membro

descrito do filo *Acidobacteria* e a primeira identificação de uma bactéria heterotrófica acidofílica, diferente de *Acidiphilium*, de um ambiente mineral ácido. Em meados dos anos 90, a coleção crescente de sequências gênicas de rRNA de amostras ambientais, permitiram a identificação de *A. capsulatum* como um membro de um grupo amplo, monofilético e de ramificação profunda dentro de *Eubacteria* (*Hiraishi* et al., 1995).

O filo *Acidobacteria* é atualmente oficialmente reconhecido pelo *Bergey's Manual of Systematic Bacteriology* e inclui três gêneros com representantes já cultivados: *Acidobacterium*, *Geothrix*, e *Holophaga* (Garrity *et al*., 2004). O filo já foi caracterizado como o filo bacteriano com mais representantes em bibliotecas de solos do Arizona (Dunbar *et al*., 2002) e foi tido como o segundo mais abundante na biblioteca de clones de rDNA 16S da microbiota do solo do cerrado *sensu stricto* de Quirino *et al*., 2009.

A diversidade de sequências atribuídas ao filo é similar à encontrada no metabolicamente diverso filo *Proteobacteria*, sugerindo uma extensa diversidade metabólica em *Acidobacteria*. Apesar do potencial metabólico de acidobacteria ser pouco descrito, sua abundância sugere um importante papel na ciclagem de nutrientes em solos (Eichorst *et al*., 2007).

Desta forma, a identificação de um clone com similaridade de sequência com o genoma da *Acidobacterium* sp. é plausível e interessante devido ao pouco conhecimento existente na literatura sobre os componentes deste filo e devido ao possível importante papel desempenhado pelo integrantes deste filo no solo.

Na proteína predita 1, foram identificadas três regiões da sequência proteica com similaridade às das assinaturas das enzimas classificadas na superfamília das α-amilases(Anexo 8). Entretanto, a sequência difere significativamente das sequências de amilases já conhecidas. Como citado anteriormente, a triagem por atividade de bibliotecas metagenômicas apesar de ser uma metodologia de baixo rendimento, é uma das poucas formas de obtermos sequências realmente novas de enzimas.

Sequências anotadas como trealose 6-fosfato hidrolases foram identificadas pelo programa BLASTx como tendo maior similaridade com a sequência da proteína predita 1. Entretanto, como pôde ser observado na filogenia construída (Figura 15), a proximidade da proteína predita 1 com a anotação da região catalítica da α-amilase de *Acidobacterium* sp. é maior do que com as demais sequências identificadas. Estas duas últimas estão mais próximas às trealose 6-fosfato hidrolases do que com as amilases utilizadas para compor a filogenia. A identificação de domínios conservados da família

de glicosil e glicosídio hidrolases pelo programa InterProScan reforçam o papel da enzima no grupo das Glicosil Hidrolases, não necessariamente como uma amilase e sim, possivelmente, como uma glicosil hidrolase pouco específica, que possui atividade sobre amido.

A trealose é um dissacarídeo composto por duas moléculas de glicose ligadas em α,α,1-1, portanto, não redutor A trealose-6-fosfato hidrolase catalisa a quebra de tralose-6-fosfato em glicose-6-fosfato e glicose e permite a utilização de trealose, como única fonte de carbono. (Rimmele & Boos, 1994). Em condições otimizadas de expressão, a atividade específica da proteína predita 1 de pA1, tanto sobre amido, como sobre trealose, identificará a principal papel biológico desta enzima.

A amplificação por PCR da ORF da proteína predita 1 na versão sem as bases do vetor e com as bases do vetor já foi realizada (Figura 16) e o tamanho do fragmento M aparenta estar um pouco abaixo do esperado. Entretanto, é dificl estimar o tamanho com o marcador utilizado. Os *amplicons* purificados e clonados no vetor pBluescript KS II + estão sendo utilizados para transformar *E. coli* XL10 Gold.

Neste contexto, a caracterização bioquímica da proteína predita 1, clonada no vetor de expressão pET21a, sob o controle do promotor forte e regulável T7, será a forma de verificarmos o real potencial da utilização desta enzima em processos biotecnológicos e sua ação sobre diferentes substratos. Apesar da alta especificidade ao substrato ser interessante e indispensável em determinados processos como na produção de xaropes e na panificação, possivelmente, a sua utilização na alimentação animal ou processos semelhantes seja viável.

Com relação às proteínas preditas 2 e 4, sugerimos que não sejam sequências protéicas funcionais devido ao seu pequeno tamanho e ausência de domínios putativos conservados.

A proteína predita 3 também teve maior similaridade com uma sequência do genoma de *Acidobacterium* sp. e parece ser bem conservada entre microrganismos por ter grande similaridade com um grande número de sequências depositadas no banco de dados do GenBank (dados não mostrados). A presença do domínio Plug receptor TonB-dependente é um indício de que esta proteína efetivamente seja o domínio de dobra dos receptores TonB dependentes. A proteína TonB interage com estes tipos de receptores de alta afinidade, ATP-dependentes, presentes na membrana externa, que desempenham o papel de permitir a entrada de substratos para o espaço periplasmático. Na ausência de TonB estes receptores ligam seus substratos porém não conseguem efetuar seu

transporte ativo (Buchanan *et al.*, 2004). Considerando a proximidade deste gene ao gene identificado como uma glicosil hidrolase e a sua possível ação sinérgica na utilização de determinado substrato, como o amido, é possível a sua relação e anotação, apesar de não estarem organizados na forma de um *operon*.

O plasmídeo pB1 teve seu tamanho aproximado identificado pela soma do tamanho dos fragmentos resultantes das digestões realizadas: 30 kb do fragmento metagenômico e de 8,0 kb do vetor de clonagem. A escolha inicial de clonar somente os fragmentos com tamanho acima de 3,5 kb foi justificada por uma tentativa de obtermos um subclone com o fenótipo original. A não identificação do fenótipo é um indício de que a digestão de pB1 com NotI pode ter gerado um fragmento contendo o gene responsável pelo fenótipo amilolítico truncado e não funcional, indicando também a necessidade de clonarmos todos os fragmentos resultantes da digestão.

O sequenciamento dos subclones de pB1 permitiram a identificação de aproximadamente 15 kb do fragmento metagenômico. Dentre os motivos que explicam o erro cometido de clonar e sequenciar um fragmento referente ao vetor pCC1FOS, temos que após muitas tentativas mal sucedidas de clonarmos a banda de aproximadamente 9,0 kb em pBluescipt II KS, pensamos não ser possível esta clonagem devido ao fragmento conter as origens de replicação do vetor pCC1FOS.

A obtenção de um plasmídeo funcional resultante da clonagem de pCC1FOS em pBluescript II KS, se deve ao fato de que as origens de replicação de ambos (ColE1 para pBluescript II KS, e OriV e Ori2 e para pCC1FOS) permitem a coexistência em uma mesma célula.

A análise das proteínas preditas dos fragmentos metagenômicos dos plasmídeos, pB1.3, pB1.4 e pB1.5 (Tabelas 3, 4 e 5) não resultou na identificação de uma proteína predita cuja função seja diretamente relacionada ao halo de hidrólise observado em placa. Dentre as proteínas preditas identificadas, algumas julgamos possuir características interessantes e abordaremos com mais detalhes.

A proteína predita 4 possui máxima similaridade atribuída pelo programa BLASTp (e-value 0.0) com sequências anotadas como subunidades α da malonato descarboxilase. Esta enzima catalisa a descarboxilação de malonato a acetato e dióxido de carbono. Os genes das subunidades que compõem a malonato descarboxilase (α, β, δ, ε e γ) normalmente são encontrados na forma do *cluster* mdc, ou seja, sequencialmente no genoma (Clark, 1981). Não foi possível identificar outra subunidade apesar da localização central da proteína predita 4 na sequência nucleotídica do fragmento. O

cladograma da proteína predita 4 (Figura 28) com sequências de maior similaridade identificada pelo BLASTp, demonstra maior semelhança da sequência da proteina predita 4 com sequências protéica do gênero *Burkholderia*.

Três outras proteínas preditas dos plasmídeos decorrentes da subclonagem de pB1 apresentaram maior similaridade com sequências do gênero *Burkholderia*. Inclusive, por estar dentre as proteínas preditas com menores e-values atribuídos pelo programa BLASTp, a proteína predita 4 do plasmídeo pB1.5, também foi utilizada para construir uma filogenia e teve uma evidente relação com o gênero *Burkholderia* (Figura 30).

Membros do gênero *Burkholderia* são muito abundantes e ocupam diversos nichos ecológicos, dentre eles: solos e hospitais. O gênero é muito conhecido por seus componentes causadores de infecções em animais e humanos, mas uma importância crescente tem sido atribuída ao gênero pela sua associação com plantas, sendo como componentes livres da rizosfera ou como membros epifíticos e endofíticos. Várias linhagens são conhecidas por aumentar a resistência de plantas, contribuírem para um melhor manejo hídrico e melhorarem a fixação de nitrogênio e adaptação generalizada do hospedeiro a estresses ambientais (Salles *et al*., 2002; Van Elsas *et al*., 2002; Coenye & Vandamme, 2003; Nowak & Shulaev, 2003; Compant *et al*., 2005; Sessitsch *et al*., 2005; Valvano *et al*., 2005; Ait Barka *et al*., 2006; Barrett & Parker, 2006; Janssen, 2006; Balandreau & Mavingui, 2007, citados por Compant *et al*., 2008). Membros do genero *Burkholderia* são utilizados em diversas aplicações biotecnológicas, dentre elas a biorremediação de compostos xenobióticos recalcitrantes e no controle biológico de doenças vegetais resultando no aumento do rendimento de plantações (Parke & Gurian-Sherman, 2001 citados por O'Sullivan & Mahenthiralingam, 2001).

O gênero é um componente do grupo ß-proteobacteria, grupo identificado por Quirino *et al*., 2009, como presente no solo de cerrado *sensu stricto*, porem pouco abundante.

Apesar da identificação inconclusiva, ainda há um fragmento de 9 kb para ser seqüenciado do plasmídeo pB1 e caso haja a identificação de genes relacionados a estes nichos ecológicos, considerando a possível relação desde fragmento com o gênero *Burkholderia*, uma melhor caracterização provavelmente se tornará interessante.

As proteínas preditas 5, 6 e 7 do plasmídeo pB1.4, foram ditas semelhantes a proteínas relacionadas ao metabolismo de ácidos graxos. Como as proteínas 5 e 6 foram identificadas como semelhantes à mesma anotação na banco de dados do GenBank, o

alinhamento local foi realizado de forma a verificar que as sequências *in tandem* não são resultados de falhas na replicação. A baixa similaridade entre as sequências indica que as sequências provavelmente possuem atividades distintas, possivelmente ainda não bem diferenciadas.

