

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL**

**PROSPECÇÃO E EXPRESSÃO DE UMA ENZIMA
PROTEOLÍTICA A PARTIR DE UMA BIBLIOTECA
METAGENÔMICA DO CONSÓRCIO MICROBIANO
DEGRADADOR DE ÓLEO DIESEL**

Bianca de Melo Silveira
Tecnóloga em Biocombustíveis

2015

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL**

**PROSPECÇÃO E EXPRESSÃO DE UMA ENZIMA
PROTEOLÍTICA A PARTIR DE UMA BIBLIOTECA
METAGENÔMICA DO CONSÓRCIO MICROBIANO
DEGRADADOR DE ÓLEO DIESEL**

Bianca de Melo Silveira

Orientadora: Profa. Dra. Eliana Gertrudes de Macedo Lemos

Dissertação apresentada à
Faculdade de Ciências Agrárias e
Veterinárias – UNESP, Câmpus de
Jaboticabal, como parte das exigências
para a obtenção do título de Mestre em
Microbiologia Agropecuária.

2015

Silveira, Bianca de Melo
5587p Prospeção e Expressão de uma enzima Proteolítica a partir de
uma Biblioteca Metagenômica do Consórcio Microbiano Degradador
de Óleo Diesel / Bianca de Melo Silveira. -- Jaboticabal, 2015
x, 78 p. : il. ; 28 cm

Dissertação (mestrado) - Universidade Estadual Paulista,
Faculdade de Ciências Agrárias e Veterinárias, 2015

Orientadora: Eliana Gertrudes de Macedo Lemos

Banca examinadora: Janete Aparecida Desidério, Rodrigo
Matheus Pereira

Bibliografia

1. Serino-protease. 2. Metagenoma. 3. Prospeção. I. Título. II.
Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 577.21

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



UNIVERSIDADE ESTADUAL PAULISTA

CAMPUS DE JABOTICABAL

FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS DE JABOTICABAL

CERTIFICADO DE APROVAÇÃO

TÍTULO: PROSPECÇÃO E EXPRESSÃO DE UMA ENZIMA PROTEOLÍTICA A PARTIR DE UMA BIBLIOTECA METAGENÔMICA DO CONSÓRCIO MICROBIANO DEGRADADOR DE ÓLEO DIESEL

AUTORA: BIANCA DE MELO SILVEIRA

ORIENTADORA: Profa. Dra. ELIANA GERTRUDES DE MACEDO LEMOS

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

Profa. Dra. ELIANA GERTRUDES DE MACEDO LEMOS

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

Prof. Dr. RODRIGO MATHEUS PEREIRA

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

Profa. Dra. JANETE APPARECIDA DESIDERIO

Departamento de Biologia Aplicada A Agropecuária / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal

Data da realização: 03 de dezembro de 2015.

DADOS CURRICULARES DO AUTOR

BIANCA DE MELO SILVEIRA – Nascida em 18 de Junho de 1992 em Matão, SP. Iniciou sua graduação em Tecnologia em Biocombustíveis em Fevereiro de 2010 na Faculdade de Tecnologia de Jaboticabal - FATEC JB na cidade de Jaboticabal, SP, concluindo seu curso em dezembro de 2012. Durante a graduação, foi integrante do projeto INTRALAB – Gerenciamento e descarte de resíduos do campus Unesp Jaboticabal, parceria da FATEC/ UNESP/ FUNEP, sob orientação do Prof. Dr. Marcelo Henrique Armoa e Profa. Dra. Luciana Maria Saran. Estagiou no Laboratório de Bioquímica de Micro-Organismos e Plantas na Faculdade de Ciências Agrárias e Veterinárias - Universidade Estadual Paulista “Júlio de Mesquita Filho” – (FCAV-UNESP), na cidade de Jaboticabal, SP sob a orientação da Profa. Dra. Eliana Gertrudes de Macedo Lemos. Em agosto de 2013 ingressou no curso de Pós - graduação em Microbiologia Agropecuária na Faculdade de Ciências Agrárias e Veterinárias - Universidade Estadual Paulista “Júlio de Mesquita Filho” – (FCAV-UNESP), na cidade de Jaboticabal, SP, também sob a orientação da Profa. Eliana Gertrudes de Macedo Lemos, como bolsista Capes.

Epígrafe

“TENTE!

Não diga ser incapaz, sem antes haver tentado.

Antes de mais nada, tente. Lute o quanto pode para conseguir.

E, se realizar um grande sonho, não se sinta surpreso(a).

O possível está ao seu alcance. A força está em VOCÊ.

No impossível DEUS lhe ajuda. Então, por que desistir?

TENTE!

Lembre-se: Não existe felicidade sem merecimento.

Não se chega a ser feliz, sem uma grande dose de sacrifícios.

A felicidade é GRAÇA que não chega de graça”. Prof. Henrique Melo

Dedicatória

Dedico, aos meus pais **Lenice** e **Airton**, e minha irmã **Gabriele** por seu amor, apoio e incentivo nesta longa caminhada. Aos meus queridos amigos que me ajudaram com palavras de estímulo nos momentos de desespero e preocupações, acreditando em meu potencial, compartilhando grandes alegrias e superações.

AGRADECIMENTOS

Agradeço primeiramente a **Deus** que me concedeu o fôlego de vida, me proporcionou saúde, iluminou meus caminhos e me deu forças para a realização deste trabalho;

À minha família, em especial aos meus pais **Airton e Lenice**, e minha irmã **Gabriele** que sempre estiveram ao meu lado em todos os momentos de felicidade e dificuldade, encorajando-me a vencer e conquistar os objetivos para um sucesso brilhante. Amo vocês!

À professora **Dra. Eliana Gertrudes de Macedo Lemos**, pela orientação, me concedendo a oportunidade de aprender e realizar este trabalho em seu laboratório;

Ao **Ms. Elwi Guillermo Machado de Sierra**, pela coorientação, persistência, auxílio na condução do experimento e transferência de conhecimentos;

Ao **Programa de Pós – Graduação em Microbiologia Agropecuária e à Faculdade de Ciências Agrárias e Veterinária – FCAV-UNESP** Jaboticabal pela estrutura, incentivo e suporte, tornando possível meu mestrado;

Ao **Departamento de Tecnologia** pela infraestrutura na execução deste projeto;

A **CAPES** pela bolsa concedida;

Ao **Dr. João Carlos Campanharo**, técnico do laboratório, pela assistência e ensinamentos transmitidos;

Ao **Prof. Dr. Manoel Victor Franco Lemos**, responsável pelo Laboratório de Genética de Bactérias, pelas células competentes cedidas;

Aos meus amigos de trabalho que me concederam informações, e opiniões para a realização deste trabalho: **Elisângela Gomes, Thaís Maester, Aliandra Malaman, Karla Goulart, Mariana Pereira, Gabriela Fernandes, Rosmeriana Garcia, Wellington Otori, Milena Tavares, Laís Sacco**, em especial a **Elisângela** pelo profissionalismo, disposição, competência e conselhos nos momentos de desespero e dificuldades. À grande amiga, companheira e irmã **Géssica Silveira**, que esteve presente em grandes

momentos de minha vida, compartilhando tristezas e alegrias. A vocês, fica o meu carinho e eterna gratidão!

Ao **Wellington Araújo**, pela amizade, companheirismo, compreensão, zelo e amor... Por estar ao meu lado me apoiando e incentivando;

A todos da equipe **LBMP**, que me acolheu e de alguma forma contribuiu para a realização deste trabalho. Obrigada pelos momentos de descontrações, comemorações, troca de experiências tanto profissionais quanto pessoais. Vocês foram importantes para o desenvolvimento deste trabalho.

A todos que contribuíram direta ou indiretamente para a realização deste projeto, que fizeram parte desta trajetória, que me alegraram, apoiaram, ensinaram, aconselharam, incentivaram, acreditaram, deixo aqui o meu **MUITO OBRIGADO!**

SUMÁRIO

LISTA DE TABELAS	iii
LISTA DE FIGURAS	iv
ABREVIACÕES.....	vi
LISTA DE UNIDADES	viii
RESUMO.....	ix
ABSTRACT	x
1. INTRODUÇÃO	11
2. REVISÃO DE LITERATURA	13
2.1 A metagenômica e o estudo da diversidade do solo	13
2.2 Proteases	16
3. OBJETIVOS	21
3.1 Objetivo geral	21
3.2 Objetivos específicos	21
4. MATERIAL E MÉTODOS	22
4.1 Prospeção e medição de atividade proteolítica	24
4.2 Extração do DNA fosmidial	25
4.3 Quantificação e análise do DNA	25
4.4 Digestão do DNA fosmidial	25
4.5 Construção de sub-biblioteca.....	26
4.5.1 Reação de nebulização	26
4.5.2 Precipitação e quantificação do DNA fragmentado	27
4.5.3 Reação de reparo das extremidades dos fragmentos	27
4.5.4 Reação de fosforilação dos fragmentos reparados	28
4.5.5 Seleção do tamanho dos insertos.....	28
4.5.6 Reação de ligação do inserto ao vetor	29
4.5.7 Transformação bacteriana da célula competente	31
4.5.8 Coleta e estoque dos clones transformados.....	32
4.6 Sequenciamento e análise do clone	32
4.7 Clonagem do gene codificador da proteína orf19 em vetor pET28a ...	33
4.7.1 Amplificação do gene de interesse	33
4.7.2 Reações de PCR	34

4.7.3	Digestão das Extremidades	35
4.7.4	Reação de Ligação do inserto ao vetor	36
4.7.5	Transformação da Célula de <i>Escherichia coli</i>	37
4.7.6	Coleta e estoque dos clones	37
4.7.7	Confirmação da clonagem e sequenciamento.....	38
4.8	Ensaio de expressão e extração da proteína ORF19.....	39
4.8.1	Eletroforese em gel de poliacrilamida SDS-PAGE	40
4.9	Zimograma	40
5.	RESULTADOS E DISCUSSÃO	42
5.1	Prospecção e medição de atividade proteolítica	42
5.2	Extração do DNA fosmidial	44
5.3	Digestão do DNA fosmidial	45
5.4	Construção da sub-biblioteca.....	47
5.5	Análise da sequência	48
5.5.1	Análise das relações filogenéticas e identificação dos sítios catalíticos	53
5.6	Clonagem do gene em vetor de expressão pET28a	56
5.6.1	Amplificação do gene de interesse	56
5.6.2	Construção pET28a-orf19 e transformação em células de <i>Escherichia coli</i> BL21(D3)	57
5.6.3	Confirmação da clonagem.....	58
5.6.4	Expressão da proteína ORF19	59
5.6.5	Zimograma	65
6.	CONCLUSÕES	66
	REFERÊNCIAS.....	67

LISTA DE TABELAS

Tabela 1. Maiores produtores mundiais de proteases (RAO <i>et al.</i> , 1998).	16
Tabela 2. Classificação e divisão das proteases (RAO <i>et al.</i> , 1998).....	18
Tabela 3 - Classificação e uso de endopeptidases (RAO <i>et al.</i> , 1998).....	19
Tabela 4. Sequências dos oligonucleotídeos e parâmetros utilizados para amplificação do gene codificador da ORF19.....	34
Tabela 5. Quantificação de DNA em NanoDrop TM 1000 Spectrophotometer dos clones que apresentaram atividade proteolítica mediante atividade com azocaseína como substrato.....	44
Tabela 6. " <i>Blast</i> " do <i>contig</i> completo do clone PL3C3 com as sequências do NCBI, através da ferramenta BLASTN. Foram colocados os quatro primeiros resultados fornecidos pelo NCBI.	49
Tabela 7. Predição das ORFs identificadas no <i>contig</i> do clone PL3C3 através do programa ORF Finder a partir do NCBI.....	50
Tabela 8. BLAST e análise da ORF19 através de banco de dados MEROPS - Sequência de peptidase.....	53
Tabela 9. Características físico- químicas da proteína recombinante ORF19 obtidas a partir da ferramenta ProtParam	59

LISTA DE FIGURAS

Figura 1. Enzimas no mercado mundial (Fonte: RAO et al., 1998).	15
Figura 2. Hidrólise enzimática de uma proteína hipotética (BERG et al., 2004).	18
Figura 3. Fluxograma metodológico demonstrando as etapas realizadas neste trabalho	23
Figura 4. Esquema ilustrativo apresentando características do vetor PUC19DNA/SmaI (Fermentas).	30
Figura 5. Esquema ilustrativo apresentando características do vetor pET28a (Novagen).....	36
Figura 6. Análise <i>in vitro</i> da biblioteca metagenômica em meio de cultura LB, suplementado com 1% de skim milk (p/v) e 12,5 ug/mL de clorafenicol (v/v). .	42
Figura 7. Estimativa da atividade proteolítica.	43
Figura 8. Perfil eletroforético em gel de agarose 0,8% da extração do DNA fosmidial do clone selecionado.....	45
Figura 9. Perfil eletroforético em gel de agarose 0,8% da restrição do DNA fosmidial dos clones..	46
Figura 10. Perfil eletroforético da nebulização do clone PL3C3 em gel de agarose 0,8%..	47
Figura 11. Perfil eletroforético do resultado da eluição do clone PL3C3, utilizando Gelase, em gel de agarose 0,8%.	48
Figura 12. Mapa físico do DNA fosmidial do clone PL3C3 gerado pelo programa DNAPlotter (CAVER et al., 2009).....	52
Figura 13. Árvore filogenética baseado em sequências de aminoácidos dos membros da família proteolítica S9 obtidos a partir do banco de dados MEROPS.....	54
Figura 14. Alinhamento múltiplo de sequências de aminoácidos da ORF19 com cinco membros representativos da família S9U, representando as regiões conservados de enzimas proteolíticas.	55
Figura 15. Perfil eletroforético em gel de agarose 1% mostrando padronização da temperatura de pareamento das reações de PCR em termociclador de gradiente (Biometra).....	56