Pirrolo quinolina quinona (PQQ) é um fator de redução-oxidação que atua como co-fator para várias desidrogenases bacterianas, dentre elas a metanol desidrogenase e glicose desidrogenase (Puehringer *et al*., 2008). Em adição ao papel de co-fator, PQQ também é considerado um fator trófico importante para o crescimento e metabolismo de bactérias, particularmente bactérias metilotróficas. As bactérias metilotróficas são em maioria enquadrada nas subclasses de α, β e γ de Proteobacteria ou no grupo de bactérias Gram-positivas e descobertas recentes identificaram a primeira espécie metilotrófica à ordem *Burkholderiales* (Chistoserdova *et al*., 2010).

PQQ já foi demonstrada como um fator chave no efeito positivo da interação entre rizobactérias e plantas no aumento de altura, peso, número de flores e frutos (Choi *et al*., 2008), atribuído, em parte, ao papel da PQQ de aumentar o consumo de fosfato pela acidificação do solo (Sato & Toriyama, 2009, citado por Rucker *et al*., 2009). Em humanos e demais animais seu papel como vitamina é discutido, mas supõem-se que seja antes de tudo, um importante fator de sinalização celular, não presente em bactérias intestinais, como *E. coli* (Rucker *et al*., 2009).

Desta forma, a presença deste gene com similaridade a sequências depositadas no banco de dados do GenBank como *Burkholderia* e seu possível papel no crescimento metilotrófico, poderiam auxiliar na obtenção de uma cultura do microrganismo de onde o DNA que compõem o plasmídeo pB1 foi isolado, caso isto se mostre interessante.

Os resultados aqui apresentados e discutidos constituem o trabalho inicial de caracterização de clones que apresentaram aparente atividade amilolíticas de bibliotecas metagenômicas da microbiota do solo do Cerrado. No conjunto, sequências novas foram obtidas com potencial de serem utilizadas para ensaios de expressão, teste de expressão e comparações de atividades, estudos filogenéticos, dentre outros. Estudos de expressão da proteína com perfil e domínios de glicosil hidrolase estão em andamento e determinaram as suas reais propriedades e aplicações.

7. CONCLUSÕES E PERSPECTIVAS

A utilização das ferramentas metagenômicas citadas neste trabalho permitiram aumentar o conhecimento sobre as enzimas hidrolíticas presentes na microbiota do solo do Cerrado. A utilização da triagem baseada em função nos permitiu identificar clones com a atividade de interesse e obter a sequência responsável pela atividade observada, independente do conhecimento que já possuímos sobre sequências hidrolíticas. Obtivemos a sequência de uma enzima hidrolítica com pouca similaridade a sequências depositadas em banco de dados, demonstrando o potencial de conseguirmos obter sequências novas com esta metodologia. A caracterização bioquímica da proteína expressa com um promotor forte, em um plasmídio de alto número de cópia por célula, está em processo de conclusão e demonstrará o real papel e potencial da enzima identificada em processos biotecnológicos.

Metade do plasmídeo pB1 foi sequenciado e genes novos possivelmente relacionados ao gênero *Burkholderia* foram identificados. Os fragmentos resultante da digestão de pB1 com *Not*I ainda não seqüenciados, serão seqüenciado na forma de uma sub-biblioteca decorrente da sua digestão com *Sau*3A1 e clonagem em pBluescript II KS.

Novas abordagens envolvendo mais clones hidrolíticos representam uma fonte de novas enzimas com diversas aplicações e a recente inauguração do Centro de genômica de alto desempenho do Distrito Federal é um estímulo para trabalhos que envolvam o sequenciamento de DNA metagenômico.

8. REFERÊNCIAS BIBLIOGRÁFICAS

Ait Barka E, Nowak J & Clement C (2006) Enhancement of chilling resistance of inoculated grapevine plantlets with a plant growth-promoting rhizobacterium, *Burkholderia phytofirmans* strain PsJN. *Appl Environ Microbiol*, 72: 7246–7252.

Allbrands (2004) Indústria de Alimentos Ltda. *Informações técnicas sobre o processamento dos produtos para alimentação infantil*. Colombo, PR.

Amann RI, Ludwig W & Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev* 59:143-69.

Astolfi-Filho, S Galembeck EV, Faria JB & Schenberg Frascino AC (1986) Stable yeast transformants that secrete functional amylase encoded by cloned mouse pancreatic cDNA. *Nat Biotech* 4: 311-315.

Balandreau J & Mavingui P (2007) Beneficial interactions of *Burkholderia* spp. with plants. *Burkholderia*: *Molecular Biology and Genomics* (Vandamme P & Coenye T, eds), pp. 129–151. Horizon Scientific Press, Norwich, UK.

Barrett CF & Parker MA (2006) Coexistence of *Burkholderia, Cupriavidus,* and *Rhizobium* sp. nodule bacteria on two *Mimosa* spp. in Costa Rica. *Appl Environ Microbiol*, 72: 1198–1206.

Bernstein JR, Bulter T, Shen CR & Liao JC (2007) Directed evolution of ribosomal protein S1 for enhanced translational efficiency of high GC *Rhodopseudomonas palustris* DNA in *Escherichia coli*. *J Biol Chem*, 282: 18929-18936.

Bertoldo C & Antranikian G (2002) Starch-hydrolyzing enzymes from thermophilic archaea and bacteria. *Curr Opin Chem Biol* 6: 151-160.

Bickerstaff GF (1987) *Enzymes in industry and medicine*. 1a ed. Edward Arnold. Londres, RU.

Bisgaard-Frantzen H, Borchert T, Svendsen A, Thellersen MH & Van Der Zee P. (1995) PCT Patent Application. WO 95/10603.

Borges TFB (2009) *Construção de um vetor integrativo de múltiplas cópias para* Saccharomyces cerevisiae *utilizando sequências delta*. 120 p. Dissertação (Mestrado em Ciências Biológicas) - Universidade de Brasília.

Borneman J (1999) Culture-independent identification of microorganisms that respond to specified stimuli. *Appl Environ Microbiol* 65: 3398-400.

Bortner CA, Miller RD & Arnold RR (1983) Effects of alpha-amylase on in vitro growth of *Legionella pneumophila*. *Infect Immun*, 41(1): 44-9.

Bruinenberg PM, Hulst AC, Faber A, Voogd RH (1996) A process for surface sizing or coating of paper. European Patent Application EP 0,690,170 A1.

Buchanan SK, Evans RW, Ghirlando R, Oke M, Sarra R, Farnaud S & Gorringe AR (2004) The plug domain of a neisserial TonB-dependent transporter retains structural integrity in the absence of its transmembrane beta-barrel. *FEBS Lett,* 564 (3): 294–300.

Buleon A, Colonna P, Planchot V & Ball S (1998) Starch granules: structure and biosynthesis. *Int J Biol Macromol* 23: 85-112.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V & Henrissat B (2008) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res in press [PMID: 18838391].

Chae JC, Song B & Zylstra GJ (2008) Identification of genes coding for hydrolytic dehalogenation in the metagenome derived from a denitrifying 4-chlorobenzoate degrading consortium. *FEMS Microbiol Lett*, 281: 203-209.

Chistoserdova L, Kalyuzhnaya MG & Lidstrom ME (2010) The expanding world of methylotrophic metabolism. *Annu Rev Microbiol,* 63: 477-499.

Choi O, Kim J, Kim JG, Jeong Y, Moon JS, Park CS & Hwang I (2008) Pyrroloquinoline quinone is a plant growth promotion factor produced by *Pseudomonas fluorescens* B16. *Plant Physiol*, 146: 657-668.

Clark D (1981) Regulation of fatty acid degradation in *Escherichia coli*: analysis by operon fusion. *J Bacteriol,* 148(2): 521-6.

Coenye T & Vandamme P (2003) Diversity and significance of *Burkholderia* species occupying diverse ecological niches. *Environ Microbiol,* 5: 719-729.

Cohen SN, Chang ACY & Hsu L (1972) Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc Natl Acad Sci,* 69: 2110-2114.

Compant S, Nowak J, Coenye T, Clement C & Ait Barka E (2008) Diversity and occurrence of *Burkholderia* spp. in the natural environment. *FEMS Microbiol Rev,* 32: 607-626.

Compant S, Reiter B, Sessitch A, Nowak J, Clement C & Ait Barka E (2005) Endophytic colonization of Vitis vinifera L. by plant growth-promoting bacterium Burkholderia sp. strain PsJN. *Appl Environ Microbiol,* 71: 1685-1693.

De Castro AP (2008) *Utilização de DNA metagenômico para caracterização e bioprospecção de comunidades microbianas de solo de cerrado*.178 p. Dissertação (Mestrado em Ciências Genômicas e Biotecnologia) – Universidade Católica de Brasília.

De Marco JL, Bataus LA, Valencia FF, Ulhoa CJ, Astolfi-filho S & Felix CR (1996) Purification and characterization of a truncated Bacillus subtilis alpha-amylase produced by Escherichia coli. *Appl Microbiol Biotechnol*, 44(6):746-52.

De Castro AP, Quirino BF, Pappas G, Kurokawa AS, Leonardecz E & Krüger RH (2008) Diversity of soil fungal communities of Cerrado and its closely surrounding agriculture fields. *Archives of Microbiology*, 190(2): 129-139.

Dickschat JS (2010) Quorum sensing and bacterial biofilms. *Nat Prod Rep*, 27(3): 343-369.

Dunbar J, Barns SM, Ticknor LO & Kuske CR (2002) Empirical and theoretical bacterial diversity in four Arizona soils. *Appl Environ Microbiol*, 68:3035-3045.

Eichorst SA, Breznak JA & Schmidt TM (2007) Isolation and characterization of soil bacteria that define *Terriglobus* gen. nov., in the Phylum *Acidobacteria. Appl Environ Microbiol,* 73(8): 2708–2717.

Elsied H, Stokes HW, Nakamura T, Kitamura K, Fuse H & Maruyama A (2007) Novel and diverse integron integrase genes and integron-like gene cassettes are prevalent in deep-sea hydrothermal vents. *Environ Microbiol*, 9: 2298-2312.

Ferrer M, Beloqui A, Golyshina OV, Plou FJ, Neef A, Chernikova TN, Fernández-Arrojo L, Ghazi I, Ballesteros A, Elborough K, Timmis KN, & Golyshin PN (2007) Biochemical and structural features of a novel cyclodextrinase from cow rumen metagenome. *Biotechnol J* 2:207-13.

Ferrer M, Chernikova TN, Timmis KN & Golyshin PN (2004) Expression of a temperature-sensitive esterase in a novel chaperone-based *Escherichia coli* strain. *Appl Environ Microbiol*, 70: 4499-4504.

Fieseler L, Hentschel U, Grozdanov L, Schirmer A, Wen G, Platzer M, Hrvatin S, Butzke D, Zimmermann K & Piel J (2007) Widespread occurrence and genomic context of unusually small polyketide synthase genes in microbial consortia associated with marine sponges. *Appl Environ Microbiol* 73: 2144-2155.