Figura 16. Perfil eletroforético em gel de agarose 1% da reação de PCR purificada com o Kit Wizard SV Plus Gel and PCR Clean-Up System (Promega).	57
Figura 17. Perfil eletroforético em gel de agarose 1% para confirmar a presença do inserto referente à ORF19.	58
Figura 18. Teste de expressão. Eletroforese em gel de poliacrilamida 10%, corado com "Coomassie Brilliant Blue".	60
Figura 19. Teste de expressão. Eletroforese em gel de poliacrilamida 10%, corado com "Coomassie Brilliant Blue"..	61
Figura 20. Análise de SDS-PAGE 10% da ORF19 expressa em <i>E.coli</i> BL21 (DE3):.....	62
Figura 21. Análise de presença de peptídeo sinal na sequência da proteína ORF 19. (SignalP 4.1 Server- CBS (http://www.cbs.dtu.dk/services/SignalP/)).	64
Figura 22. Comparação do perfil proteico em gel SDS-PAGE e análise da atividade proteolítica em Zimograma.	65

ABREVIações

aa- Aminoácidos

DNA- ácido desoxirribonucleico

dNTP- desoxirribonucleotídeos fosfatados

EDTA- ácido etilenodiaminotetracético

IPTG – isopropil- β -D-tiogalactopiranosídeo

p/v – peso por volume

TEB- solução tampão Tris-ácido bórico- EDTA

Tris- tris[hidroximetil]aminometano

UV- luz ultravioleta

v/v- volume por volume

BLAST – *Basic Local Alignment Sequence Tool*

DO600 – densidade ótica a 600 nm de comprimento de onda

GenBank – banco de sequências de genes e proteínas do *NCBI*

Kb – mil pares de base

NCBI – *National Center for Biotechnology Information*

ORF – *Open Reading Frame* (região codificadora de proteína)

pb – pares de bases

PCR – *Polymerase Chain Reaction* (reação em cadeia da polimerase)

MEROPS – *The Peptidase Database*

pH – potencial hidrogeniônico $\log [H^+]$

pI – ponto isoelétrico

rpm – rotações por minuto

SAP – *Shrimp Alkaline Phosphatase* (fosfatase alcalina de camarão)

SDS – dodecil sulfato de sódio

SDS-PAGE – Eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio

IUBMB (União Internacional de Bioquímica e Biologia Molecular)

mRNA – ácido ribonucleico mensageiro

G – glicina

S – serina

L – leucina

W – triptófano

Y – tirosina

H – histidina

A – alanina

F – fenilalanina

M – metionina

T – treonina

LISTA DE UNIDADES

g – aceleração da gravidade

g – grama

kb – kilobase

L – litro

M – molar

mg – miligrama

mL – mililitro

mM – milimolar

μg – micrograma

μL – microlitro

μM – micromolar

ng – nanogramas

pb – pares de bases

U – unidades

V- volt

KDa- Quilo Daltons

PROSPECÇÃO E EXPRESSÃO DE UMA ENZIMA PROTEOLÍTICA A PARTIR DE UMA BIBLIOTECA METAGENÔMICA DO CONSÓRCIO MICROBIANO DEGRADADOR DE ÓLEO DIESEL

RESUMO - A partir de uma biblioteca metagenômica de um consórcio microbiano degradador de óleo diesel, um clone (PL3C3) expressando um novo gene de protease foi selecionado dentre vinte e seis positivos para degradação de azocaseína. Através do sequenciamento, foi analisado um contig e identificado uma ORF que codifica a proteína, denominada ORF19 composta por 966 pb, 321 aminoácidos, exibindo 69% de identidade com uma alfa/beta hidrolase de *Parvibaculum lavamentivorans* DS-1 (número de acesso ABS62095.1). A sequência da ORF19 após análise em um banco de dados de protease – MEROPS – mostrou 74,3% de identidade com uma serino-protease da família S9U (número de acesso MER095859). Através da análise filogenética, foi possível sugerir que a ORF19 é um novo membro desta família exibindo a tríade catalítica e motivos conservados característicos. A massa molecular da proteína foi estimado em 35 kDa, tendo por base a sua composição em aminoácidos e estimativas junto ao programa Protparam, e considerando o padrão de migração eletroforética da proteína em condições desnaturantes. O gene da ORF19 foi clonado no vetor de expressão pET28a, porém, a proteína recombinante não foi super-expressa em *Escherichia coli* BL21 (DE3). Contudo, a troca de hospedeiro, vetor, e estudos para uma super-expressão é necessário para a obtenção da proteína, visando sua futura caracterização. Este estudo contribui para a descoberta de uma nova serino-protease utilizando abordagens moleculares, como a metagenômica.

Palavras Chaves: Serino-protease. Metagenoma. Prospeção.

PROSPECT AND EXPRESSION OF A PROTEOLYTIC ENZYME FROM A METAGENOMIC LIBRARY DEGRADING MICROBIAL CONSORTIUM OIL DIESEL

ABSTRACT - From a metagenomic library of a microbial consortium that degrades diesel oil, a clone (PL3C3) expressing a novel protease gene was selected from twenty six positive for azocasein degradation. Through sequencing, an analysis was made contig and identified an ORF encoding a protein, called ORF19 composed of 966 bp, 321 amino acids, exhibiting 69% identity with an alpha / hydrolase beta *Parvibaculum lavamentivorans* DS-1 (accession number ABS62095. 1). The sequence of ORF19 after analysis of a protease database - MEROPS - showed 74.3% identity with a serine protease family S9U (access number MER095859). Through phylogenetic analysis, it was possible to suggest that ORF19 is a new member of this family displaying the catalytic triad and characteristic conserved motifs. The molecular mass of the protein was estimated at 35 KDa, based on its amino acid composition and estimates ProtParam by the program, and considering the pattern of electrophoretic migration of the protein under denaturing conditions. The ORF19 gene was cloned into the pET28a expression vector, but not the recombinant protein was overexpressed in *E. coli* BL21 (DE3). However, the exchange of host, vector, and studies for overexpression is required to obtain the protein, for their further characterization. This study contributes to the discovery of a new serine protease using molecular approaches, as metagenomic.

Key Words: Serino-protease. Metagenome. Prospection.

1. INTRODUÇÃO

O solo é um ambiente complexo sendo um extraordinário reservatório da diversidade genética microbiana e carbono orgânico, onde os procariotos são elementos essenciais no sistema de decomposição (ROBE et al., 2003; WHITMAM et al., 1998). O número de micro-organismos e sua biomassa coletiva variam dentro e entre os diferentes tipos e condições de solos, apresentando uma imensa diversidade genética, além de contribuir para a determinação da riqueza e complexidade desses organismos para utilização em processos biotecnológicos (BEARE et al., 1995; WHITMAM et al., 1998).

Durante muitos anos, a diversidade do solo tem sido explorada baseada no cultivo e isolamento de espécies microbianas (DANIEL, 2005), no entanto, técnicas de biologia molecular aplicadas ao estudo do meio ambiente têm colaborado significativamente para amplos avanços do conhecimento sobre a diversidade microbiana em amostras ambientais a partir de dados genômicos, sem a necessidade de isolamento e cultivo, através da clonagem direta de DNA (HUGENHOLTZ e PACE, 1996).

Nas últimas décadas o desenvolvimento de métodos para isolamento de ácidos nucleicos de amostras ambientais, permitiu o acesso a uma diversidade microbiana desconhecida. A análise de ácidos nucleicos diretamente extraídos do ambiente permitiu aos pesquisadores estudos de comunidades microbianas naturais (RIESENFIELD et al., 2004). Portanto, a metagenômica explora o fato de que enquanto alguns organismos são cultiváveis e outros incultiváveis, todos eles são formas vivas fundamentadas em DNA como portador de informação genética. Deste modo, a metagenômica disponibiliza ferramentas para acessar, armazenar e analisar este DNA ambiental total (LEVEAU, 2007).

Através da utilização da técnica metagenômica, a biotecnologia procura por novos biocompostos, possibilitando a ampliação de novas abordagens para encontrar e identificar novos genes codificadores de produtos de interesse biotecnológico, como novas enzimas, por meio da construção de uma biblioteca metagenômica a fim de estudar as vias metabólicas secundárias da

comunidade microbiana, utilizando o genoma de micro – organismos complexos de habitats naturais (LIAW et al., 2010).

Empregada em diversos setores, a tecnologia enzimática (enzimas adquiridas pela biotecnologia para aplicação em escala industrial) busca por novos produtos biodegradáveis, que não ataquem o meio ambiente, possuindo diversas características para sua utilização, como a alta especificidade, e seu processo de obtenção mais econômico, estimulando o crescimento industrial de enzimas (MOREIRA et al., 2002; RAO et al., 1998).

A protease é o tipo de enzima mais importante do ponto de vista industrial. São enzimas que hidrolisam ligações peptídicas entre resíduos de aminoácidos (RAO et al., 1998). Podem ser utilizadas em diversas atividades industriais, tais como processamento de bebidas, alimentos, processamento de couro e pele, indústrias têxteis, formulação de detergentes, no amaciamento de carne e formulação de medicamentos (GODFREY e WEST, 1996; LUNA et al., 2002; GHORBEL et al., 2003; HAKI e RAKSHIT, 2003; GENCKAL e TARI, 2006; SILVA NEVES et al., 2006; MUKHERJEE et al., 2008; PATIL e CHAUDHARI, 2009). Os micro-organismos são a fonte mais empregada para a obtenção de protease de uso industrial, obtidos através de processos fermentativos (POZA et al., 2001; RAO et al., 1998). As bactérias, fungos e leveduras proteolíticas são pesquisados a fim de alcançar novos genes que sejam ótimos produtores de proteases, além de aumentar a produtividade e a estabilidade enzimática daqueles que já são admitidos como micro-organismos proteolíticos (DURAND-POUSSEREAU e FEVRE, 1996; BRAGA et al., 1998; KITANO et al., 2002; BEG e GUPTA, 2003; AZEREDO et al., 2004).

2. REVISÃO DE LITERATURA

2.1 A metagenômica e o estudo da diversidade do solo

A diversidade presente em populações microbianas do solo vem despertando aos pesquisadores estudos da riqueza genética presente nos micro – organismos com potencial na aplicação industrial (MANFIO, 2002). A complexidade das comunidades microbianas são o resultado das interações de fatores como estrutura do solo, pH, variação climática, conteúdo da água, e atividade biótica (ROBE et al., 2003). A diversidade microbiana pode então ser definida como a riqueza de elementos biológicos, tais como espécies ou gêneros, e genes, presentes num ecossistemas (MARTA et al., 2004).

O solo é um extraordinário habitat para micro – organismos, formado por minerais, água, ar e materiais orgânicos (como carbono), elementos indispensáveis para a biota terrestre. Dentre os papéis dos micro – organismos, estão a decomposição de plantas e animais, fixação de nitrogênio atmosférico, conversão de elementos como fósforo, manganês e ferro (WHITMAN et al., 1998) em formas que possam ser utilizadas para a nutrição e crescimento de plantas, fertilidade do solo, representando a principal porção da diversidade genética viva e produzindo importantes componentes da atmosfera terrestre. Assim, todas as abordagens toleram a afirmação de que o solo representa um dos mais diversos habitats de micro-organismos (PACE, 1997; TORSVIK et al., 2002a). Segundo Torvisk e Ovreas (2002b), um grama de solo pode conter até 4000 espécies distintas, dentre as quais a maior parte não é cultivável. O número e o tipo de micro-organismos variam de acordo com as diferentes condições de solo, sendo as bactérias as mais abundantes (WHITMAN et al., 1998).

Tem se tornado cada vez mais difícil a obtenção de culturas puras de micro – organismos devido ao desconhecimento de seus metabolismos, ou das interações em seu ambiente natural (VENTER et al., 2004). Ferramentas moleculares e tecnologias baseadas em sequenciamento genético chegam a diminuir essas restrições e mostrar novas perspectivas sobre a variedade

microbiana. Embora uma pequena parcela (1%) de micro-organismos seja cultivável em laboratório por técnicas padrões de cultivo, deixando desconhecido seu potencial metabólico, econômico e industrial (NIKOLAKI & TSIAMIS, 2013), os avanços com ferramentas moleculares revelaram que a variedade genética de micro-organismos na natureza é abundantemente maior que a refletida nas coleções e bancos de linhagens (AMANN et al., 1995).

A ampliação de novas técnicas moleculares como a metagenômica permitiu uma maior exploração através do banco de DNA de comunidades microbianas, despertando a busca por novos genes. O termo metagenoma proposto em 1998 apresenta o campo estudado como ambiente genômico ou comunidade genômica (HANDELSMAN et al., 1998). Esta técnica consiste na extração de DNA diretamente do ambiente, (eliminando a etapa de isolamento e cultivo prévio) e a construção de uma biblioteca realizando a clonagem em vetor adequado (fosmídeo, plasmídeo, cosmídeo), gerando assim as bibliotecas de DNA, permitindo desta forma, a descoberta e o estudo de novos genes e/ou vias metabólicas de interesse biotecnológico (KNIGHT et al., 2003; HANDELSMAN, 2004; LIAW et al., 2010). Portanto, a metagenômica é um campo de pesquisa de acelerado crescimento que evolui ao entendimento da função, evolução e diversidade dos micro-organismos presentes em ambientes como solo, água e sistemas digestivos de animais e humanos (HUSON, RICHTER, et al., 2009).

A partir de análises genômicas de uma população microbiana pode-se acessar uma quantidade quase infinita de genes de organismos incultiváveis (KNIGHT et al., 2003), além da combinação tecnológica molecular do último século, que permitiu aos cientistas a investigação procariótica, habilitando a entrada ao vasto potencial biotecnológico escondido nestes micro - organismos (STEELE e STREIT, 2005).

Os micro-organismos do solo são capazes de produzir enzimas extracelulares, constituindo o principal meio pelo qual degradam os complexos maiores em moléculas pequenas para serem assimiladas (ALLISON e VITOUSEK, 2004). A utilização da técnica metagenômica permite encontrar novos genes que codifiquem produtos biotecnológicos como enzimas com múltiplas aplicações, sendo as proteases (ACEVEDO *et al.*, 2008; ZHANG *et al.*, 2011), lipases (GARCIA *et al.*, 2015; JEON *et al.*, 2009; HARDEMAN &

SJOLING, 2007; RHEE *et al.*, 2005), amilases (LAMMLE *et al.*, 2007; RICHARDSON *et al.*, 2002), esterases (FERRER *et al.*, 2005; JEON *et al.*, 2009), e quitinases (LECIEIR *et al.*, 2004; COTTRELL *et al.*, 1999). Estas enzimas e outros compostos obtidos a partir de micro-organismos apresentam grande importância para várias áreas da indústria, sendo utilizada como chaves de ativação de componentes em vários alimentos industrializados, manufatura de papel, detergentes, compostos agroindustriais, biocombustíveis entre outros (VAN BEILEN e LI, 2002).

Aumentar o conhecimento e maior compreensão da diversidade e interações dos micro-organismos em ambientes diversos é importante para a obtenção de benefícios econômicos estratégicos, como aqueles que podem ser adquiridos com a descoberta de micro-organismos que oferecem potencial para serem utilizados em processos biotecnológicos (PAIXÃO *et al.*, 2010).

Estima-se que o mercado mundial de enzimas atinja a 7 bilhões de dólares em 2017 com um crescimento anual do 6,7% (ADRIO e DEMAÏN, 2014). Na atualidade, das 4000 enzimas prospectadas, 200 são de origem microbiana utilizada comercialmente, na qual 70% destas enzimas são hidrolases do tipo lipase, carboxilase e protease (LI *et al.*, 2012).

Uma das enzimas de maior valor industrial são as proteases, estas são adquiridas principalmente dos micro-organismos, como o *Aspergillum* entre os fungos, e os *Bacillus* entre as bactérias, (JOO e CHANG, 2006). A utilização desta enzima em escala industrial apresenta aproximadamente 48% a 60% do mercado mundial/ano de enzimas, (Figura 1) (ANWAR e SALEEMUDDIN, 1998; GUPTA *et al.*, 2002a) por um valor de 5,1 bilhões de dólares (P. SINGHAL, P., NIGAM, V., e VIDYARTHI, A. 2012), onde só 4 empresas do setor privado controlam o mercado (Tabela 1).

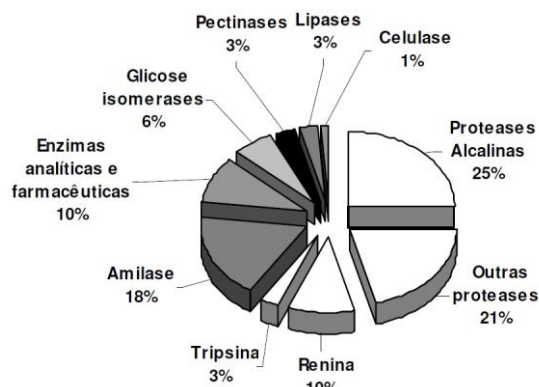


Figura 1. Enzimas no mercado mundial (Fonte: RAO *et al.*, 1998).

Tabela 1. Maiores produtores mundiais de proteases (RAO *et al.*, 1998).

Empresa	País	Fração no mercado
Novo Industries	Dinamarca	40%
Gist Brocades	Holanda	20%
Genecor International	Estados Unidos	10%
Mites Laboratories	Estados Unidos	10%
Outras		20%

2.2 Proteases

São um grupo de enzimas importantes também denominadas peptidases ou enzimas proteolíticas podem ser encontradas em diversos organismos (plantas, animais e micro-organismos) (SAJID e MCKERROW, 2002). Esta catalisa processos biológicos e hidrólises (quebra da molécula) de uma ligação peptídica que levaria muitos anos para ocorrer. Embora as proteases de plantas e animais não atendam à demanda mundial de enzimas, o interesse pelas proteases de origem microbiana tem sido cada vez maior, pois os micro-organismos representam uma excelente fonte devido ao seu rápido crescimento, espaço limitado exigido para seu cultivo, a facilidade de manipulação genética e a diversidade bioquímica (RAO *et al.*, 1998; HAKI E RAKSHIT, 2003).

A atividade proteolítica é uma capacidade própria de todos os micro-organismos, considerando seu papel energético e estrutural do metabolismo primário dos seres vivos. Entretanto, apenas aqueles que produzem quantidades relevantes de enzimas extracelulares são de importância industrial. Em organismos nativos, esta produção é fortemente influenciada pelas condições em que o micro-organismo se encontra, tais como fonte de carbono, nitrogênio, açúcares, dentre outros fatores quanto a temperatura, pH, densidade de inóculo e tempo de incubação, que podem afetar a quantidade de protease extracelular produzida (LÓPEZ-OTIN e BOND, 2008; FERRERO *et al.*, 1996; HAMEED *et al.*, 1999; NEHETE *et al.*, 1985). Contudo, com o

advento da tecnologia do de DNA recombinante é possível super-expressar uma proteína oriunda de micro-organismos desconhecidos, clonando seu DNA em vetor de expressão e utilizando um hospedeiro heterólogo como a *E. coli*, obtendo assim, uma produção significativa com baixo custo.

As proteases microbianas realizam importantes funções fisiológicas como a hidrólise de proteínas no meio externo (proteases extracelulares), admitindo a absorção de nutrientes pelos micro-organismos, enquanto que as proteases intracelulares são utilizadas em vários processos metabólicos e celulares (GUPTA *et al.*, 2002b).

Onipresentes na biologia, as peptidases possuem um envolvimento bioquímico ou fisiológico em muitos aspectos da célula e do organismo, incluindo a adaptação, crescimento, esporulação e germinação, nutrição, remodelação da superfície de células e tecidos, doença e morte (LODISH *et al.*, 2004), digestão de alimentos, transporte, coagulação e composição do sangue, ativação de zimogênios, ativação de proenzimas, processos de inflamação, liberação de peptídeos fisiologicamente ativo, ativação do sistema do complemento, entre outros, sendo consideradas altamente relevantes no contexto biotecnológico (LEUNG – TOUNG *et al.*, 2002; MAURER, 2004).

As proteases inicialmente estavam associadas ao catabolismo proteico; no entanto, tem-se admitido que a alta especificidade de hidrólise de peptídeos pode regular diversos processos biológicos (FREDERIKS e MOOK, 2004). Representante de um grande grupo de enzima hidrolíticas, são classificadas por sua estrutura, sítio ativo, local de ação, mecanismos de reação específica e faixa de pH, na qual apresentam atividade máxima, diferenciando – se em 3 formas: proteases ácidas, neutras ou alcalinas (RAO *et al.*, 1998). De acordo com a IUBMB (União Internacional de Bioquímica e Biologia Molecular) estão inseridas no subgrupo 4 do grupo 3 (Hidrolases), que por uma reação de hidrólise, clivam uma ligação peptídica de uma proteína adicionando uma molécula de água (Figura 2) (BERG *et al.*, 2004).

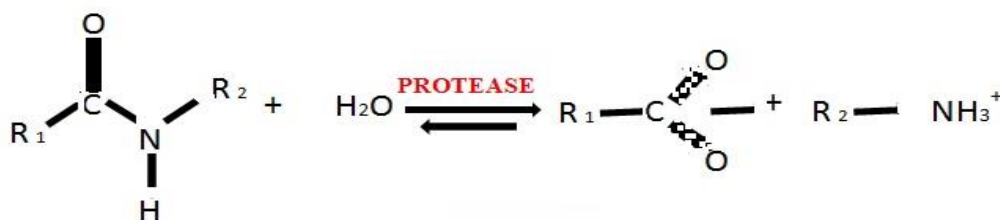


Figura 2. Hidrólise enzimática de uma proteína hipotética (BERG et al., 2004).

De maneira geral, as proteases são divididas em dois principais grupos: **exopeptidases** que atuam nas extremidades da cadeia, catalisando a hidrólises das ligações peptídicas próximas às extremidades C-terminal ou N-terminal liberando apenas um resíduo de aminoácido; e as **endopeptidases** que atuam preferencialmente nas regiões internas das cadeias polipeptídicas (Tabela 2) (BARRETT, 1994; RAO *et al.*, 1998). Ao gerar hidrolisados proteicos as endopeptidases são usadas em diversas funções (Tabela 3).

Tabela 2. Classificação e divisão das proteases (RAO *et al.*, 1998).

Local de clivagem no substrato	Sítio ativo da enzima	Número de resíduos de aminoácidos removidos
Exopeptidases	Aminopeptidases	Aminopeptidases
		Aminodipeptidases
		Aminotripeptidases
	Carboxipeptidase	Sítio ativo da carboxipeptidase
		Serinocarboxipeptidases
		Metalocarboxipeptidases
		Cisteínocarboxipeptidases
Endopeptidases	Serinoproteases	
	Aspartatoproteases	
	Cisteinoproteases	
	Metaloproteases	

Tabela 3 - Classificação e uso de endopeptidases (RAO *et al.*, 1998).

Grupo	Ph	Sítio Ativo	Exemplos de aplicações	Fontes
Serinoproteases	Alcalino	Ser (His)	Hidrólise proteica	Tripsina (pâncreas)
			Detergentes	Subtilisina (bacillus)
Metaloproteases	Neutro	Zn ²⁺	Cerveja	Bacillus
		Ca ²⁺	Couro	Aspergillus oryzae
Carboxiproteases	Ácido	Asp	Queijo	Quimosina (bezerro)
Tiolproteases	Tiol	Cys	Amaciamento de carne	Papaína (Papaya)

As famílias de proteases foram designadas de acordo com o resíduo de aminoácido presente no sítio catalítico, classificando - as em nove famílias com base na similaridade de sequência e mecanismo catalítico sendo: Aspártato (A), Cisteína (C), Glutâmina (L), Asparagina (N), Serina (S), Treonina (T)-proteases: são famílias que apresentam pelo menos um destes aminoácidos no sítio catalítico; as Metaloproteases (M): são as famílias dependentes de íons metálicos bivalentes para sua atividade; as Proteases mistas (P): famílias que apresentam pelo menos dois aminoácidos (C, S, T) no sítio catalítico; e as Proteases de mecanismos Desconhecido (L): ainda não elucidadas (RAWLINGS *et al.*, 2014). Diante desta classificação, quatro grupos são mais relevantes: serino - proteases, cisteíno - proteases, aspártico - proteases e metalo - proteases.

Serino-protease (SP): Enzimas mais estudadas são encontradas em eucariotos, bactérias e vírus, (PAGE E DI CERA, 2008), indicando uma participação vital no metabolismo desses organismos. Os representantes mais conhecidos são a quimiotripsina e a tripsina (WILSON *et al.*, 1997; GORMAN *et al.*, 2000a e 2000b). Ativas em pH alcalino e neutro, (RAO *et al.*, 1998) sua massa molecular geralmente está na faixa de 18 a 35 kDa, entretanto outros tipos de SP possuem massas moleculares maiores (ROMERO *et al.*, 2001). As SP são distinguidas por serem inibidas irreversivelmente por 3,4 - dicloroisocoumarina (3,4 - DCI) L-3-carboxitrans-2,3-epoxipropil-leucilamido (4-

guanidina) butano (E-64), diisopropilfluorofosfato (DFP), e tosil-L-lisina clorometil cetona (TLCK), e tiol *p* - cloromercuribenzoato (PCMB), em razão da presença de um resíduo de cisteína na região do centro ativo (RAO *et al.*, 1998).

Cisteíno – proteases: São enzimas encontradas em todos os organismos, sendo representadas principalmente pela papaína (nome da família) extraída a partir do látex de *Carica papaya*, constituindo a primeira cisteíno protease com estrutura tridimensional determinada (BELITZ e GROSCH, 1999; KAMPHUIS *et al.*, 1984; RAO *et al.*, 1998). Seu sítio ativo é composto por um resíduo de cisteína, por um de asparagina e um de histidina na tríade catalítica (BAIRD *et al.*, 2006). Agem via ataque nucleofílico do ânion sulfeto no sítio ativo, porém são ativas em uma larga faixa de pH (6,0 – 7,5), mas exercem uma melhor atividade dependendo do substrato, com pH entre 4,5 a 10 (KUNAKBAEVA *et al.*, 2003). Esta enzima é dividida em clãs que não seguem uma identidade de estrutura tridimensional, nem de sequência e provavelmente surgiram de diferentes linhas evolutivas (BARRETT e RAWLINGS, 2001; RAWLINGS e BARRETT, 1999).

Aspartil – proteases: Pertence à família pepsina, incluindo enzimas digestivas, catepsinas lisossomais e enzimas de processamento, obtidas principalmente a partir de fungos. São ativas em pH ácido, dependendo do substrato (BELITZ e GROSCH, 1999; RAO *et al.*, 1998). Seu sítio ativo esta localizado entre dois lobos homólogos, em que cada lobo contribui com um resíduo de aspartato, o qual é necessário para a atividade desta enzima (BEYNON e BOND, 1989).

Metalo – proteases: Obtidas a partir de bactérias, fungos e organismos superiores, as metalo-proteases são ativas em pH 6,0 – 9,0, sendo sua especificidade geralmente baixa (BELITZ e GROSCH, 1999; RAO *et al.*, 1998). Apresentam grande diferença em suas sequências e estruturas, porém a maioria destas enzimas contém um átomo de zinco cataliticamente ativo. Em alguns casos o zinco pode ser substituído por outro metal, como o níquel ou cobalto sem perda de atividade (BLUNDELL, 1994).

3. OBJETIVOS

3.1 Objetivo geral

Utilização de uma biblioteca metagenômica do consórcio microbiano degradador de óleo diesel (PAIXÃO et al., 2010), para detectar clones produtores de enzimas proteolíticas a fim de expressar seus genes codificadores.

3.2 Objetivos específicos

- ✓ Realizar análise *in vitro*, em meio de cultivo sólido com *skim milk* e medição da atividade proteolítica utilizando azocaseína, nos clones da biblioteca metagenômica para detecção de enzimas proteolíticas;
- ✓ Realizar a subclonagem do(s) clone(s) positivo(s) selecionado(s);
- ✓ Sequenciar o(s) clone(s) selecionado(s);
- ✓ Analisar as sequências do clone e obter a sequência completa do fragmento metagenômico;
- ✓ Analisar as sequências de aminoácidos e construir a árvore filogenética;
- ✓ Identificar a família proteolítica a qual o(s) gene(s) encontrado(s) pertence;
- ✓ Subclonar ORFs codificadoras de enzimas proteolíticas em pET28a;

4. MATERIAL E MÉTODOS

O material utilizado para esta pesquisa é oriundo de uma biblioteca metagenômica do consórcio microbiano degradador de óleo diesel, construída no Laboratório de Bioquímica de Microrganismos e Plantas (LBMP), localizado na Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias (FCAV) de Jaboticabal. As etapas foram realizadas de acordo com o fluxograma (Figura 3).