Fischer E (1894) Einfluss der Configuration auf die Wirkung der Enzyme". *Ber Dt Chem Ges,* 27: 2985–2993.

Fujiwara S, Takagi M & Imanaka T (1998) Archaeon *Pyrococcus kodakaruensis* KOD1: application and evolution. *Biotechnol Annu Rev*, 4: 259-284.

Fuqua WC, Winans SC & Greenberg EP (1994) Quorum sensing in bacteria: the luxr-luxi family of cell density-responsive transcriptional regulator. *J Bacteriol*, 176(2): 269-275.

Fuwa H (1954) A new method for microdetermination of amylase activity by the use of amylose as the substrate.*The journal of Biochemistry,* 41: 583-603.

Gabor EM, Alkema WB & Janssen DB (2004) Quantifying the accessibility of the metagenome by random expression cloning techniques. *Environ Microbiol,* 6: 879-886.

Garrity GM (2005) *Bergey's Manual of Systematic Bacteriology*, 2 ed, vol. New York: Springer.

Garrity GM, Bell JA, & Lilburn TG (2004) Taxonomic outline of the prokaryotes. *Bergey's manual of systematic bacteriology*, 2nd ed., release 5.0. New York: Springer-Verlag.

Goldberg SM, Johnson J, Busam D, Feldblyum T, Ferriera S, Friedman R, Halpern A, Khouri H, Kravitz SA, Lauro FM, Li K, Rogers YH, Strausberg R, Sutton G, Tallon L, Thomas T, Venter E, Frazier M & Venter JC (2006) A Sanger/pyrosequencing hybrid approach for the generation of high-quality draft assemblies of marine microbial genomes. *Proc Natl Acad Sci,* 103: 11240-11245.

Greener AL (1996) Process of producing highly transformable bacterial cells and cells produced thereby. Patent application 5512468.

Guan C, Ju J, Borlee BR, Williamson LL, Shen B, Raffa KF & Handelsman J (2007) Signal mimics derived from a metagenomic analysis of the gypsy moth gut microbiota. *Appl Environ Microbiol*, 73:3669-3676.

Gupta R, Gigras P, Mohapatra H, Goswami VK & Chauhan B (2003) Microbial a-amylases: a biotechnological perspective. *Process Biochemistry* 38: 1599-1616

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nuc Acids Symp*, 41:95-98.

Handelsman J, Rondon MR, Brady SF, Clardy J & Goodman RM (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* 5: R245-9.

Head IM, Saunders JR & Pickup RW (1998) Microbial evolution, diversity, and ecology: a decade of ribosomal RNA analysis of uncultivated microorganisms. *Microb Ecol*, 35:1–21

Hehre EJ, Genghof DS & Okada G (1971) The alpha-amylases as glycosylases, with wider catalytic capacities than envisioned or explained by their representation as hydrolases. *Arch Biochem Biophys*,142(1):382-93.

Hendriksen HV, Pedersen S & Bisgard-Frantzen H (1999) A process for textile warp sizing using enzymatically modified starches. Patent Application WO 99/35325

Henrissat B & Davies G (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Struc Biol* 7: 637-644.

Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* 280: 309-316.

Hiraishi A, Kishimoto N, Kosako Y, Wakao N & Tano T (1995) Phylogenetic position of the menaquinone-containing acidophilic chemo-organotroph *Acidobacterium capsulatum*. *FEMS Microbiol Lett*, 1;132(1-2):91-4.

Hugenholtz P & Pace NR (1996) Identifying microbial diversity in the natural environment: a molecular phylogenetic approach. *Trends Biotechnol* 14:190-7.

Imanaka T & Kuriki T (1989) Pattern of action of *Bacillus steurothermophilus* neopullulanase on pullulan. *J Bacterial*, 171: 369-374.

IUBMB: *Enzyme nomenclature. Recomendations*. 1992. San Diego: Academic Press

Janecek S & Sevcik J (1999) The evolution of starch-binding domain. *FEBS Letters*, 456: 119-125.

Janssen PH (2006) Identifying the dominant soil bacteria taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl Environ Microbiol,* 72: 1719–1728.

Jenkins PJ, Cameron RE & Donald AM (1993) *Stärke/Starch*. 45: 417-420.

Kazimierczak KA, Scott KP, Kelly D & Aminov RI (2009) Tetracycline resistome of the organic pig gut. *Appl Environ Microbiol*, 75: 1717-1722.

Kent AD & Triplett EW (2002) Microbial communities and their interactions in soil and rhizosphere ecosystems. *Annu Rev Microbiol* 56: 211–36.

Kitahata S & Okada S (1982) Comparison of action of cyclodextrin glucanotransferase from *Bacillus megaterium*, *B. circulans*, *B. stearothermophiius* and *B. macerans*. *J. Jpn Sot Starch Sci*, 29: 13-18.

Koenig JE, Sharp C, Dlutek M, Curtis B, Joss M, Boucher Y & Doolittle WF (2009) Integron gene cassettes and degradation of compounds associated with industrial waste: the case of the Sydney tar ponds. *PLoS ONE*, 4(4): e5276.

Kulp K (1993) Enzymes as dough improvers. *In*: Kamel BS, Stauffer CE. Advances in baking technology. New York: Blackie Academic and Professional, p. 152-178.

Kuriki T & Imanaka T (1999) The concept of the a-amylase family: structural similarity and common catalytic mechanism. *J Biosci Bioeng,* 87(5): 557-565.

Liu Z, Zhang G & Liu S (2004) Constructing an amylolytic brewing yeast *Saccharomyces pastorianus* suitable for accelerated brewing. *J Biosci Bioeng,* 98(6): 414–419

Lo HF, Lin LL, Chiang WY, Chie MC, Hsu WH & Chang CT (2002) Deletion analysis of the C-terminal region of the alpha-amylase of *Bacillus* sp. strain TS-23. *Arch Microbiol*, 178(2): 115-23.

MacGregor EA, Janecek S & Svensson B (2001) Relationship of sequence and structure to specificity in the α-amylase family of enzymes. *Biochim Biophys Acta*, 1549(1): 1-20.

Machovic M & Janecek S (2006) The evolution of putative starch-binding domains. *FEBS Lett,* 580: 6349–6356.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437: 376–380

Martinez A, Kolvek SJ, Yip CLT, Hopke J, Brown KA, MacNeil IA, & Osburne MS (2004) Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts. *Appl Environ Microbiol,* 70: 2452–2463.

Marzorati M, de Ferra F, Van Raemdonck H, Borin S, Allifranchini E, Carpani G, Serbolisca L, Verstraete W, Boon N & Daffonchio D (2007) A novel reductive dehalogenase, identified in a contaminated groundwater enrichment culture and in *Desulfitobacterium dichloroeliminans* strain DCA1, is linked to dehalogenation of 1,2-dichloroethane. *Appl Environ Microbiol*, 73: 2990-2999.

Matsuura Y, Kusunoki M., Harada W & Kakudo M (1984) Structure and possible catalytic of Taka-amylase. *A J Biochem*, 95: 697-702.

Mayumi D, Akutsu-Shigeno Y, Uchiyama H, Nomura N & Nakajima-Kambe T (2008) Identification and characterization of novel poly(DL-lactic acid) depolymerases from metagenome. *Appl Microbiol Biotechnol* 79:743-750.

Mellersh A, Clark A & Hafiz S (1979) Inhibition of Neisseria gonorrhoeae by normal human saliva. *Br J Vener Dis,* 55: 20-23

Moraes LMP (2004) Amilases. *In:* Said S & Pietro RCL. *Enzimas como agentes biotecnológicos*. Ribeirão Preto: Legis Summa, p.223-242.

Moraes LMP, Astolfi-Filho S & Ulhoa CJ (1999) Purification and some properties of an α-amylase glucoamylase fusion protein from *Saccharomyces cerevisiae*. *World J Microb Biotech,* 15: 561-564.

Moraes LMP, Astolfi-Filho S & Oliver SG (1995) Development of yeast strains for the efficient utilization of starch: evaluation of constructs that express α-amylase and glucoamylase separately or as bifuntional fusion proteins. *Appl Microb Biotech,* 43 (6): 1067-1076.

Mori T, Mizuta S, Suenaga H & Miyazaki K (2008) Metagenomic screening for bleomycin resistance genes. *Appl Environ Microbiol,* 74: 6803-6805.

Morimoto S & Fujii T (2009) A new approach to retrieve full lengths of functional genes from soil by PCR-DGGE and metagenome walking. *Appl Microbiol Biotechnol*, 83: 389-396.

Nagashima T, Toda S, Kitamoto K, Gomi K, Kumagai C & Toda H (1992) Site-directed mutagenesis of the catalytic active-site residues of Taka-amylase A. *Biosci Biotech Biothem*, 56: 207-210.

Nelson DL, Cox MM (2006) *Lehninger Princípios de bioquímica*. 4ed. São Paulo: Sarvier. 1100 p.

Newman JR & Fuqua C (1999) Broad-host-range expression vectors that carry the L-arabinose-inducible Escherichia coli araBAD promoter and the araC regulator. *Gene,* 227: 197-203.

Novick RP (1962) Staphylococcal Penicillinase and the New Penicillins. *Biochem J,* 83: 229-235.

Novozymes (2004) Latin America Ltda. Enzimas para todas as Aplicações. Disponível em: http://www.novozymes.com/produtos.pdf. Acesso em 1 de abril de 2010.

Nowak J & Shulaev V (2003) Priming for transplant stress resistance in in vitro propagation. *In Vitro Cell Dev Biol-Plant,* 39: 107–124.

Ohdan K, Kuriki T, Kaneko H, Shimada J, Takada T, Fujimoto Z, Mizuno H & Okada S (1999) Characteristics of two forms of α-amylases and structural implication. *Appl Envir Microbiol*, 65(10): 4652-4658.

Okada, S. & Mizokami, K (1980) Action on pullulan by *Streptococcns bovis* α-amylase. *J Jp. Sot Starch Sci*, 27, 127-133 (em japonês).

Ono A, Miyazaki R, Sota M, Ohtsubo Y, Nagata Y & Tsuda M (2007) Isolation and characterization of naphthalene-catabolic genes and plasmids from oil-contaminated soil by using two cultivation-independent approaches. *Appl Microbiol Biotechnol*, 74: 501-510.

O'Sullivan LA & Mahenthiralingam E (2005) Biotechnological potential within the genus *Burkholderia*. *Lett Appl Microbiol*, 41(1): 8-11.

Pace NR (1997) A molecular view of microbial diversity and the biosphere. *Science,* 276: 734-4079.

Pace NR, Stahl DA, Lane DJ & Olsen GJ (1985) Analyzing natural microbial populations by rRNA sequences. *ASM News,* 51: 4-12

Pace NR, Stahl DA, Lane DJ & Olsen GJ (1986) The analysis of natural populations by ribosomal RNA sequences. *Adv Microbial Ecol,* 9: 1-55

Pandey A, Nigam P, Soccol CR, Soccol VT, Singh D & Mohan R (2000) Advances in microbial amylases. *Biotechnol Appl Biochem*, 31: 135-152.