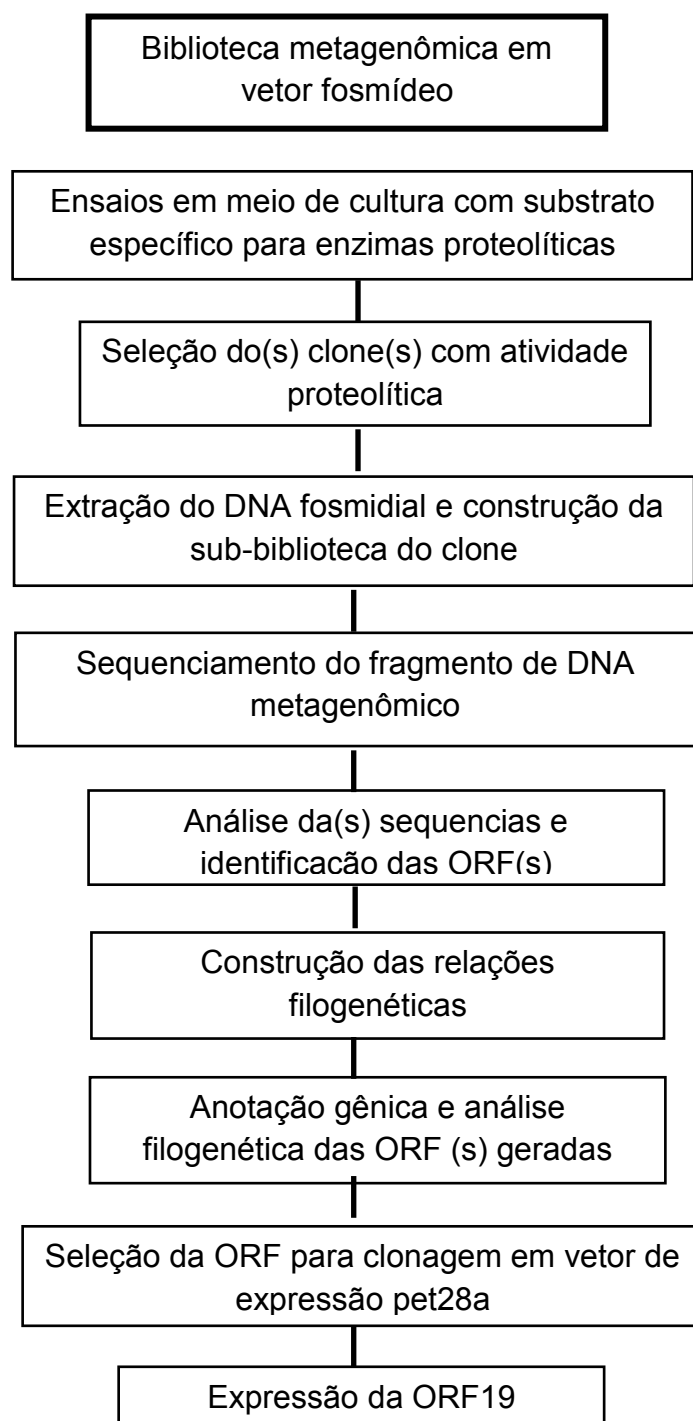


Figura 3. Fluxograma metodológico demonstrando as etapas realizadas neste trabalho

4.1 Prospecção e medição de atividade proteolítica

Os 4224 clones da biblioteca metagenômica do consórcio microbiano degradador de óleo diesel foram submetidos a um teste bioquímico no intuito de identificar os possíveis produtores de enzimas proteolíticas, utilizando o meio de cultura Luria-Bertani (Tryptona 10 g/L, Extrato de Levedura 5g/L, NaCl 10 g/L e Agar 15 g/L) suplementado com *Skim – Milk* a 1% (Difco) e clorafenicol (12,5 µg/mL), sendo posteriormente incubados por 72 horas à 37°C. Para os clones que apresentaram halo de hidrólise (possíveis proteolíticos) foram determinadas a atividade proteolítica na biomassa bacteriana, aplicando azocaseína como substrato (MORRIS et al., 2012). Os clones foram cultivados meio LB com clorafenicol, por 24 horas a 37°C e 200 r.p.m. Após o cultivo, as cepas foram centrifugadas a 2254 x g, 20°C por 10 minutos.

Inicialmente o *pellet* foi lavado duas vezes com tampão PBS [50 mM NaH₂PO₄, 300 mM de NaCl, 10 mM Imidazol, pH8] para eliminar possíveis resíduos do meio de cultura, em seguida foi ressuspenso em tampão de lise (Tris-HCl 0,1M pH8,8 + Gliserol 15%) e submetido a sonicação *Brason Sonifier 250* por 1 minuto, com *dyty cycle* a 10%, *output control* – 5 a 4°C. Em seguida, as amostras foram centrifugadas por 10 minutos a 20°C, 2254 x g. Após a centrifugação, o sobrenadante foi recuperado. A atividade proteolítica foi determinada usando como substrato 150 µL de azocaseína (5 mg/ml em Tris–HCl 50 mM), e 150 µL do sobrenadante seguindo de incubação por 4 horas a 37°C. Posteriormente foi adicionado 400 µL de ácido tricloroacético – TCA (10%) para interromper a reação e precipitar as proteínas, sendo a amostra centrifugada por 30 minutos, 20°C, 2254 x g. O sobrenadante foi transferido para tubos estéreis contendo 700 µL de NaOH (525 mM) e a absorbância da amostra foi medida a 442 nm (MORRIS et al., 2012). Os ensaios foram realizados em triplicata, definindo uma unidade (U) de protease como a quantidade de enzima requerida para aumentar em 0,01 a absorbância em uma hora.

4.2 Extração do DNA fosmidial

Os clones selecionados através da medição de atividade proteolítica foram submetidos à extração de DNA fosmidial através do kit Wizard® Plus SV Minipreps DNA Purification System (Promega, Madison, Wisconsin, Estados Unidos da América), seguindo as instruções do fabricante.

4.3 Quantificação e análise do DNA

O DNA fosmidial foi quantificado por espectrofotometria no NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, Delaware, Estados Unidos da América) e a relação 260/280 nm foi utilizada para avaliar a qualidade do material genético (DNA) e a presença de proteínas.

A análise da qualidade dos DNAs foi realizada em gel de agarose 0,8% contendo brometo de etídio (0,5 mg/mL). A eletroforese foi realizada em uma cuba horizontal e conduzida em tampão TBE 1X [Tris 89 mM, Ácido Bórico 89 mM e EDTA (ácido etilenodiaminotetracético) 2,5 mM, pH 8,3] a 80 V por 1 hora e 30 minutos. As amostras foram preparadas contendo 5 µL de DNA adicionada a 3 µL do carregamento [0,025% de azul de bromofenol (p/v) e 50% de glicerol (p/v)]. Foi aplicado o padrão de tamanho molecular 1 kb DNA Ladder (Fermentas, Burlington, Ontário, Canadá) e a imagem foi visualizada e documentada sob luz UV em aparelho fotodocumentador Gel Doc 1000 (Bio-Rad, Hercules, Califórnia, Estados Unidos da América), através do software Quantity One^R (Bio-Rad).

4.4 Digestão do DNA fosmidial

Uma alíquota de DNA fosmidial foi digerida com 3 enzimas de restrição – *AluI*, *EcoRI* e *PstI*, a fim de comparar o perfil de restrição e estimar o tamanho do inserto clonado no vetor fosmídial.

As reações de digestão foram feitas separadamente para cada clone, e em cada reação, 1 µg do DNA foi acrescentado à 10 U da enzima, 2 de µL tampão FastDigest® 10X (cat. n°; FD0504, Fermentas) e água deionizada ultra-pura para completar um volume de 20 µL no final da reação. As reações foram realizadas separadamente sendo submetidas à incubação a 37°C por 4 horas e em seguida, para inativação térmica das enzimas, a 70°C por 15 minutos.

O volume total de cada digestão foi aplicado em gel de agarose 1% (p/v), sendo o preparo realizado conforme metodologia descrita no item 4.3, porém a voltagem consistiu em 80 V por 3 horas com um marcador molecular 1 Kb DNA Ladder, e visualizada as diferentes restrições dos clones. Posteriormente, foi selecionado 1 clone para a construção das sub – biblioteca.

4.5 Construção de sub-biblioteca

O DNA fosmidial extraído do clone selecionado conforme descrito no item 4.3, foi utilizado para construção de sub-biblioteca *shotgun* em vetor pUC19, para possibilitar o acesso à informação genética por meio do sequenciamento do DNA.

4.5.1 Reação de nebulização

Com o objetivo de gerar fragmentos com tamanho de 1 a 3 kb para ser sub-clonado, o DNA fosmidial do clone foi inserido em nebulizador comercial pré-esterelizado em etanol 70% e luz UV por 2 minutos. Foi adicionado ao nebulizador aproximadamente 1 µg de DNA, 50 mM de Tris-HCl, pH 8,0, 15 mM de MgCl₂, ambas soluções filtradas em membrana de 0,22 µm, 25% de glicerol a 50% (v/v) e água deionizada estéril para completar o volume de 2 mL. O nebulizador foi conectado a um tubo de gás nitrogênio e mantido sob pressão 3Kgf/cm² durante 10 segundos a 4°C.

O nebulizador foi ajustado a um tubo de centrífuga para ser centrifugado a 1000 rpm., a 4°C, por 2 minutos, a fim de precipitar o material. O volume recuperado foi transferido em alíquotas de 500 µL para tubos de 1,5 mL.

4.5.2 Precipitação e quantificação do DNA fragmentado

A cada alíquota do DNA nebulizado foram adicionados 50 µL (1/10 do volume) de acetato de sódio 3M, pH 5,2 e 1000 µL (2 volumes) de etanol absoluto gelado. As amostras foram homogeneizadas e incubadas por 22h a -20°C. Após a incubação, os tubos contendo DNA precipitado foram centrifugados a 4°C, 3624 x g, por 25 minutos. Posteriormente, ao sedimento obtido foi adicionado 1 mL de etanol 70% (v/v) gelado para limpeza do precipitado e centrifugado novamente nas condições já descritas por 10 minutos. O sedimento foi incubado em temperatura ambiente para eliminar o excesso de etanol e ressuspendido com 20 µl de água Milli-Q estéril. Em seguida, o DNA foi incubado a 37°C por 1 hora e 30 minutos. O volume das alíquotas foi reunido em tubo único de 1,5 mL.

Uma alíquota (3µl) do DNA nebulizado foi visualizada e quantificada por eletroforese em gel de agarose 0,8% (p/v) a 80V por 1 hora de 30 minutos, conforme descrito no item 4.3.

4.5.3 Reação de reparo das extremidades dos fragmentos

Após a obtenção do fragmento de DNA, este foi submetido ao reparo de sua extremidade. Em cada reação foram utilizados aproximadamente 1 µg de DNA, 10 µL de Tampão NEBuffer2 1X (50mM NaCl, 10 mM Tris-HCl, 10 mM MgCl₂ e 1 mM Dithiothreitol, pH 5,9) (New England BioLabs, *Ipswich, Massachusetts, EUA*), 1 µL de BSA 100X (20 mM KPO₄, 50 mM NaCl, 0,1 mM EDTA e 5% de Glicerol, pH 7,0, New England BioLabs), 1 µL de dNTP (10 mM, Fermentas), 3 U de T4 DNA Polimerase (New England BioLabs) e o volume foi

completado para 100 μ L com água deionizada estéril. Uma unidade da enzima T4 DNA polimerase é definida como a quantidade de enzima requerida para incorporar 10 nmol de dNTP em material precipitado.

Em seguida, a reação foi submetida à incubação por 20 minutos a 12°C. Posteriormente foi adicionado 5 U de Klenow Fragment (New England BioLabs) e 0,5 μ L dNTP (10 mM, Fermentas), sendo a reação incubada a 37°C por 30 minutos, seguido de outra incubação para inativação térmica da enzima a 75°C por 20 minutos.

A enzima T4 DNA polimerase e Klenow DNA polimerase foram empregadas para incorporação de nucleotídeos complementares livres, nos terminais 3' da dupla fita de DNA.

4.5.4 Reação de fosforilação dos fragmentos reparados

Depois de reparado, realizou-se a fosforilação dos fragmentos inseridos para obter a clonagem em vetor escolhido. Para isto, aos fragmentos de DNA reparados (101,5 μ L) foram adicionados 12 μ L de tampão T4 DNA Ligase 10X, que contém a quantidade necessária de ATP para a reação, (50 mM Tris-HCl, 10 mM $MgCl_2$, 10 mM Dithiothreitol e 1 mM de ATP, pH 7,5, New England BioLabs), 10 U de T4 Polinucleotídeo Quinase (New England BioLabs) e água deionizada estéril para completar o volume da reação para 120 μ L. Posteriormente, a reação foi incubada a 37 °C por 30 minutos e logo após a 65 °C por 20 minutos, para a inativação térmica da enzima.

4.5.5 Seleção do tamanho dos insertos

Para a seleção do tamanho dos insertos a serem clonados, retirou-se o material de interesse do gel de agarose. Para isso, todo o material reparado e fosforilado foi submetido à eletroforese preparativa em gel de agarose de baixo ponto de fusão a 1% (p/v), isento de brometo de etídio. Através deste gel foi

possível selecionar o tamanho do DNA entre 1 a 3 kb, para posterior recuperação.

Foram aplicados, na seguinte ordem no gel de agarose: 5 µL da amostra com 3 µL de tampão de carregamento, 6 µL do padrão molecular 1 kb DNA Ladder (Fermentas), o volume total da amostra (120 µL) adicionado de 10 µL de azul de bromofenol 6X concentrado (3,3,5,5-tetrabromofenolsulfonftaleína) em três canaletas unidas. A eletroforese foi conduzida em tampão TEB 1 X isento de brometo de etídio, com voltagem inicial de 100 V por 30 minutos e posteriormente a 80 V durante 2 horas.

Após a eletroforese, o gel de agarose por meio de um corte vertical, foi dividido em duas porções, resultando em uma porção com os insertos a serem recuperados e outra porção contendo o marcador e a amostra comparativa para serem corados. A porção do gel com o marcador foi corada com brometo de etídio (5mg/mL) e observada sob luz UV com um transiluminador, com a finalidade de guiar a obtenção do inserto na porção não corada colocada ao lado, sendo esta cortada e retirada do restante do gel. Esta parte retirada foi dividida em tubos de 1,5 mL com aproximadamente 400 mg de gel por tubo e o restante do gel foi corado para ser observado se a banda foi selecionada corretamente. A visualização foi feita em luz UV e documentado em fotodocumentador Gel Doc 1000 (Bio-Rad).

A eluição da banda selecionada foi realizada utilizando gelase seguindo instruções do fabricante, posteriormente o DNA foi precipitado (ver item 4.5.2) e a quantificação realizada em aparelho de espectrofotometria. Uma alíquota de 3 µL de cada amostra eluída foi quantificada em gel de agarose 0,8% contendo brometo de etídeo (0,5 mg/mL), utilizando o padrão molecular 1 kb DNA Ladder (Fermentas), conforme já descrito no item 4.3.

4.5.6 Reação de ligação do inserto ao vetor

O DNA fosmidial extraído do clone selecionado foi clonado em vetor pUC19 desfosforilado e digerido com SmaI para permitir o acesso à informação

genética. O vetor pUC19DNA/SmaI (Fermentas) possui 2.686 pares de bases e é ilustrado na Figura 4.

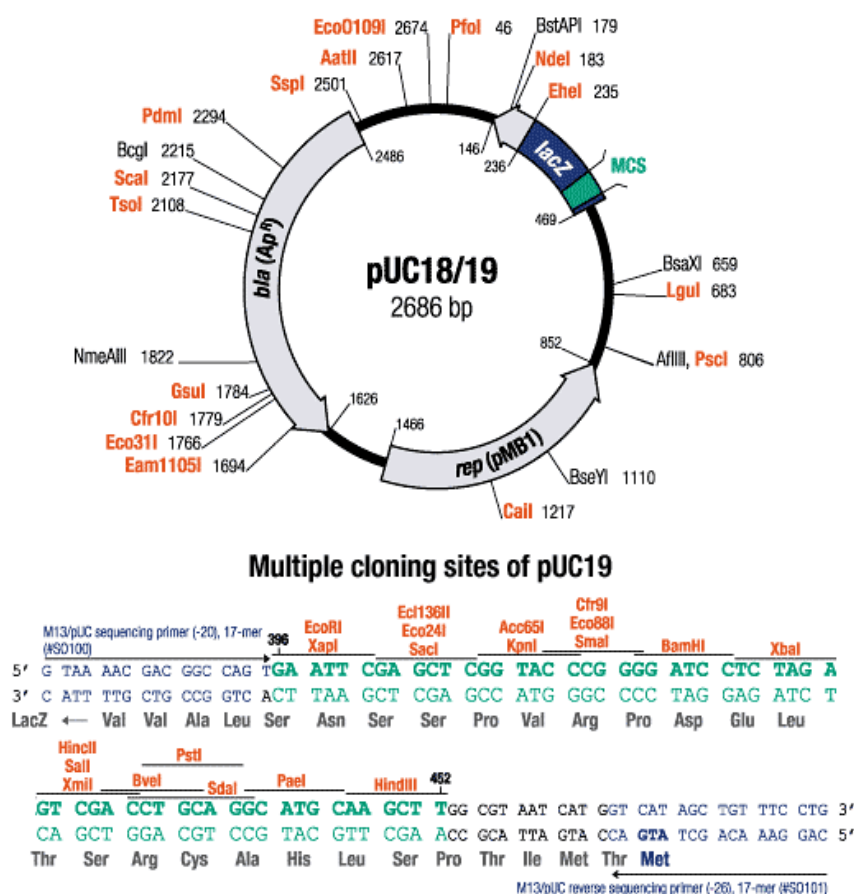


Figura 4. Esquema ilustrativo apresentando características do vetor PUC19DNA/SmaI (Fermentas).

O cálculo para concentração inserto: vetor foi realizado de acordo com o Technical Manual pGEM-T and pGEM-T Easy Vector Systems (Promega), conforme demonstrado abaixo:

$$\frac{\text{ng vetor} \cdot \text{kb inserto}}{\text{kb vetor}} \cdot \frac{3 \text{ inserto}}{1 \text{ vetor}} = \text{ng inserto}$$

Deste modo, a reação de ligação constituiu-se de 13 µL de DNA a 130 ng/ µL, reparado e fosforilado, 2 µL do Tampão T4 DNA Ligase 10X com 10

mM de ATP (BioLabs), 1,5 µL da enzima T4 DNA Ligase (BioLabs), 1 µL do vetor pUC19DNA/Sma a 500 ng/µL (Fermentas) e água deionizada estéril para completar o volume de 20 µL. A reação foi incubada a 16°C por 16 horas.

A enzima T4 DNA ligase catalisa o desenvolvimento de uma ligação fosfodiéster entre terminais 5'-fosfato e 3'-hidroxila de DNA dupla fita. A enzima repara o DNA dupla fita, uni os fragmentos com terminais coesivos ou sem corte, mas não possui nenhuma atividade em ácidos nucleicos fita única. Esta enzima necessita ATP como cofator.

4.5.7 Transformação bacteriana da célula competente

As transformações foram realizadas com células competentes de *Escherichia coli*, DH10β, cedidas pelo Prof. Dr. Manoel Victor Franco Lemos, responsável pelo Laboratório de Genética de Bactérias, do Departamento de Biologia Aplicada à Agropecuária da FCAV, UNESP, Campus de Jaboticabal.

As células competentes foram previamente retiradas do freezer a -80°C e descongeladas em banho de gelo por aproximadamente 5 minutos antes de serem usadas. Na reação foram utilizados 10 µL do DNA ligado, e 200 µL da célula competente DH10β. A transformação foi feita por meio de choque térmico e, para isso, a reação foi colocada por 20 minutos em banho de gelo e logo em seguida, submetida a 42°C, precisamente por 90 segundos, sendo recolocada no banho de gelo por mais 2 minutos.

Após a transformação, nos tubos falcon, foram adicionados 970 µL de meio SOC [2% de triptona (p/v), 0,5 % de extrato de levedura (p/v), 1 mL de NaCl 1 M, 0,25 mL de KCl 1 M, 1 mL de Mg²⁺ 2 M filtrado em membrana de poro 0,22 µm e 1 mL de Glicose 2 M filtrada sob as mesmas condições] para propiciar o crescimento das células competentes, sendo submetidas a agitação de 200 r.p.m. à 37°C, por uma hora e trinta minutos.

Após incubação das células transformadas, alíquotas de 100 µL da cultura foram distribuídas em placas de Petri contendo o meio LB sólido seletivo. Durante a preparação do meio, foram adicionados 50 mg/mL de

ampicilina e após a solidificação, a fim de promover a seleção das células transformadas, foram espalhados sobre o meio com o auxílio da alça de Drigalski estéril, 100 µL de IPTG 0,1M [2,4 % em água (p/v)] e 20 µL X-GAL [5-bromo-4-cloro-3-indolil-β-D-galactopiranosídeo, 50 mg dissolvidos em 1 mL de DMF (N,N'dimethyl-formamida)].

As placas de Petri com o meio LB e as células transformadas foram incubadas em estufa B.O.D. por 16 horas a 37°C.

4.5.8 Coleta e estoque dos clones transformados

Após o período de incubação em BOD, as colônias brancas (transformadas) foram coletadas com palitos de madeira esterilizados, um para cada subclone.

Os clones foram inoculados em placas estéreis de poliestireno (microplaca de 96 poços) contendo 120 µL de meio LB líquido suplementado com ampicilina, em uma concentração de 50 mg/mL, sendo então incubados em BOD a 37°C por 22 horas. Após esse período, foram adicionados 110 µL de glicerol 40% (v/v) estéril em cada poço da placa. Posteriormente, as placas foram seladas com adesivos e estocadas a -80°C.

4.6 Sequenciamento e análise do clone

O DNA obtido foi sequenciado em plataforma Illumina. Para avaliação de qualidade e montagem da sequência foi utilizado o pacote de programas Phred/ Phrap/Consed sendo que foram selecionadas somente sequências que apresentaram mais de 350 bases com qualidade Phred ≥ 20 , através do programa contGEN. As sequências foram montadas e o “contig” formado analisado, sendo as ORFs (open reading frames) identificadas com o programa ORF Finder (WHEELER et al., 2003) do banco de dados do NCBI - “*National Center for Biotechnology Information*”

(<http://www.ncbi.nlm.nih.gov/projects/gorf>), em que as sequências de aminoácidos foram comparadas com o banco de dados do NCBI através do algoritmo BLASTX (ALTSCHUL et al., 1997).

As sequências identificadas como possíveis genes codificadores de enzimas proteolíticas foram comparadas com a base de dados de Peptidases (MEROPS) (RAWLING et al., 2014). Foi realizado um alinhamento múltiplo global através do programa “BioEdit Sequence Alignment Editor v.7” (HALL, 1999) entre a sequência estudada e as similares presentes no banco de dado MEROPS, buscando os aminoácidos do sítio catalítico, domínios e motivos conservados próprios de cada família de proteases. O possível produto do gene codificador de protease foi analisado com a ferramenta ProtParam para estimar as características teóricas da proteína (web.expasy.org/protparam) (GASTEIGER et al., 2005) e a ferramenta SignalP 4.1 (PETERSEN et al., 2011) para identificar a possível presença de peptídeo sinal. As sequências de aminoácidos utilizadas para as análises filogenéticas e evolutivas, foram obtidas no MEROPS (RAWLING et al., 2014), e alinhadas com ClustalW2 (LARKIN et al., 2007). O alinhamento foi submetido ao MrBayes 3.2 (RONQUIST et al., 2011) e Mega 6.0 (TAMURA et al., 2013) para a estimativa do modelo evolutivo adequado para a construção da árvore filogenética usando o método máxima verossimilhança - “Maximum Likelihood” com “bootstrap” de 1000 réplicas, e a matriz de substituição de aminoácidos Whelan And Goldman (WAG) (ALTENHOFF et al., 2012) (com frequência e 5 categorias) usando todos os sítios do alinhamento.

4.7 Clonagem do gene codificador da proteína orf19 em vetor pET28a

4.7.1 Amplificação do gene de interesse

A partir da ORF19 identificada no clone PL3C3, um par de oligonucleotídeos iniciadores de síntese foi construído para permitir a clonagem do gene codificador da orf19 em vetor de expressão pET28a (Novagen,

Gibbstown, New Jersey, Estados Unidos da América), utilizando as enzimas de restrição NdeI e XhoI, ilustrados na Tabela 4.

Tabela 4. Sequências dos oligonucleotídeos e parâmetros utilizados para amplificação do gene codificador da ORF19.

Proteína	Enzima de Restrição	Oligonucleotídeos (Sequência 3' - 5')	TM*	Tamanho do fragmento
ORF19	Forward	GTAACATATGCCCATGACCGACAGAA	58°C	966pb
	NdeI			
	Reverse	GGACTCGAGCGATCAGCCCGTAA	61°C	
	XhoI			

*TM = Temperatura de pareamento

4.7.2 Reações de PCR

No intuito de definir a temperatura de pareamento, o gene da orf19 foi amplificado com reações de PCR utilizando os oligonucleotídeos específicos, nas seguintes condições: 2 µL de Tampão 10X [Tris-HCl 200 mM, pH 8,8; 100 mM (NH₄)₂SO₄; 100 mM KCl; 1% (v/v) Triton X-100, 1 mg/mL BSA (Fermentas, Burlington, Ontário, Canadá)]; 0,4 µL de dNTP (4 mM); 0,6 µL de MgCl₂ (16mM); 100 pmoles de cada oligonucleotídeo (forward e reverse); 1U de Taq Polimerase; 60 ng de DNA fomial do clone PL3C3; e água Milli-Q ultra pura para totalizar o volume final de reação de 20 µL. As reações foram submetidas ao termociclador de gradiente (Biometra, Göttingen, Niedersachsen, Alemanha) de acordo com o programa: 1 passo de 95°C por 1 minutos, 30 ciclos a 95 °C por 2 minutos, 54°C – 66 °C por 1 minuto, 72 °C por 1 minutos, e um passo final de 72°C por 5 minutos. Com intuito de abranger todo o gradiente do termociclador foram realizadas doze repetições, que tiveram uma diferença de 1°C de uma amostra para outra na temperatura de pareamento.

Definida a temperatura de pareamento, reações de PCR foram feitas com a *Pfu* DNA Polimerase, sendo que esta enzima apresenta menor taxa de erro nas amplificações. Assim, para as reações com a enzima *Pfu* DNA

Polimerase foram realizadas as seguintes condições: 2 µL de Tampão 10X [200 mM Tris-HCl (pH 8,8), 100 mM (NH₄)₂SO₄, 100 mM KCl, 1 mg/mL BSA, 1% Triton X-100 (v/v)] 1,2 µL de MgSO₄ (25 mM, Thermo Scientific); 0,4 µL de dNTP (16mM); 60 ng de DNA fosmidial do clone CRP16D11; 100 pmoles de cada oligonucleotídeo (forward e reverse); 1 µL da *Pfu* DNA Polimerase (2,5 U/mL, Thermo Scientific); e água ultra pura para totalizar o volume final de reação de 20 µL. As reações foram submetidas ao programa descrito acima, no entanto, com temperatura de pareamento de 65 °C para a ORF19.

Uma alíquota de 2 µL das reações de PCR foi aplicada em gel de agarose 1% contendo brometo de etídeo (0,5 mg/µL), conforme descrito no item 4.3, a fim de visualizar a amplificação dos produtos de PCR.

4.7.3 Digestão das Extremidades

Os produtos de PCR purificados foram submetidos à dupla digestão com as enzimas de restrição *Nde*I e *Xho*I.