Parke JL & Gurian-Sherman D (2001) Diversity of the *Burkholderia cepacia* complex and implications for risk assessment of biological control strains. *Annu Rev Phytopathol,* 39, 225–258.

Pinto MN (1990) Cerrado, caracterização, ocupação e perspectivas. Editora UnB, Brasília, D.F.

Puehringer S, Metlitzky M & Schwarzenbacher R(2008) The pyrroloquinoline quinone biosynthesis pathway revisited: A structural approach *BMC Biochem,* 9: 8.

Quaglia G (1991) *Ciência y Tecnologia de la Panificación. Zaragoza*: Editorial Acribia, 485 p.

Quirino BF, Pappas GJ, Tagliaferro AC, Collevatti RG, Leonardecz E, Da Silva MRSS, Bustamante MMC & Krüger RH (2009) Molecular phylogenetic diversity of bacteria associated with soil of the savanna-like Cerrado vegetation. *Microbiological Research*, 164(1): 59-70.

Ratter JA, Ribeiro JF & Bridgewater S (1997) The Brazilian cerrado vegetation and threats to its biodiversity. *Ann Bot*, 80(3): 223-230.

Ribeiro-dos-Santos MGG (1988) Melhoramento genético e caracterização de novas cepas amilolíticas de leveduras obtidas por engenharia genética. 93 p. Dissertação (Mestrado em Ciências Biológicas) - Universidade se São Paulo.

Richardson TH, Tan X, Frey G, Callen W, Cabell M, Lam D, Macomber J, Short JM, Robertson DE & Miller C (2002) A novel, high-performance enzyme for starch liquefaction. *J Biol Chem,* 277: 26501–26507.

Riesenfeld CS, Schloss PD & Handelsman J (2004) Metagenomics: Genomic analysis of microbial communities. *Annu Rev Genet*, 38: 525-52.

Rimmele M & Boos W (1994) Trehalose-6-phosphate hydrolase of *Escherichia coli. J Bacteriol,* 176(18): 5654–5664.

Robe P, Nalin R, Capellano C, Vogel TM, C & Simonet, P (2003). Extraction of DNA from soil. *Eur J Soil Biol,* 39: 183-190.

Rodriguez-Sanoja R., Oviedo N & Sanchez S (2005) Microbial starch-binding domain. *Curr Opin Microbiol,* 8: 260-267.

Rondon MR, Raffel SJ, Goodman RM & Handelsman J (1999) Toward functional genomics in bacteria: analysis of gene expression in Escherichia coli from a bacterial artificial chromosome library of *Bacillus cereus*. *Proc Natl Acad* Sci, USA 96: 6451-5.

Rondon, MR., August PR, Bettermann AD, Brady SF, Grossman TH, Liles MR, Loiacono KA, Lynch BA, MacNeil IA, Minor C, Tiong CL, Gilman M, Osburne MS, Clardy J, Handelsman J & Goodman RM (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl Environ Microbiol,* 66: 2541-7.

Rucker R, Chowanadisai W & Nakano M (2009) Potential physiological importance of pyrroloquinoline quinone. *Altern Med Rev*,14: 268-277.

Sakano Y, Fukushima J & Kobavashi T (1983) Hvdrolvsis of an α-1,4- and α-1,6-glucosidic linkages in trisaccharides by *Thermoactinomyces vulgaris* α -amylase. *Agric Biol Chem*, 47: 2211-2216.

Salles JF, De Souza FA & Van Elsas JD (2002) Molecular method to assess the diversity of *Burkholderia* species in environmental samples. *Appl Environ Microbiol* 68: 1595-1603.

Sambrook J & Russel DW (2001) Molecular Cloning: A laboratory Manual. 3ª Edição. Cold Spring Harbor Laboratory Press, New York.

Sánchez OJ &Cardona CA (2008) Trends in biotechnological production of fuel ethanol from different feedstocks. *Bioresour Technol,* 99(13): 5270-5295.

Sato K & Toriyama M (2009) Effect of pyrroloquinoline quinone (PQQ) on melanogenic protein expression in murine B16 melanoma. *J Dermatol Sci,* 53: 140-145.

Sawai T, Takahashi I & Yamagishi S (1978) Iodometric assay method for beta-lactamase with various beta-lactam antibiotics as substrates. *Antimicrob Agents Chemother,* 13(6): 910-913.

Saxena RK, Dutt K, Agarwal L & Nayyar P (2007) A highly thermostable and alkaline amylase from a *Bacillus* sp. PN5. *Bioresour Technol,* 98: 260-265.

Schloss PD & Handelsman J (2005). Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol*, 71(3): 1501-1506.

Schmeisser C, Steele H & Streit WR (2007) Metagenomics, biotechnology with non-culturable microbes. *Appl Microbiol Biotechnol*, 75:955–962.

Schmeisser C, Steele H & Streit WR. (2007) Metagenomics, biotechnology with non-culturable microbes. *Appl Microbiol Biotechnol,* 75: 955–962.

Sessitsch A, Coenye T, Sturz AV, Vandamme P, Ait Barka E, Faure D, Reiter B, Glick BR, Wang-Pruski G & Nowak J (2005) *Burkholderia phytofirmans* sp. nov., a novel plant-associated bacterium with plant beneficial properties. *Int J Syst Evol Microbiol,* 55: 1187–1192.

SF Altschul, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, & Lipman DJ. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res,* 25:3389-402.

Sharma S, Khan FG & Qaki GN (2010) Molecular cloning and characterization of amylase from soil metagenomic library derived from Northwestern Himalayas. *Appl Microbiol Biotechnol*, 86 (6): 1821-1828.

Silva JMC & Bates JM (2002) Biogeographic patterns and conservation in the South American Cerrado: a tropical savanna hotspot. *BioScience,* 52(3): 225-233.

Simon C, Herath J, Rockstroh S, Daniel R (2009) Rapid identification of genes encoding DNA polymerases by function-based screening of metagenomic libraries derived from glacial ice. *Appl Environ Microbiol*, 75: 2964-2968.

Sleator RD, Shortall C & Hill C (2008) Metagenomics. *Lett Appl Microbiol,* 47:361-366.

Smith AM (2001) The biosynthesis of starch granules. *Biomacromolecules*, 2: 335-341.

Souza MBNS (1986) *Clonagem molecular e expressão do gene de α-amilase de Bacillus sp.* 80 p. Dissertação (Mestrado em Biologia Molecular) - Universidade de Brasília.

Spier MR (2008) Produção de enzimas amilolíticas fúngicas α-amilase e amiloglucosidase por fermentação no estado sólido. 155 f. Dissertação (Mestrado em Tecnologia de alimentos) - Universidade Federal do Paraná.

Stahl DA, Lane DJ, Olsen GJ, Pace NR (1984) Analysis of hydrothermal vent-associated symbionts by ribosomal RNA sequences. *Science,* 224: 409-411.

Steele HL, Jaeger KE, Daniel R & Streit WR (2009) Advances in Recovery of Novel Biocatalysts from Metagenomes. *J Mol Microbiol Biotechnol,* 16: 25-37.

Streit WR, Daniel R & Jaeger KE (2004) Prospecting for biocatalysts and drugs in the genomes of non-cultured microorganisms. *Curr Opi Biotechnol,* 15: 285-290.

Suenaga H, Ohnuki T & Miyazaki K (2007) Functional screening of a metagenomic library for genes involved in microbial degradation of aromatic compounds. *Environ Microbiol*, 9: 2289-2297.

Svendsen A & Bisgaard-Frantzen H (1994) PCT patent publication. WO 94/0.

Swift HJ, Brady L, Derewenda ZS, Dodsen EJ, Dodsen GG, Turkenburg JP & Wiimson AJ (1991) Structure and molecular model refinement of *Aspergilius oryzae* (TAKA) α- amylase: an application of the simulated-anneling method. *Acta Crystallogr*, Sec B. 47: 535-544.

Tamura K, Dudley J, Nei M & Kumar S (2007) *MEGA4*: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol,* 24: 1596-1599.

Tang K, Kobayashi RS, Champreda V, Eurwilaichitr L & Tanapongpipat S (2008) Isolation and characterization of a novel thermostable neopullulanase-like enzyme from a hot spring in Thailand. *Biosci Biotechnol Biochem*, 72: 1448-1456.

Tubb RS (1986) Amylolitic yeast for commercial applications. *Trends Biotechnol*, 4: 98-104.

Tyson W, Lo I, Baker BJ, Allen EE, Hugenholtz P & Banfield JF (2005) Genome-directed isolation of the key nitrogen fixer *Leptospirillum ferrodiazotrophum* sp. nov. from an acidophilic microbial community. *Appl Environ Microbial*, 71(10): 6319–6324.

Uchiyama T & Miyazaki K (2009) Functional metagenomics for enzyme discovery: challenges to efficient screening. *Curr Opin Biotechnol*, 20:616-622.

Uchiyama T, Abe T, Ikemura T & Watanabe K (2005) Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes. *Nat Biotechnol*, 23:88-93.

Valvano MA, Keith KE & Cardona ST (2005) Survival and persistence of opportunistic *Burkholderia* species in host cells. *Curr Opin Microbiol,* 8: 99–105.

Van Dam HW & Hille JDR (1992) Yeast and enzymes in bread making. *Cereal Foods World*, St. Paul, v.37, n.3, p. 245-252.

Van der Maarel MJEC, Van der Veen B, Uitdehaag JCM, Leemhuis H & Dijkhuizen L (2002) Properties and aplications of starch-converting enzymes of the α-amylase family. *J Biotechnol,* 94:137-155.

Van Ee JH, van Rijswijk WC & Bollier M (1992) Enzymatic automated dishwash detergents. *Chim Oggi*, 10: 21-24.

Van Elsas JD, Garbeva P & Salles J (2002) Effects of agronomical measures on the microbial diversity of soils as related to the suppression of soil-borne plant pathogens. *Biodegradation* 13: 29–40.

Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers Y & Smith HO (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, 304: 66–74.

Voget S, Leggewie C, Uesbeck A, Raasch C, Jaeger KE & Streit WR (2003) Prospecting for novel biocatalysts in a soil metagenome. *Appl Environ Microbiol,* 69: 6235-6242.

Wang C, Meek DJ, Panchal P, Boruvka N, Archibald FS, Driscoll BT & Charles TC (2006) Isolation of poly-3-hydroxybutyrate metabolism genes from complex microbial communities by phenotypic complementation of bacterial mutants. *Appl Environ Microbiol*, 72:384-391

Ward D, Weller MR & Bateson MM (1990) 16s Rrna Sequences Reveal Numerous Uncultured Microorganisms In A Natural Community. *Nature,* 345: 63-65.

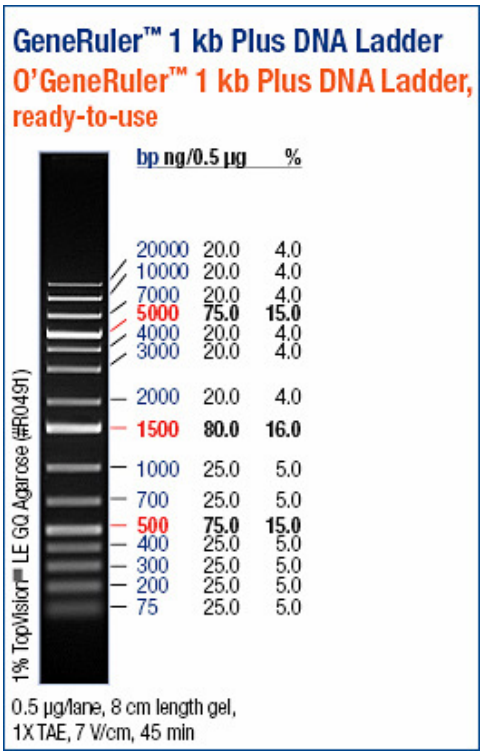
Warren RL, Freeman JD, Levesque RC, Smailus DE, Flibotte S & Holt RA (2008) Transcription of foreign DNA in *Escherichia coli*. *Genome Res*, 18: 1798-1805.

Williamson LL, Borlee BR, Schloss PD, Guan C, Allen HK & Handelsman J (2005) Intracellular screen to identify metagenomic clones that induce or inhibit a quorum-sensing biosensor. *Appl Environl Microbiol,* 71: 6335-6344.

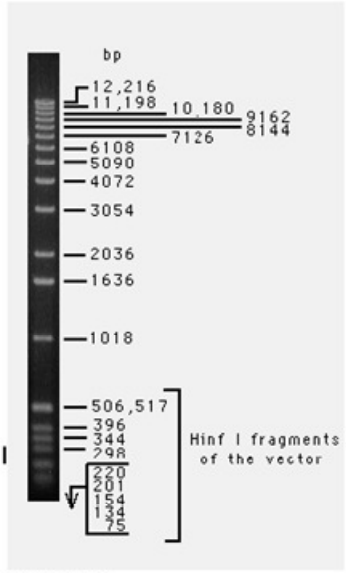
Xu M, Xiao X & Wang F (2008) Isolation and characterization of alkane hydroxylases from a metagenomic library of Pacific deep-sea sediment. *Extremophiles*, 12:255-262.

Yun J, Kang S, Park S, Yoon H, Kim MJ, Heu S & Ryu S (2004) Characterization of a Novel Amylolytic Enzyme Encoded by a Gene from a Soil-Derived Metagenomic Library. *Appl Environ Microbiol,* 70:7229-35.

Anexo 1. Marcadores de DNA utilizados: O'Gene Ruler 1 kb Plus DNA Ladder (Fermentas) e 1 kb Ladder (Gibco BRL).

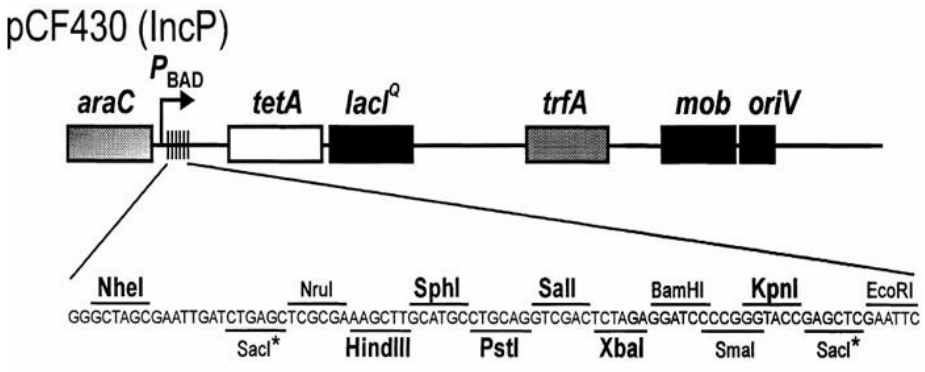


1 kb Ladder (Gibco)

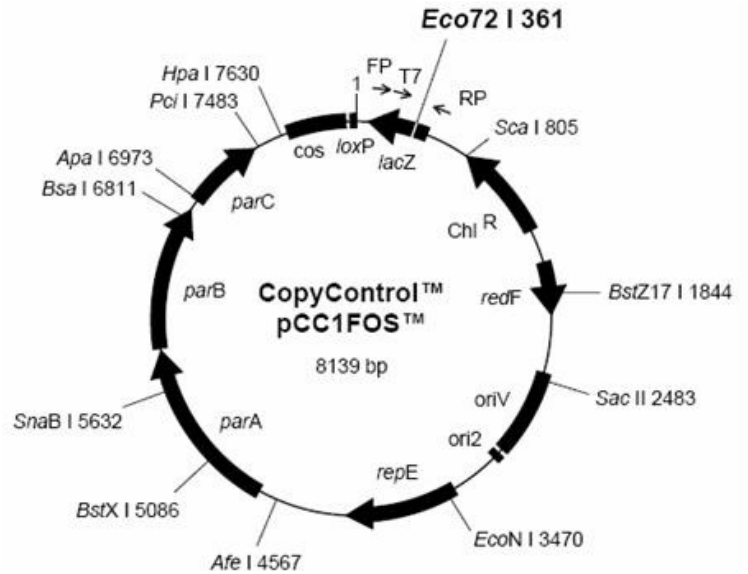


1 Kb DNA Ladder
0.5 µg/lane
0.9% agarose gel
stained with ethidium bromide

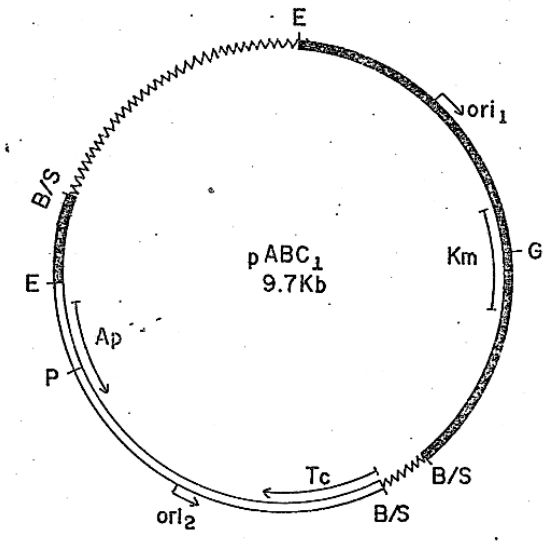
Anexo 2. Representação esquemática do mapa físico do vetor pCF430 e seus sítios múltiplo de clonagem. Sítios de restrição em fonte maior são únicos e sítios em letras menores ocorrem mais de uma vez no vetor (Extraído de Newman & Fuqua, 1999).



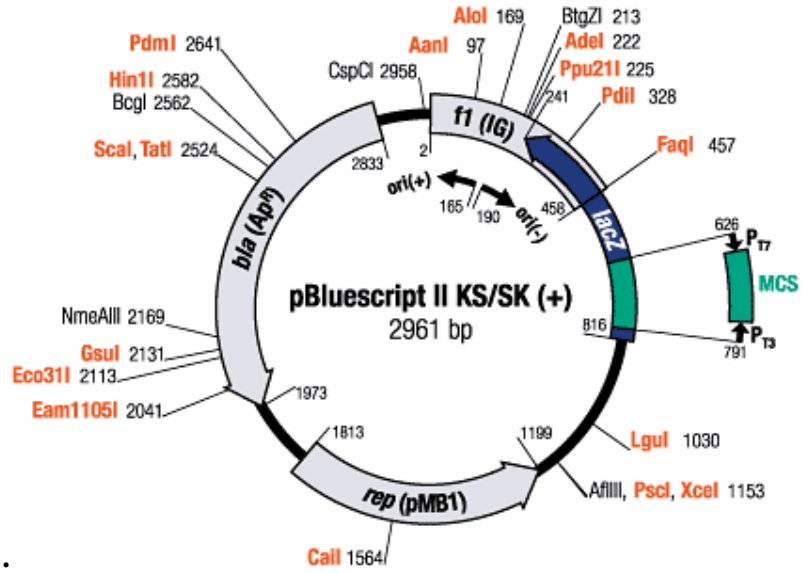
Anexo 3. Representação esquemática do mapa físico do vetor do fosmídio CopyControl™ pCCFOS™ (Extraído de www.epibio.com - acessado em 09/05/10).



Anexo 4. Representação esquemática do mapa físico do vetor pABC1. E - *Eco*RI; P - *Pst*I; G - *Bgl*II; B/S - *Bam*HI/*Sau*3A. (/\/\/\/\) - Fragmento clonado de *Bacillus* sp. com atividade amilolítica (Extraído de Souza,1986).



Anexo 5. Representação física dos vetores pBluescript II KS/SK + (Extraído de http://fermentas.com/- acessado em 09/05/10).



Anexo 6. Sequência dos p*rimers* utilizados em reações PCR com suas respectivas temperaturas de anelamento e sítios de restrição.