O produto de PCR da ORF19 foi submetido à seguinte reação: 8 µL do produto de PCR [125,5 ng/µL]; 1 µL da *Nde*I FastDigest (Thermo Scientific); 2 µL da *Xho*I FastDigest (Thermo Scientific); 2 µL de Tampão 10X FastDigest (Thermo Scientific); e água Milli-Q filtrada para completar 20 µL.

A reação de digestão foi submetida a 37°C por 2 horas e a 80 °C por 20 minutos para inativação da enzima. Posteriormente uma alíquota de 3 µL das digestões foi analisada em gel de agarose 1% contendo brometo de etídeo (0,5 mg/µL), conforme descrito no item 4.3.

Após a digestão e análise em gel de agarose, o volume restante da digestão foi purificado com o Kit DNA Clean & ConcentratorTM (Epigenetics Company), seguindo instruções do fabricante, sendo posteriormente quantificado no NanoDropTM 1000 Spectrophotometer (Thermo Fisher Scientific).

4.7.4 Reação de Ligação do inserto ao vetor

4.7.5 Transformação da Célula de *Escherichia coli*

O produto da reação de ligação foi utilizado para transformação da linhagem celular *Escherichia coli* BL21(D3) cedida pelo Prof. Dr. Manoel Victor Franco Lemos. Na reação foram utilizado 10 µL do DNA ligado molecularmente ao vetor e 200 µL da célula competente BL21(D3). A transformação foi realizada através de choque térmico, assim, a reação permaneceu 5 minutos no gelo, e logo em seguida foi submetida a 42 °C por 90 segundos e recolocada no gelo por 2 minutos. Após a transformação bacteriana foi adicionado 970 µL de meio SOC (Tryptona, Extrato de levedura, NaCl 1M, KCl 1M, Mg²⁺ 2M e Glicose 2M) e o cultivo foi conduzido à agitação de 250 rpm a 37 °C, por 1 hora e 30 minutos. Após este período, uma alíquota de 110 µL da cultura foi distribuída, com alça de Drigalski esterelizada, em placas de Petri contendo meio LB com 50 µg/mL de canamicina. As placas foram incubadas por 22 horas a 37 °C.

4.7.6 Coleta e estoque dos clones

Foram coletados aleatoriamente dez clones da clonagem pET28a-*orf19*. A coleta foi feita com palitos de madeira e inseridos em tubo do tipo falcon contendo 5 mL de meio LB com 50 µg /mL canamicina. Os tubos foram incubados sob agitação constante de 250 rpm, 37 °C, por 20 horas. Após este período, uma alíquota de 800 µL da cultura foi transferida para tubo criogênico contendo 200 µL de glicerol 40% esterelizado. Os tubos foram estocados a -80 °C.

A extração foi feita com o kit Wizard® Plus SV Minipreps DNA Purification System (Promega) seguindo as orientações do fabricante e quantificado em gel de agarose 0,8% conforme item 4.3.

4.7.7 Confirmação da clonagem e sequenciamento

A confirmação da clonagem foi feita através da reação de PCR em colônias. Os clones coletados com palitos de madeira foram depositados em microtubos estéreis e adicionados 20 µL de tampão de lise celular, sendo submetidos a 99°C por 30 minutos. Após a lise celular, 4 µL do sobrenadante foi utilizado para a reação de PCR utilizando os oligonucleotídeos iniciadores específicos, descritos no item 6.2 e a visualização do resultado foi feita em gel de eletroforese à 1%, ver item 4.3.

Os fragmentos de DNA plasmidial foram sequenciados em microplacas nas seguintes condições: 100 ng de DNA plasmidial, 3,0 µL de tampão de sequenciamento (Save money 2,5X), 1,0 µL de BigDye v3.1 (Applied Biosystems Carlsbad, Califórnia, Estados Unidos da América), 10 pmoles do oligonucleotídeo sendo que as reações foram feitas separadamente para *Forward* e *Reverse*, e água deionizada estéril para completar o volume de 10 µL.

As placas foram submetidas ao termociclador, com os seguintes ciclos: desnaturação a 96°C por 1 minuto; 39 ciclos de 96°C por 15 segundos; 65°C por 15 segundos, 60°C por 4 minutos e ciclo final a 4 °C até serem retiradas.

Após a reação, as amostras serão preparadas para o sequenciamento. O DNA amplificado foi precipitado com 80 µL de isopropanol 75% adicionados a cada reação. As placas foram incubadas em temperatura ambiente por 15 minutos, sendo posteriormente centrifugadas a 3040 x g, por 30 min. a 20 °C. Após precipitação do DNA, o sobrenadante foi descartado e as amostras foram lavadas duas vezes com 200 µL de etanol 70% seguido de centrifugação por 10 minutos na mesma temperatura e força centrífuga descritas anteriormente. O sobrenadante foi descartado, sendo que o excesso de etanol foi retirado por inversão da placa em papel absorvente. Para secar o líquido restante das placas, as amostras foram secas a vácuo por 5 minutos em aparelho com função dessecador para soluções alcoólicas, a 45°C.

As amostras foram ressuspensas em 10 µL de formamida, a fim de obter o DNA em fita simples, desnaturadas por 5 minutos a 95 °C e submetidas ao sequenciador automático ABI 3730 XL DNA Analyzer (Applied Biosystems,

Foster City, California) seguindo recomendações do fabricante. Após o sequenciamento, os eletroferogramas foram analisados conforme descrito no item 4.6.

4.8 Ensaios de expressão e extração da proteína ORF19

Foram realizados testes de expressão da proteína recombinante pET28a-*orf19*, na célula hospedeira *Escherichia coli* BL21(D3). Para isso uma colônia foi isolada em 50 mL de meio LB com canamicina a 50 µg/mL que permaneceu sob agitação por 18 horas, 37°C, 230 r.p.m. Uma alíquota de 1 mL foi inoculada em frasco contendo 100 mL de meio LB com canamicina 50 µg/mL, que permaneceu sob agitação a 200 rpm, 37 °C, até atingir a fase logarítmica de crescimento, com D.O de 600nm entre 0,4 e 0,6. A indução foi feita adicionando 0,1 mM de IPTG (isopropil-β-D-tiogalactopiranosídeo) à cultura, que permaneceu nas condições descritas por 6 horas.

A análise de expressão da proteína de interesse em comparação às proteínas totais, foi realizada através da coleta de 1mL da cultura no tempo zero (T 0 -antes da indução), no tempo 2 (T2 – 2 horas após a indução), no tempo 4 (T4 – 4 horas após a indução) e no tempo 6 (T6 - 6 horas após a indução) e centrifugadas por 10 min a 14000 rpm. Após confirmar a expressão, condições diferentes foram testadas, com variações na concentração de IPTG, temperatura, e tempo de incubação. Determinada a melhor condição para a expressão da proteína, ensaios em volumes maiores foram realizados, seguindo o procedimento descrito anteriormente.

Outras condições testadas foram a 25 e 30°C, 200 rpm e após atingir fase logarítmica de crescimento, com D.O de 600nm entre 0,4 e 0,6 a indução foi feita adicionando 0,4 – 0,8 – e 1,0 mM de IPTG à cultura, que permaneceu nas condições descritas por 20 horas. Alíquotas de 1 mL da cultura foram coletadas no tempo zero (T0 - antes da indução), no tempo 20 (T20 – 20 horas após a indução) e centrifugadas para posterior análise da expressão.

Para a extração da proteína, o meio de cultura foi centrifugado a 3624 x g por 20 minutos a 4°C. O *pellet* celular foi ressuspendido em 10 mL do tampão

de lise por litro de cultura [50 mM Tris-HCl pH 7,0, 100 mM de NaCl, 10 % (v/v) Glicerol] acrescido de 1 mg/mL de lisozima. As amostras permaneceram em gelo por 1 hora, e posteriormente submetidas à lise com sonicador em aparelho Branson Sonifier 250, com 20% da razão cíclica por até 15 minutos com intervalos de 20 segundos. As amostras foram centrifugadas por 20 minutos, a 4°C, 10000 rpm, para separação de extratos solúveis e insolúveis.

4.8.1 Eletroforese em gel de poliacrilamida SDS-PAGE

Através da eletroforese desnaturante em gel de poliacrilamida 12% - SDS-PAGE (LAEMMLI, 1970), foi realizada a análise da expressão das proteínas utilizando o sistema mini-protean, e documentada sob luz UV em aparelho fotodocumentador Gel Doc 1000 (Bio-Rad, Estados Unidos da América). A corrida foi realizada a 90 V, com o tampão de corrida Tris-Glicina [50 mM de Tris-HCl; 150 mM de Glicina e 0,1% de SDS (p/v)]. O preparo das amostras foi realizado com a adição de 20 µL do tampão de amostra 2X e fervura por 5 minutos a 100 °C e o gel foi corado com “Coomassie Brilliant Blue” [40% de metanol (v/v); 10% de ácido acético glacial (v/v) e “Coomassie” R-250 a 0,1% (p/v)] e descorados com a solução 1 [10% de ácido acético (v/v) e 45% de metanol (v/v)] e solução 2 [8% de ácido acético (v/v) e 21% de etanol (v/v)].

4.9 Zimograma

O zimograma foi utilizado para verificar a atividade enzimática perante o consumo de substrato, adaptando a metodologia proposta por Heussen e Dowdle (1980). As amostras foram aplicadas em gel de poliacrilamida SDS-PAGE 10% contendo 1 mg/ mL de albumina. O preparo da amostra foi realizado com a adição de 10 µL do tampão de amostra [0,025% de azul de

bromofenol (p/v) e 50% de glicerol (p/v)] isento de β -mercaptoetanol. A amostra aplicada no gel não foi previamente fervida. Após a eletroforese o SDS foi removido com agitação suave, sendo lavado com Tris-HCl (50 mM, pH 7,0) contendo Triton X-100 2,5% por 30 minutos duas vezes. Em seguida, o gel foi incubado com Tris-HCl (50 mM, pH 7,0) a 37°C durante 20 horas. Posteriormente o gel foi corado com Commassie blue e descorados com solução 1 e 2 conforme descrito no item 4.8.1.

5. RESULTADOS E DISCUSSÃO

5.1 Prospeção e medição de atividade proteolítica

O teste bioquímico *in vitro* dos 4224 clones da biblioteca metagenômica do consórcio microbiano degradador de óleo diesel foi realizado em placas de petri contendo meio LB com *skim milk* como substrato. A evolução dos clones foi observada diariamente com o objetivo de verificar a formação de halos ao redor das colônias.

Foram identificados 26 clones que apresentaram uma possível atividade proteolítica com base na formação de halos claros (Figura 6), em meio Luria-Bertani suplementado com *Skim-Milk*. O clone CRP16D11 foi selecionado para ser subclonado.

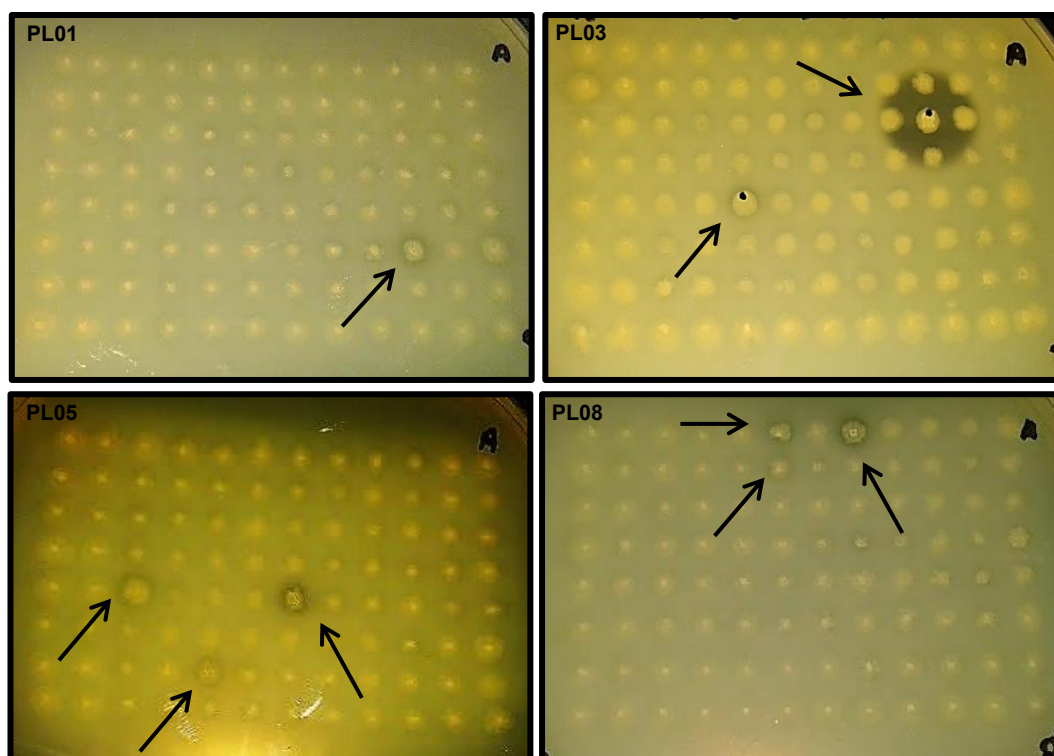


Figura 6. Análise *in vitro* da biblioteca metagenômica em meio de cultura LB, suplementado com 1% de skim milk (p/v) e 12,5 ug/mL de clorafenicol (v/v) apresentando possível atividade proteolítica (seta) no terceiro dia de cultivo.

No entanto, nem todos os clones que apresentam formação de halos possuem atividades proteolíticas. Em alguns casos a aparição de halo é devido à produção de ácido a partir da degradação de monossacarídeos (JONES et al., 2007). Este meio de cultura pode ser apropriado para a identificação de clones degradadores de lactose e galactose, porém não é uma técnica segura para identificar atividades proteolíticas, mas facilita a prospecção de clones que apresentam tais atividades em bibliotecas metagenômica (MORRIS et al., 2012).