Primer	Sequência (5'→ 3')	Tm (°C)	Sítio
pCF430F	CTGTTTCTCCATACCCGTT	53	-
pCF430R	TGCAAGGCGATTAAGTTGG	58	-
A1F31	ACGCTGTCCTCGGGACTCTG	66	-
A1F32	*GCCGCGACGCTGCGCTACAT*	68	-
A1F2R1	ATGTCAAGTTACGTTCGGTT	60	-
A1F2R2	GATGTTACGCAGCTTCTAGC	57	-
A1R31	AAGAGGTGGTGTCACACCGA	62	-
A1R32	GTGAACGTTGTCGTTGAAGG	60	-
A1CLOF	CATATGAGTGGCGAGCCCTGGT	61	*Nde*I
A1CLOR1	GAGCTCTTAAGCTGATGCGAGACCGATC	57	*Sal*I
A1CLOR2	GAGCTC TTATGCCATAGCATTTTTATCCATAA	60	*Sal*I

Anexo 7. Sequência nucleotídica e de aminoácidos da proteína predita 1 do plasmídeo pA1. Em vermelho está indicado o sítio de restrição da enzima *Pst*I.

```
  1   M  S  G  E  P  W  W  R  H  A  I  L  Y  Q  L  D  P  K  S  F
  1  ATGAGTGGCGAGCCCTGGTGGAGACACGCAATCCTCTACCAACTCGACCCGAAGAGCTTC

 21   S  P  E  G  L  R  G  V  S  R  R  L  D  Y  I  Q  S  L  G  A
 61  TCCCCTGAGGGCCTGCGCGGCGTCTCCAGGCGCCTCGACTACATCCAGTCTCTCGGCGCC

 41   D  A  I  L  L  D  G  I  T  S  G  T  I  D  S  S  I  G  T  L
121  GACGCCATCCTCCTCGACGGAATCACGTCCGGCACCATCGACTCGTCTATCGGCACGTTG

 61   D  D  L  D  D  L  I  H  Q  A  S  S  R  N  I  R  I  L  I  G
181  GATGATCTCGACGACCTCATCCACCAGGCCAGCAGCCGAAACATCCGAATCCTCATCGGC

 81   L  D  P  K  A  S  D  L  T  S  V  A  R  L  W  L  T  H  G  V
241  CTCGATCCCAAAGCATCCGATCTCACCTCAGTCGCCCGCCTCTGGCTCACTCACGGCGTC

101   A  G  F  R  A  P  G  A  T  P  E  Q  F  A  E  L  R  K  L  L
301  GCAGGATTCCGCGCGCCAGGCGCTACTCCAGAACAGTTCGCCGAGCTCAGAAAGCTTCTC

121   N  T  F  V  G  Q  R  I  L  V  D  D  L  D  P  S  N  A  Q  Q
361  AACACCTTCGTCGGTCAGCGCATACTTGTTGACGATCTCGACCCGTCCAACGCGCAGCAG

141   P  N  H  G  P  D  V  T  Q  L  L  A  D  P  R  P  G  T  V  P
421  CCGAACCACGGGCCCGATGTTACGCAGCTTCTAGCCGACCCTCGCCCGGGCACCGTGCCC

161   Q  L  T  A  A  T  V  R  P  A  L  E  A  S  Q  A  L  I  E  R
481  CAACTCACCGCCGCCACCGTGCGCCCCGCGCTTGAGGCATCGCAGGCGTTAATCGAGCGC

181   G  R  V  T  T  V  L  F  S  D  G  P  D  R  P  R  S  F  T  R
541  GGCCGCGTCACCACTGTACTCTTTAGCGACGGTCCCGACCGCCCGCGTAGCTTCACACGC

201   Y  G  D  G  S  H  D  I  A  I  A  K  T  L  A  A  I  L  L  A
601  TACGGCGACGGATCTCACGACATCGCCATCGCGAAGACGCTTGCCGCCATCCTACTCGCA

221   T  R  S  G  S  L  L  Y  Y  G  Q  E  I  G  L  A  S  A  P  A
661  ACACGCTCCGGATCGCTCCTCTATTACGGCCAGGAGATCGGTCTCGCATCAGCTCCTGCA

241   G  M  Q  A  F  A  S  S  T  L  A  Q  K  N  G  Y  G  E  T  V
721  GGCATGCAAGCTTTCGCGAGCTCGACGCTAGCCCAAAAAAACGGGTATGGAGAAACAGTA

261   E  S  C  D  K  K  R  Q  V  G  S  A  N  L  M  D  K  N  A  M
781  GAGAGTTGCGATAAAAAGCGTCAGGTAGGATCCGCTAATCTTATGGATAAAAATGCTATG

281   A  -
841  GCATAG
```