Desta maneira, houve a necessidade de uma metodologia complementar para a identificação dos clones produtores de proteases. Jones et al., (2007), propôs a azocaseína como substrato genérico para a medição da atividade proteolítica. Dos 26 clones avaliados, apenas 11 clones (43,7%) apresentaram atividade proteolítica com azocaseína como substrato (Figura 7). Os demais clones (57,4%) não apresentaram atividade relevante quando comparada com a atividade da *E. coli* BL21 não transformada (clones negativos ou falsos positivos). O Clone PL3C3 foi o que apresentou uma maior atividade e por isso foi selecionado para realizar as análises genéticas posteriores.

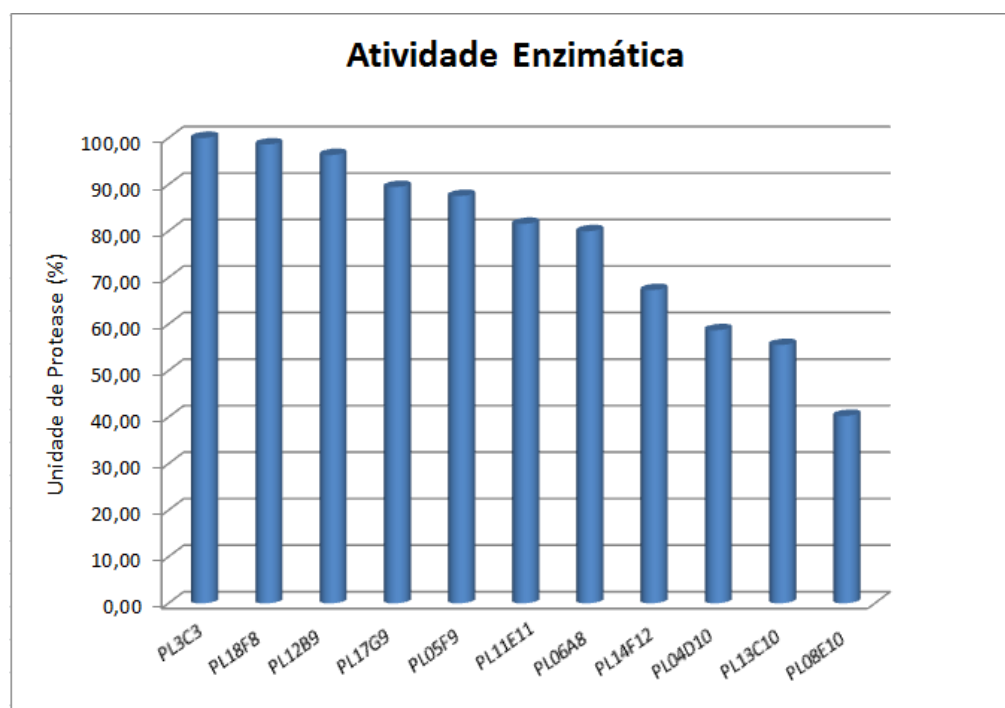


Figura 7. Estimativa da atividade proteolítica. Amostras do sobrenadante dos clones selecionados com base na formação de halos foram incubadas com azocaseína a 37°C, 4h. Clones com 0% da atividade não foram representados no gráfico.

5.2 Extração do DNA fosmidial

Mediante a atividade proteolítica com azocaseína como substrato, onze clones tiveram seu DNA fosmidial extraído. Após a extração, o mesmo foi submetido ao gel de agarose 0,8% podendo ser analisado na Figura 8. Os valores da concentração e purezas seguem descritos na tabela 5.

Tabela 5. Quantificação de DNA em NanoDrop™ 1000 Spectrophotometer dos clones que apresentaram atividade proteolítica mediante atividade com azocaseína como substrato.

	Amostra	ng/μL	260/280 nm
1	PL3C3	208,0	1,92
2	PL18F8	140,5	1,83
3	PL12B9	102,4	1,80
4	PL17G9	178,5	1,90
5	PL05F9	163,2	1,86
6	PL11E11	112,3	1,72
7	PL06A8	178,4	1,88
8	PL14F12	109,6	1,79
9	PL04D10	139,1	1,85
10	PL13C10	100,2	1,72
11	PL08E10	184,0	1,90

Pode-se verificar que o DNA não estava fragmentado e não apresentava impurezas, apontando a ausência de contaminação por proteínas.

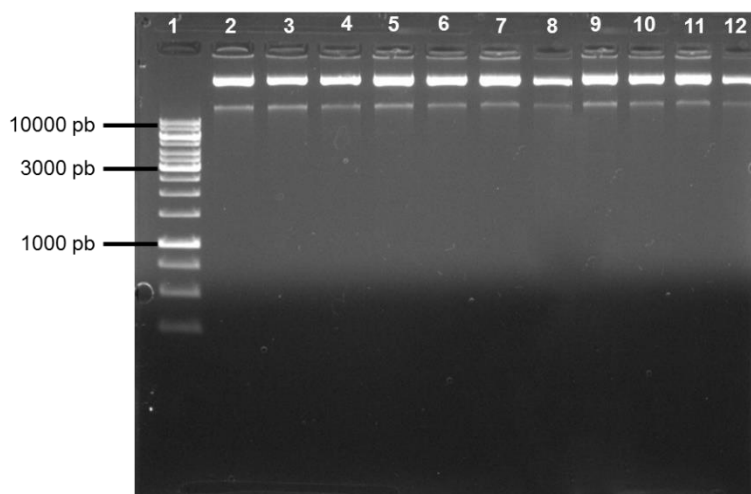


Figura 8. Perfil eletroforético em gel de agarose 0,8% da extração do DNA fosmidial do clone selecionado. 1 - Marcador de tamanho molecular 1kb DNA Ladder (Fermentas); 2 - DNA extraído do clone PL3C3. 3 - DNA extraído do clone PL18F8; 4 - DNA extraído do clone PL12B9; 5 - DNA extraído do clone PL17G9; 6 - DNA extraído do clone PL05F9; 7 - DNA extraído do clone PL11E11; 8 - DNA extraído do clone PL06A8; 9 - DNA extraído do clone PL14F12; 10 - DNA extraído do clone PL04D10; 11 - DNA extraído do clone PL13C10; 12 - DNA extraído do clone PL08E10.

5.3 Digestão do DNA fosmidial

A digestão fosmidial dos clones com as enzimas de restrição *Alu1*, *Pst1* e *EcoR1*, permitiram afirmar que os clones são distintos entre si, visto que apresentaram padrões diferentes de restrição, possibilitando calcular o tamanho aproximado do inserto clonado no vetor, sendo de aproximadamente 30 kb para o clone PL3C3 (Figura 9).

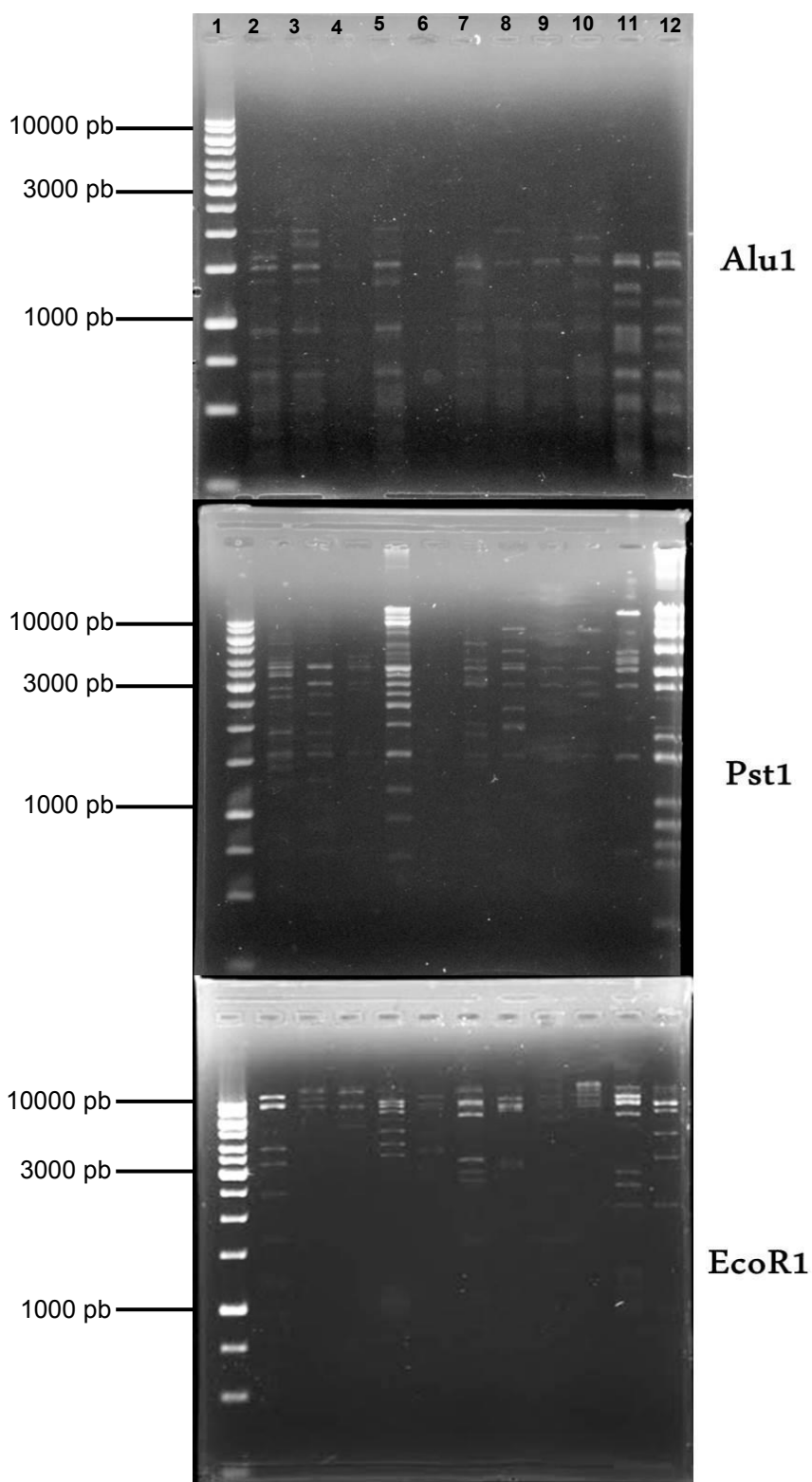


Figura 9. Perfil eletroforético em gel de agarose 0,8% da restrição do DNA fosmidial dos clones. 1 - Marcador de tamanho molecular 1kb DNA Ladder (Fermentas); 2 - clone PL3C3. 3 - clone PL18F8; 4 - clone PL12B9; 5 - clone PL17G9; 6 - clone PL05F9; 7 - clone PL11E11; 8 - clone PL06A8; 9 - clone PL14F12; 10 - clone PL04D10; 11 - clone PL13C10; 12 - clone PL08E10.

5.4 Construção da sub-biblioteca

Foram selecionados fragmentos do clone PL3C3, variando seu tamanho de 1 a 3 kb. A partir da análise em eletroforese em gel de agarose, uma alíquota de cada material nebulizador, foi possível verificar que a nebulização foi bem sucedida, visto que os fragmentos gerados estavam mais concentrados na região de interesse (Figura 10). A concentração do material nebulizador foi de 269,7 ng/ μ L, aproximadamente, de acordo com a relação ng/ μ L do marcador de tamanho molecular utilizado 1 kb DNA Ladder (Fermentas).

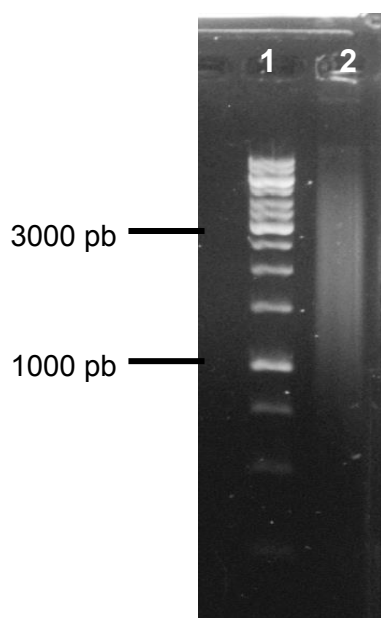


Figura 10. Perfil eletroforético da nebulização do clone PL3C3 em gel de agarose 0,8%. 1 - Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas). 2 - DNA nebulizado do clone PL3C3.

A seleção da região no gel de agarose de baixo ponto de fusão foi bem sucedida (Figura 11) e os fragmentos de 1 a 3 kb, já reparados e fosforilados, puderam ser selecionados. Após a recuperação dos fragmentos, foi realizada a leitura de concentração por espectrofotometria, resultando em 129,3 ng/ μ L.

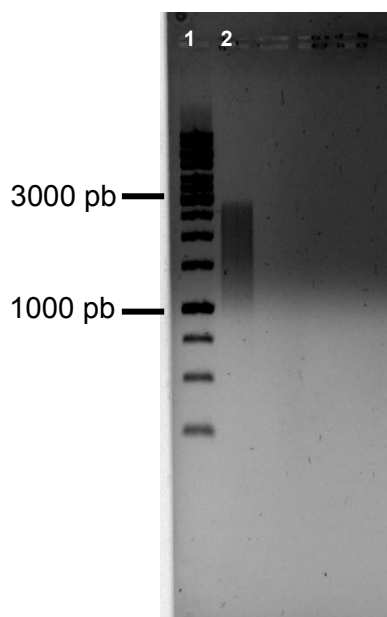


Figura 11. Perfil eletroforético do resultado da eluição do clone PL3C3, utilizando Gelase, em gel de agarose 0,8%. 1 - Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2 - DNA fosmidial eluído do clone PL3C3.

Os fragmentos de DNA fosmidial recuperado da eluição, foi inserido em vetor pUC19 desfosforilado e digerido com *Sma*I. Posteriormente, os plasmídeos recombinantes foram inseridos por transformação em células competentes de *Escherichia coli* DH10 β . Foram coletados 768 subclones, ou seja, oito placas de poliestireno (microplacas de 96 poços).

5.5 Análise da sequência

As sequências foram analisadas com o auxílio de programas do pacote Phred/ Phrap/ Consed e através do programa ContGEN foi possível construir um “contig” de 29492 pb.