Anexo 8. Alinhamento múltiplo da sequência de aminoácidos da proteína predita 1 de pA1 com sequências similares. As assinaturas da Taka-amilase A de *Aspergillus oryzae* estão em destaque amarelo estão. Em letras vermelhas estão as possíveis assinaturas da proteína predita 1.

```
Proteina-predita-1-pA1              ---------------------------------------MSGEPWWRHAI 11
Eubacterium-trealose-6-fosfato      ------------------------------------MSKNNNEKWWKKAV 14
Ruminococcus-trealose-6-fosfat      -----------------------------------MEGTRKEHWWRDAV 14
Lactobacillus-trealose-6-fosfa      ------------------------------------------MTPWWKKAV 9
Ecoli-trealose-6-fosfato-hidro      --------------------------------------MTNLPHWWQNGV 12
Shigella-trealose-6-fosfato-hi      --------------------------------------MTNLPHWWQNGV 12
Citrobacter-trealose-6-fosfato      --------------------------------------MNTLPHWWQNGV 12
Salmonella-trealose-6-fosfato-      -------------------------------------MTIPLWWQNGV 11
Yersinia-trealose-6-fosfato-hi      ------------------------------------MNNPIPWWQNGV 12
Mobiluncos-trealose-6-fosfato-      ----------------------MTDSHRNYLANGLQANGISDNPWWINAV 28
Hirschia-amilase                    ------------------------MNIVTPSEIELGHKSSDKAPWWKGAV 26
Acidobacterium                      MQILNQRTILLCSVFLLCVPPAIPVAAQTLARPGWAGSGLTQETWWQNAV 50
Aspergillus-Taka-amylase-A          ----------------------------------------ATPADWRSQS 10
                                                                                *

Proteina-predita-1-pA1              LYQLDPKSFSPE------------------GLRGVSRRLDYIQSLGADAI 43
Eubacterium-trealose-6-fosfato      VYQIYPKSFQDSNGDGF------------GDLQGIIKRLDYLETLGINAI 52
Ruminococcus-trealose-6-fosfat      IYQIYPRSFQDSNGDGI------------GDLRGIIQRLDYLKELGIDAI 52
Lactobacillus-trealose-6-fosfa      IYQIYPKSFQDSNGDGI------------GDLPGIITRLDYLAKLGIDAI 47
Ecoli-trealose-6-fosfato-hidro      IYQIYPKSFQDTTGSGT------------GDLRGVIQRLDYLHKLGVDAI 50
Shigella-trealose-6-fosfato-hi      IYQIYPKSFQDTTGSGT------------GDLRGVIQRLDYLHKLGVDAI 50
Citrobacter-trealose-6-fosfato      IYQIYPKSFQDTTGSGT------------GDLRGVTQRLDYLQKLGVDAI 50
Salmonella-trealose-6-fosfato-      IYQIYPKSFQDTTGSGT------------GDLRGVTQRLDYLQRLGVDAI 49
Yersinia-trealose-6-fosfato-hi      IYQIYPKSFQDSTGNGY------------GDLAGVTQRLDYIRQLGVDAI 50
Mobiluncos-trealose-6-fosfato-      VYQIYPRSFQDSNGDGV------------GDIKGIISKLDYIANLGADVI 66
Hirschia-amilase                    IYHIYPRSFYDSNGDGI------------GDLRGCIEKLDHIASLGVDAI 64
Acidobacterium                      FYRVQPKNFQDSDGDGT------------GDLRGVAQRLDYLQTLGVDAI 88
Aspergillus-Taka-amylase-A          IYFLLTDRFARTDGSTTATCNTADQKYCGGTWQGIIDKLDYIQGMGFTAI 60
                                    .* : .  *                       *   :**::  :*  .*

Proteina-predita-1-pA1              LLDGITSG--------------------TIDSSIGTLDDLDDLIHQASSR 73
Eubacterium-trealose-6-fosfato      WLSPVFKSPQADNGY-------DISDYRDIDPTFGSLDDMEELINEAKKH 95
Ruminococcus-trealose-6-fosfat      WLSPVCKSPQDDNGY-------DISDYQDIDPMFGSLDDMEELIKEAKKR 95
Lactobacillus-trealose-6-fosfa      WLSPVYQSPGIDNGY-------DISDYQAIDPQYGTMDDMDNLILEAKKR 90
Ecoli-trealose-6-fosfato-hidro      WLTPFYVSPQVDNGY-------DVANYTAIDPTYGTLDDFDELVTQAKSR 93
Shigella-trealose-6-fosfato-hi      WLTPFYVSPQVDNGY-------DVANYTAIDPTYGTLDDFDELVTQAKSR 93
Citrobacter-trealose-6-fosfato      WLTPFYVSPQIDNGY-------DVANYTAIDPAYGTLDDFDELAAQAKAR 93
Salmonella-trealose-6-fosfato-      WLTPFYISPQVDNGY-------DVANYTAIDPTYGTLDDFDELVAQAKAR 92
Yersinia-trealose-6-fosfato-hi      WLTPIYLSPQVDNGY-------DVADYCAIDPAYGTMDDFKAFIHAAHQR 93
Mobiluncos-trealose-6-fosfato-      WLNPVYKSPQDDNGY-------DISDYQDIDPLFGTLDDLDELIYQSHRR 109
Hirschia-amilase                    WLSPFFKSPQADYGY-------DISDYKEIDPTFGTLDDFDDLVEQAHSR 107
Acidobacterium                      VLESPFDDAG----------------------------FDDLLSEATSR 109
Aspergillus-Taka-amylase-A          WITPVTAQLPQDCAYGDAYTGYWQTDIYSLNENYGTADDLKALSSALHER 110
                                     :                                      :. :      :

Proteina-predita-1-pA1              NIRILIGLDPKAS------------------------------------ 86
Eubacterium-trealose-6-fosfato      NIRIMMDLVLNHSSNEHRWFKEAKKSKDNPYHDYYIWR-------DGDEG 138
Ruminococcus-trealose-6-fosfat      NIRIIMDLVLNHSSDEHRWFQEAKKSKDNPYHDYYVWR-------DGKEG 138
Lactobacillus-trealose-6-fosfa      NIKIIMDLVVNHTSDQHAWFVEARKNKENQYRDYYIWR-------DPVDG 133
Ecoli-trealose-6-fosfato-hidro      GIRIILDMVFNHTSTQHVWFREALN-KESPYRQFYIWR-------DGEPE 135
Shigella-trealose-6-fosfato-hi      GIRIILDMVFNHTSTQHAWFREALN-KESPYRQFYIWR-------DGEPE 135
Citrobacter-trealose-6-fosfato      GIRIILDMVFNHTSTQHAWFHNALN-KESPHREFYIWR-------DGTPD 135
Salmonella-trealose-6-fosfato-      GIRIILDMVFNHTSTQHAWFREALN-KESPYRQFYIWR-------DGTPD 134
Yersinia-trealose-6-fosfato-hi      GIRIVMDMVFNHTSTEHAWFQAAKD-PASPYRQFYIWR-------DGEGE 135
Mobiluncos-trealose-6-fosfato-      GIRIIMDLVVNHTSDEHSWFQASRN-PQDPRADWYWWRPARSGYVPGEPG 158
Hirschia-amilase                    GLKIIIDQVYAHTSDQHAWFEESRQSRKNVKSDWYVWAD------PKIDG 151
Acidobacterium                      HVRIVVELAPHHEGGAQQALAEARS------------------------- 134
Aspergillus-Taka-amylase-A          GMYLMVDVVANHMGYDGAGSSVDYSVFKPFSSQDYFHP------------ 148
                                     : :::

Proteina-predita-1-pA1              ---------------------------------------------DLTSV 91
Eubacterium-trealose-6-fosfato      VPPSDMKACFGGSAWEYVPEIGQYYFHQFLPEQPDLNWENPKVRRAIYDM 188
Ruminococcus-trealose-6-fosfat      VYPNDMRSVFGGPAWEWVPELEQYYFHQFSVKQPDLNWENPKVRREIYDM 188
Lactobacillus-trealose-6-fosfa      HEPNGLQSAFSGSAWQYDETTKQYYLHFFAKQQPDLNWSNSIVRDKIYQM 183
Ecoli-trealose-6-fosfato-hidro      TPPNNWRSKFGGSAWRWHAESEQYYLHLFAPEQADLNWENPAVRAELKKV 185
Shigella-trealose-6-fosfato-hi      TPPNNWRSKFGGSAWRWHAESEQYYLHLFAPEQADLNWENPAVRAELKKV 185
Citrobacter-trealose-6-fosfato      APPNNWRSKFGGSAWRWHAQSEQYYLHLFAPEQADLNWENPAVRAELKKV 185
Salmonella-trealose-6-fosfato-      VCPNNWQSKFGGSAWRWHSQSEQYYLHLFAPEQADLNWENPAVRAELKKV 184
Yersinia-trealose-6-fosfato-hi      QLPNNWRSKFGGNAWQWHAESGQYYLHLFAPEQADLNWEHPPVRDELKKV 185
```

```
Mobiluncos-trealose-6-fosfato-      AEPNKWGSYFGGSAWTYDTLRKEYFLHQYSAKQPDLNWNNPEVRRAIYSM 208
Hirschia-amilase                    MPPNNWQSVFGGPAWEWDARRQQYFMHNFLKEQPQLNVHNLDVQNAILEI 201
Acidobacterium                      --------------------------------------------------
Aspergillus-Taka-amylase-A          --------FCFIQNYEDQTQVEDCWLGDNTVSLPDLDTTKDVVKNEWYDW 190

Proteina-predita-1-pA1              ARLWLTHG-VAGFRAPGATP------------------------------ 110
Eubacterium-trealose-6-fosfato      ILWWMDKG-VGGFRLDVIDQIAKEP---------DKRIT----------- 217
Ruminococcus-trealose-6-fosfat      ILWWMEKG-AGGFRLDVIDQIAKEP---------DQKIT----------- 217
Lactobacillus-trealose-6-fosfa      MNYWIDKG-IGGFRMDVIELIGKDP---------DKMVR----------- 212
Ecoli-trealose-6-fosfato-hidro      CEFWADRG-VDGLRLDVVNLISKDPRFPDDLDGDGRRFY----------- 223
Shigella-trealose-6-fosfato-hi      CEFWADRG-VDGLRLDVVNLISKDPRFPEDLDGDGRRFY----------- 223
Citrobacter-trealose-6-fosfato      CEFWADRG-VDGLRLDVVNLISKDQRFPNDLDGDGRRFY----------- 223
Salmonella-trealose-6-fosfato-      CEFWADRG-VDGLRLDVVN------------------------------- 202
Yersinia-trealose-6-fosfato-hi      CEFWADLG-VDGLRLDVINLVSKQQDFPSDPTGDGRRFY----------- 223
Mobiluncos-trealose-6-fosfato-      MNWWMDRG-INGFRMDVITQVSKITDESGHLPGEEGSAIPDLPAGDDGYS 257
Hirschia-amilase                    MRFWIDRG-VDGFRMDAINFAMHNRAMTDNPPSNLERSR----------- 239
Acidobacterium                      ---WLTRG-AAGLYAKSSLFAGPAG------------------------- 155
Aspergillus-Taka-amylase-A          VGSLVSNYSIDGLRIDTVKH------------------------------ 210
                                              *:

Proteina-predita-1-pA1              ------EQFAELRKLLNTFVGQRILVDDLDPSNAQQPNHGPDVTQLLADP 154
Eubacterium-trealose-6-fosfato      ------INGPRLQEYFKELSRETFQKGD-LITVGEAWGADTERAKLYSNP 260
Ruminococcus-trealose-6-fosfat      ------NNGPRLHEFIQELSRETFQKGD-LITVGEAWGADIERAKLYSNP 260
Lactobacillus-trealose-6-fosfa      ------ENGKMLHPYLQEMNQKTFGQSD-LLTVGETWNATPKIAEDYSDP 255
Ecoli-trealose-6-fosfato-hidro      ------TDGPRAHEFLHEMNRDVFTPRG-LMTVGEMSSTSLEHCQRYAAL 266
Shigella-trealose-6-fosfato-hi      ------TDGPRAHEFLHEMNRDVFTPRG-LMTVGEMSSTSLEHCQRYAAL 266
Citrobacter-trealose-6-fosfato      ------TDGPRAHEFLHEMNRDVFTPRN-LMTVGEMSSTTLENCQRYAAL 266
Salmonella-trealose-6-fosfato-      --------------------------------------------------
Yersinia-trealose-6-fosfato-hi      ------TDGPRIHEFLQEMSREVFQPRG-LMTVGEMSSTHLEHCQRYAAL 266
Mobiluncos-trealose-6-fosfato-      SPFPFCSDGPRLDEFLHEMRCEVFANRDGYLVVGEAPGISASRNQFITDP 307
Hirschia-amilase                    -----VTRPFDMQKRIHNQSQPEIPLFMERIRKELDKHGEIFTVAEIGGD 284
Acidobacterium                      ----TADPLQPLRQLVNGFPGDRILLTDNDPAAATARHGQRGASHELVDF 201
Aspergillus-Taka-amylase-A          ---------------VQKDFWPGYNKAAGVYCIGEVLDGDPAYTCPYQNV 245

Proteina-predita-1-pA1              RPG----------------------TVPQLTAATVRPALEASQALIERGR 182
Eubacterium-trealose-6-fosfato      DGSEFSMVFQFEHIGLDQKEGGEKWDLAPLPFKKLKKIMAHWQNEL-YNC 309
Ruminococcus-trealose-6-fosfat      DGSEFSMVFQFEHIGLDQQEGKSKWDVAPLPVVKLKQVLAKWQREL-HGC 309
Lactobacillus-trealose-6-fosfa      ARHELSMVFQFENQTLDQKDGCSKWDLKPIDLVELKKILVKWQTELDFKH 305
Ecoli-trealose-6-fosfato-hidro      TGSELSMTFNFHHLKVDYPGG-EKWTLAKPDFVALKTLFRHWQQGM-HNV 314
Shigella-trealose-6-fosfato-hi      TGSELSMTFNFHHLKVDYPGG-EKWTLAKPDFVALKTLFRHWQQGM-HNV 314
Citrobacter-trealose-6-fosfato      DGSELSMTFNFHHLKVDYPGG-EKWTLAKPDYVALKALFRHWQQGM-HNV 314
Salmonella-trealose-6-fosfato-      --------------------------------------------------
Yersinia-trealose-6-fosfato-hi      AGNELSMTFNFHHLKVDYLNG-EKWSLAEPNLVELKEIFRTWQQGM-HNR 314
Mobiluncos-trealose-6-fosfato-      KHNELDMLFLFDHVEVDCSNG-TKWNPKPLYLPDLKKCMTKQQEAV-SCR 355
Hirschia-amilase                    EALSEMRDYTYGDKRLNSAYSFDFLDGPPVSADKVRGTLGNWIGED--GE 332
Acidobacterium                      NLGKGAATLENYRSELAAVQA-------SPDIPLVLQLDEQTPPEASLAK 244
Aspergillus-Taka-amylase-A          MDGVLNYPIYYPLLN--------AFKSTSGSMDDLYNMINTVKSDC-PDS 286

Proteina-predita-1-pA1              VTTVLFSDGPDRPRSFTRYGDGSHDI---AIAKTLAAILLATRSGSLLYY 229
Eubacterium-trealose-6-fosfato      GWNSLFWDNHDLPRIVSRWGNDREYR--VESAKMLAILLHGMQGTPYIYQ 357
Ruminococcus-trealose-6-fosfat      GWNSLFWNNHDLPRIVSRWGNDQKYR--VESAKMLATLLHGMQGTPYIYQ 357
Lactobacillus-trealose-6-fosfa      AWNSLFWENHDIPRVISRWGNDQKYR--VVCAKMFAIVLHLMHGTPYIFN 353
Ecoli-trealose-6-fosfato-hidro      AWNALFWCNHDQPRIVSRFGDEGEYR--VPAAKMLAMVLHGMQGTPYIYQ 362
Shigella-trealose-6-fosfato-hi      AWNALFWCNHDQPRIVSRFGDEGEYR--VPAAKMLAMVLHGMQGTPYIYQ 362
Citrobacter-trealose-6-fosfato      AWNALFWCNHDQPRIVSRFGDEGELR--VPAAKMLAMVLHGMQGTPYIYQ 362
Salmonella-trealose-6-fosfato-      --------------------------------------------------
Yersinia-trealose-6-fosfato-hi      AWNALFWCNHDQPRIVSRFGDEGSLR--VPAAKMLGMVLQGMQGTPYIYQ 362
Mobiluncos-trealose-6-fosfato-      GWTSLYFNNHDQPRIVSRWGDDKSVDSVAKSAKALALLLHMHRGTPYIFQ 405
Hirschia-amilase                    GWPSWAFENHDAPRAVSRWAKPENRW---FAARNYMMLLLSLRGNAFIYQ 379
Acidobacterium                      IAATMLFASHG-PASLHDGQEIGIER---------ASTMQWTPGNITLPK 284
Aspergillus-Taka-amylase-A          TLLGTFVENHDNPRFASYTND-------IALAKNVAAFIILNDGLPIIYA 329

Proteina-predita-1-pA1              GQE---IGLASAPAGMQAFASSTLAQ------------------------ 252
Eubacterium-trealose-6-fosfato      GEE---LGMTNVQY-DIEDYKDCEIINMYHERLEKG--YSKDEIMKSIYA 401
Ruminococcus-trealose-6-fosfat      GEE---LGMTNVRFADISQYQDIETLNMYKERLDEG--YSKEEIMNSIYA 402
Lactobacillus-trealose-6-fosfa      GEE---IGMTNSPVKNINEVEDIESINMYNERLDAG--YDKDELLHAINV 398
Ecoli-trealose-6-fosfato-hidro      GEE---IGMTNPHFTRITDYRDVESLNMFAELRNDG--RDADELLAILAS 407
Shigella-trealose-6-fosfato-hi      GEE---IGMTNPHFTRITDYRDVESLNMFAELRNDG--RDADELLAILAS 407
Citrobacter-trealose-6-fosfato      GEE---IGMTNPHFSHITDYRDVESHNMFAGLRAQG--REANDLLAILAS 407
Salmonella-trealose-6-fosfato-      --------------------------------------------------
Yersinia-trealose-6-fosfato-hi      GEE---IGMTNPSFSTIEQYRDLESLNMFAELSASG--CQPDELLAILAS 407
Mobiluncos-trealose-6-fosfato-      GEE---LGMTNAHFNSLSQYRDLEALNMYHQRVEQAGIQSPQSMLEALGR 452
Hirschia-amilase                    GEE---LGLP-QGEVAFEHLKDPEAIANWPR------------------T 407
```

```
Acidobacterium                      PTE---PVAEAPKPTPPPQREDVYTAFVPYVPPPTK------------PK 319
Aspergillus-Taka-amylase-A          GQEQHYAGGNDPANREATWLSGYPTDSELYKLIASAN-----AIRNYAIS 374

Proteina-predita-1-pA1              -----------------KNGYGET------VESCDKKRQVGSANLMDKNA 279
Eubacterium-trealose-6-fosfato      KGRDNARTPMQWDD-SANAGFTTG---TPWIKVNDNYDKINAKSQVNDPD 447
Ruminococcus-trealose-6-fosfat      KGRDNARTPMQWSG-EVNAGFTKG---TPWLEVNPNYTKINAESELSDPD 448
Lactobacillus-trealose-6-fosfa      KGRDNARTPMQWSD-ALNAGFTSG---KPWLRVNSNYTDINVKQCLADPD 444
Ecoli-trealose-6-fosfato-hidro      KSRDNSRTPMQWSN-GDNAGFTAG---EPWIGLGDNYQQINVEAALADDS 453
Shigella-trealose-6-fosfato-hi      KSRDNSRTPMQWSN-GDNAGFTAG---EPWIGLGDNYQKINVEAALADDS 453
Citrobacter-trealose-6-fosfato      KSRDNSRTPMQWDS-SANAGFTQG---EPWINLCDNYADINVAAALNDED 453
Salmonella-trealose-6-fosfato-      --------------------------------------------------
Yersinia-trealose-6-fosfato-hi      KSRDNGRTPMQWDS-SANAGFTVG---TPWISPCTNYPEVNVSAALADKN 453
Mobiluncos-trealose-6-fosfato-      RARDNSRTPMQWDA-SPYAGFTWRGATEPWISVNSNYLEINASSQINDDD 501
Hirschia-amilase                    LGRDGARVPMPWNKNADHAGFTTG---TPWIPVDPKQAELAVDTQTDDIP 454
Acidobacterium                      PRPIDPRTGEPSADPDSLMGFTSG--KLPAEASVPGHESLNVAIEDLNPD 367
Aspergillus-Taka-amylase-A          KDTGFVTYKNPYIKDDTTIAMRKGTDGSQIVTILSNKGASGDSYTLSLSG 424

Proteina-predita-1-pA1              MA------------------------------------------------ 281
Eubacterium-trealose-6-fosfato      SIFSCYKKLVQLRKDYPVFVDGKFTLLLEDDENIFAYSRKN----EEKTM 493
Ruminococcus-trealose-6-fosfat      SVFYYYQKLIRLRKEHSVFVNGEFTLLMEEDPQVFAYIRKE----GDTEV 494
Lactobacillus-trealose-6-fosfa      SIFYTYQKLIELRHHNQIVVDGDFQLIEETDAQILAYYRIL----GDQKW 490
Ecoli-trealose-6-fosfato-hidro      SVFYTYQKLIALRKQEAILTWGNYQDLLPNSPVLWCYRREW----KGQTL 499
Shigella-trealose-6-fosfato-hi      SVFYTYQKLIALRKQEAILTWGNYQDLLPNSPVLWCYRREW----KGQTL 499
Citrobacter-trealose-6-fosfato      SVFYTYQRLIALRKTEPVLTWGNYQDLLPDSPHVWCYRREW----QGQRL 499
Salmonella-trealose-6-fosfato-      --------------------------------------------------
Yersinia-trealose-6-fosfato-hi      SVFYAYQRLISLRKQHDIFTYGDYQDLCPEHPSLWCYLRSW----RNQRL 499
Mobiluncos-trealose-6-fosfato-      SVYNFYRRLIHLRHAQPIVATGDWNLLDASDLNVYSFTRSL----NNEKL 547
Hirschia-amilase                    CMLRFTKYAMGVRKRTLALRLGSMTFMDTHNSSVLAFMRRH----DEEHI 500
Acidobacterium                      SLLNFYRRLSQLHHGSPALQTGSLSVLDHNADNALVWVRQAPGVRETGTV 417
Aspergillus-Taka-amylase-A          ASYTAGQQLTEVIGCTTVTVGSDGNVPVPMAGGLPRVLYPT------EKL 468

Proteina-predita-1-pA1              --------------------------------------------------
Eubacterium-trealose-6-fosfato      IVVCNFFDKE------------IPMP-LAKECEGM--------------E 516
Ruminococcus-trealose-6-fosfat      LVCANFSETP------------AACD-LLSEWKDG--------------E 517
Lactobacillus-trealose-6-fosfa      LVVANLTDKE------------KPFN-CHEQYQ----------------K 511
Ecoli-trealose-6-fosfato-hidro      LVIANLSREI------------QPWQPGQMR-GNW--------------Q 522
Shigella-trealose-6-fosfato-hi      LVIANLSREI------------QPWQPGQMR-GNW--------------Q 522
Citrobacter-trealose-6-fosfato      LVVANLSNAC------------QAWLPEHHE-GNW--------------Q 522
Salmonella-trealose-6-fosfato-      --------------------------------------------------
Yersinia-trealose-6-fosfato-hi      LVVANLSDKP------------VQWQSGDLERGNW--------------S 523
Mobiluncos-trealose-6-fosfato-      LVIVNLSKETSLIPKETITLLRNAWLSKKSTCSNLDCSDMANSMPLLSSK 597
Hirschia-amilase                    MCVYNFSDKH------------SKWNHHLGK---G-------------WK 522
Acidobacterium                      IIACNLSSKTTTLS------LQNDIERLHLRPGG--------------LR 447
Aspergillus-Taka-amylase-A          AGSKICSDSS------------------------------------- 478

Proteina-predita-1-pA1              ---------------------------------
Eubacterium-trealose-6-fosfato      VLISNYKDTSDMS-----VLRPYEARMYIR--- 541
Ruminococcus-trealose-6-fosfat      VLIWNYKEKGETG-----VLRPYEAMMIRR--- 542
Lactobacillus-trealose-6-fosfa      VLVENYTTPSNLQN-I--ILKPYQAFVVTIA-- 539
Ecoli-trealose-6-fosfato-hidro      LVMHNYEEASPQPCAM--NLRPFEAVWWLQK-- 551
Shigella-trealose-6-fosfato-hi      LVMHNYEEASPQPCAM--NLRPFEAVWWLQK-- 551
Citrobacter-trealose-6-fosfato      ALMHNYDEVATRPAAM--TLRPFEAVWWLQK-- 551
Salmonella-trealose-6-fosfato-      ---------------------------------
Yersinia-trealose-6-fosfato-hi      LLMSSYDDSQSQPESL--MLRPYEGIYWMKD-- 552
Mobiluncos-trealose-6-fosfato-      ILISTYNKSHAVHSLVRECLSPWEGIAIML--- 627
Hirschia-amilase                    VMFSHDGMSDGWEVPH--DLPAYSAYWAKLTK- 552
Acidobacterium                      MLLVSDKANMRFQTTQQLTLAPWSVFIGELHSY 480
Aspergillus-Taka-amylase-A    -     ----------------------------
```

Anexo 9. Alinhamento local entre a proteína predita 1 com a sequência do domínio catalítico de amilase *Acidobacterium* sp.

```
prot-pred1-pA      1 MSGEPWWRHAILYQLDPKSFSPE------GLRGVSRRLDYIQSLGADAIL     44
                     ::.|.||::|:.|::.||:|...      .||||::||||:|:||.|||:
amy-acidobact     40 ltqetwwqnavfyrvqpknfqdsdgdgtgdlrgvaqrldylqtlgvdaiv     89

prot-pred1-pA     45 LDGITSGTIDSSIGTLDD--LDDLIHQASSRNIRILIGLDPK----ASDL     88
                     |:           ...||  .|||:.:|:||::||::.|.|.    |...
amy-acidobact     90 le-----------spfddagfddllseatsrhvrivvelaphheggaqqa    128

prot-pred1-pA     89 TSVARLWLTHGVAG-------FRAPGATPEQFAELRKLLNTFVGQRILVD    131
                     .:.||.|||.|.||       |..|..|.:....||:|:|.|.|.|||:.
amy-acidobact    129 laearswltrgaaglyaksslfagpagtadplqplrqlvngfpgdrillt    178

prot-pred1-pA    132 DLDPSNAQQPNHGP-DVTQLLADPRPGTVPQLTAATV---RPALEASQAL    177
                     |.||: |....||. ..:..|.|...|.    .|||:   |..|.|.||
amy-acidobact    179 dndpa-aatarhgqrgashelvdfnlgk----gaatlenyrselaavqa-    222

prot-pred1-pA    178 IERGRVTTVLFSDGPDRPRSFTRYGDGSHDIAIAKTLAAILLATRSGSLL    227
                                 .||.|...........:.::||..|.:|.|:...:.|
amy-acidobact    223 -------------spdiplvlqldeqtppeaslakiaatmlfashgpasl    259

prot-pred1-pA    228 YYGQEIGLASA    238
                     :.|||||:..|
amy-acidobact    260 hdgqeigiera    270
```

Livros Grátis

(http://www.livrosgratis.com.br)

Milhares de Livros para Download:

Baixar livros de Administração
Baixar livros de Agronomia
Baixar livros de Arquitetura
Baixar livros de Artes
Baixar livros de Astronomia
Baixar livros de Biologia Geral
Baixar livros de Ciência da Computação
Baixar livros de Ciência da Informação
Baixar livros de Ciência Política
Baixar livros de Ciências da Saúde
Baixar livros de Comunicação
Baixar livros do Conselho Nacional de Educação - CNE
Baixar livros de Defesa civil
Baixar livros de Direito
Baixar livros de Direitos humanos
Baixar livros de Economia
Baixar livros de Economia Doméstica
Baixar livros de Educação
Baixar livros de Educação - Trânsito
Baixar livros de Educação Física
Baixar livros de Engenharia Aeroespacial
Baixar livros de Farmácia
Baixar livros de Filosofia
Baixar livros de Física
Baixar livros de Geociências
Baixar livros de Geografia
Baixar livros de História
Baixar livros de Línguas

Baixar livros de Literatura
Baixar livros de Literatura de Cordel
Baixar livros de Literatura Infantil
Baixar livros de Matemática
Baixar livros de Medicina
Baixar livros de Medicina Veterinária
Baixar livros de Meio Ambiente
Baixar livros de Meteorologia
Baixar Monografias e TCC
Baixar livros Multidisciplinar
Baixar livros de Música
Baixar livros de Psicologia
Baixar livros de Química
Baixar livros de Saúde Coletiva
Baixar livros de Serviço Social
Baixar livros de Sociologia
Baixar livros de Teologia
Baixar livros de Trabalho
Baixar livros de Turismo