Utilizando a ferramenta BlastN (nucleotide collection nr/nt), o resultado foi de 83% de similaridade da sequência de nucleotídeos com o micro-organismo *Parvibaculum lavamentivorans* (número de acesso CP000774.1), 39% de *query coverage* e 0.0 de E-value, seguido de *Mesorhizobium huakuii*

(número de acesso CP006581.1), com 75% de similaridade, 8% de *query coverage* e 0.0 de E-value, conforme demonstrado na Tabela 6.

Tabela 6. "Blast" do *contig* completo do clone PL3C3 com as sequências do NCBI, através da ferramenta BLASTN. Foram colocados os quatro primeiros resultados fornecidos pelo NCBI.

Organismo	Número de acesso	Identidade (%)	Query coverage (%)	E-value
<i>Parvibaculum lavamentivorans</i>	CP000774. 1	83	39	0.0
<i>Mesorhizobium huakuii</i>	CP006581. 1	75	8	0.0
<i>Mesorhizobium loti</i>	BA000012. 4	75	10	0.0
<i>Caulobacter crescentus</i>	CP001340. 1	81	14	0.0

A compreensão do resultado indica que o *contig* obtido do clone PL3C3 ainda não foi estudado, visto que o mesmo apresentou pequenas regiões com similaridade. A possível presença deste organismo no consórcio microbiano degradador de óleo diesel, corrobora com outros estudos que identificaram este gênero em solos contaminados com hidrocarbonetos (WUNSCH et al., 1995; HIGASHIIOKA et al., 2009; PIRÔLLO et al., 2008). De acordo com Paixão et al. (2010) o micro-organismo *Parvibaculum lavamentivorans* foi detectado na mesma biblioteca do consórcio microbiano degradador de óleo diesel. Este gênero *Parvibaculum* tem a capacidade de usar os biosurfactantes produzidos por outros micro-organismos como fonte de carbono e energia (SCHLEHECK et al., 2003), permitindo que o mesmo se desenvolva em ambientes contaminados com hidrocarbonetos.

Após a análise *contig* completo, foi feita a predição das ORFs com o programa ORF Finder (Open Reading Finder) do NCBI. O programa gerou 156

ORFs e todas foram analisadas, principalmente quanto a *Max Score*, *Query coverage* e *E-value*. Através da análise foi possível selecionar 29 ORFs, ou possíveis genes para anotação, determinando o início e término da ORF, localizar a frame de leitura, aferir o tamanho em nucleotídeos e aminoácidos, prever a sua função biológica e identificar o possível micro-organismo (Tabela 7). Todas as ORFs foram representadas no mapa físico para o *contig* ilustrado na Figura 12.

Tabela 7. Predição das ORFs identificadas no *contig* do clone PL3C3 através do programa ORF Finder a partir do NCBI

ORF N.	ORF INÍCIO-FINAL	ORF (pb)	ORF (aa)	Descrição	Micro-organismo/número de acesso	Identidade. (%)	E-value
1	309 – 1025	717	238	Regulador de transcrição putativa – família MerR	<i>Parvibaculum lavamentivorans</i> (ABS62084.1)	69	1e-62
2	1098 – 2882	1785	594	Acetil- CoA desidrogenase	<i>Parvibaculum lavamentivorans</i> (WP_011995376.1)	79	0.0
3	2947 – 4104	1158	385	Acetil- CoA desidrogenase	<i>Parvibaculum lavamentivorans</i> (WP_011995377.1)	82	0.0
4	4155 – 5360	1206	401	Acetil- CoA acetiltransferase	<i>Parvibaculum lavamentivorans</i> (WP_011995378.1)	90	0.0
5	5368 – 7626	2259	752	3-hidróxiacil-CoA desidrogenase	<i>Parvibaculum lavamentivorans</i> (WP_011995379.1)	86	0.0
6	7811 – 8434	624	207	Proteína hipotética	<i>Parvibaculum lavamentivorans</i> (WP_041535811.1)	48	3e-46
7	8436 – 9164	729	242	Proteína hipotética	<i>Sphingomonas changbaiensis</i> (WP_046349414.1)	51	5e-58
8	9278 – 9703	426	141	Fator de liberação da cadeia peptídica	<i>Parvibaculum lavamentivorans</i> (WP_011995548.1)	71	8e-66
9	10160 – 11254	1095	364	Glutathione – S Transferase	<i>Parvibaculum lavamentivorans</i> DS-1 (ABS62108.1)	65	6e-166
10	11804 – 13987	2184	727	Receptor putativo TonB-dependente	<i>Novosphingobium tardaugens</i> NBRC 16725 (GAD48813.1)	51	0.0
11	13987 – 14658	672	223	Transdutor de energia TonB – várias espécies	<i>Sphingomonadaceae</i> (WP_037446514.1)	48	2e-15
12	14845 – 15228	384	127	Enoil – CoA hidratase/isomerase	<i>Parvibaculum lavamentivorans</i> (WP_011995395.1)	67	5e-42
13	15268 – 16443	1176	391	alfa/beta hidrolase	<i>Parvibaculum lavamentivorans</i> (WP_011995394.1)	83	0.0
14	16479 – 17372	893	297	Proteína hipotética	<i>alpha proteobacterium</i> MA2 (WP_045446911.1)	64	3e-126

ORF N.	ORF INÍCIO-FINAL	ORF (pb)	ORF (aa)	Descrição	Micro-organismo/ número de acesso	Identidade. (%)	E-value
15	17463 – 17708	246	81	Proteína hipotética	<i>Thalassospira australica</i> - <i>Thalassospira</i> sp. NP3b2 (WP_033068967.1)	56	1e-15
16	(17705 - 18625)	(921)	307	Proteína hipotética	<i>Parvibaculum lavamentivorans</i> (ABS62100. 1)	71	7e-139
17	18667 – 19569	903	300	Proteína tipo tioesterase	<i>Parvibaculum lavamentivorans</i> (WP_011995390.1)	67	7e-145
18	19566 – 20039	474	157	Tioesterase	<i>Bradyrhizobium</i> sp. Cp5.3 (WP_027552344. 1)	39	2e-28
19	20048 – 21013	966	321	alfa/beta hidrolase	<i>Parvibaculum lavamentivorans</i> DS-1 (ABS62095. 1)	69	2e-161
20	21221 – 21842	621	268	Biossíntese de proteínas tiazole ThiJ	<i>Parvibaculum lavamentivorans</i> (WP_011995384. 1)	76	4e-109
21	21831 – 22874	1044	347	Regulador de transcrição, família AraC	<i>Parvibaculum lavamentivorans</i> DS-1 (ABS62091.1)	63	8e-139
22	22976 – 23833	858	285	3-alfa, 7-alfa, 12-alfa-tri-hidroxi-5-beta-colest-24-enoil-CoA hidratase	<i>Alpha Proteobacterium</i> MA2 (WP_045446940. 1)	81	1e-173
23	23867 – 24280	414	137	Proteína hipotética	<i>Fischerella muscicola</i> (WP_016867545.1)	52	2e-40
24	24406 – 24867	462	153	Superfamília proteína tioesterase	<i>Parvibaculum lavamentivorans</i> (ABS62114. 1)	65	2e-57
25	24883 – 25155	273	90	Proteína hipotética	<i>Chelatococcus</i> sp. GW1 (WP_026015080. 1)	44	1e-14
26	25723 – 26136	414	137	Superfamília proteína tioesterase	<i>Parvibaculum lavamentivorans</i> (ABS62115. 1)	65	1e-48
27	26426 – 27628	1203	532	Acetil-CoA-acetiltransferase	<i>Alpha Proteobacterium</i> Mf 1.05b.01 (WP_029640547. 1)	84	0.0
28	27836 – 28276	441	146	Desidratase	<i>Parvibaculum lavamentivorans</i> (WP_011995409.1)	88	8e-90
29	28415 – 29398	984	327	Acil-CoA desidrogenase	<i>Parvibaculum lavamentivorans</i> (WP_011995410. 1)	78	0.0

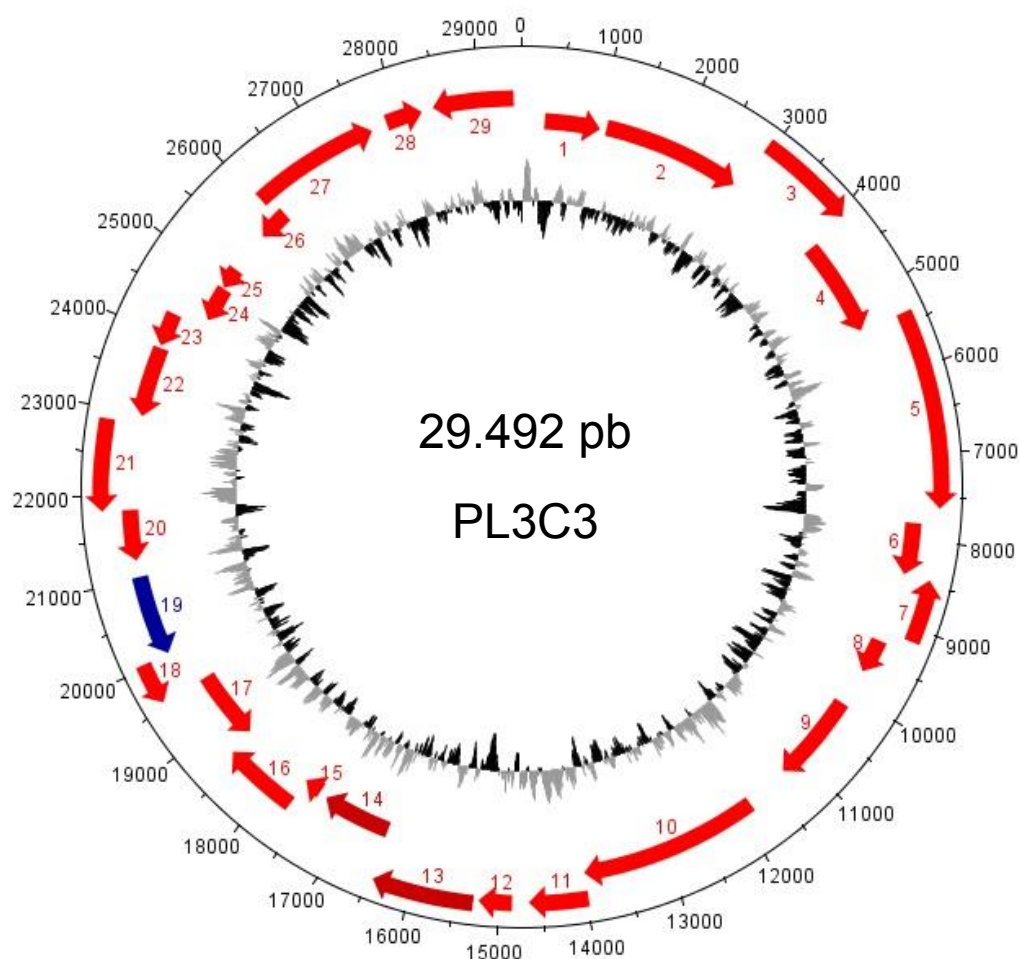


Figura 12. Mapa físico do DNA fosmidial do clone PL3C3 gerado pelo programa DNAPlotter (CAVER et al., 2009). A ORF19 está ilustrada como uma seta azul e é responsável pela atividade proteolítica do clone.

As ORFs identificadas como proteínas hipotéticas (6, 7, 14, 15, 16, 23 e 25) e alfa/beta hidrolases (13 e 19), foram adicionalmente analisadas com auxílio de outros programas e bases de dados (MEROPS) visando identificar domínios conservados que indicassem genes codificadores de proteases. Para as ORFs codificadoras de proteínas hipotéticas, não foram detectados genes codificadores de proteases.

A partir das análises averiguadas das ORFs através do programa PRODOM e a base de dados MEROPS (SERVANT et al., 2002; RAWLINGS et al., 2014), a ORF19 apresentou maiores indícios ao codificar uma protease putativa, composta por 966pb (321 aminoácidos), apresentando uma

similaridade de 74,38% com uma Prolil-oligopeptidase da família das Serino peptidase não classificada (S9U) de *Parvibaculum lavamentivorans* (Tabela 8).

Tabela 8. BLAST e análise da ORF19 através de banco de dados MEROPS - Sequência de peptidase.

Predição	Organismo	Número de acesso	Identidade (%)	E-value
Família peptidases S9 não atribuídas/classificadas	<i>Parvibaculum lavamentivorans</i>	MER095859	74.38	2.50e-98
Família peptidases S9 não atribuídas/classificadas	<i>Starkeya novela</i>	MER215046	59.51	4.00e-75
Família peptidases S9 não atribuídas/classificadas	<i>Methylobacterium Populi</i>	MER078431	55.56%	7.70e-72

5.5.1 Análise das relações filogenéticas e identificação dos sítios catalíticos

Após a predição das ORFs e localização das mesmas no mapa gênico, foram feitos alinhamentos usando o ClustalW através do programa BioEdit Sequence Alignment Editor.

As relações filogenéticas foram constituídas com o auxílio do programa Mega 6 utilizando sequências distintas que representam as famílias proteolíticas. As proteases são classificadas de acordo com o resíduo que apresenta em seu sítio catalítico (A, C, G, M, N, P, S, T, U) (ver tabela 2 e 3 – página 25), sendo que algumas famílias podem estar subdivididas em subfamílias.

Por meio da análise da árvore (Figura 13), é possível observar que a ORF19 encontra-se localizada na família das Serino – peptidases (S9). A família S9 apresenta três resíduos de aminoácidos conservados (Ser, Asp, His) que formam a tríade catalítica, e está subdividida em 5 subfamílias: S9A - Prolil oligopeptidase, S9B Dipeptidil peptidase, S9C – Acil-amino-acilo peptidase e S9D - Glutamilendopeptidase (RAWLINGS et al., 1991), e a subfamília S9U,

corresponde as proteases não atribuídas ou ainda não classificadas que apresentem estes três resíduos catalíticos. Os membros representativos da família S9 apresentam estrutura alfa/beta/alfa hidrolase e são semelhantes estruturalmente às acetil-colino-esterase, esterases e lipases (KISS et al., 2004). Assim por meio desta análise filogenética é possível observar que a ORF19 permaneceu em um único ramo, entre os membros da subfamília S9U, sugerindo que ela pertence a esta família.

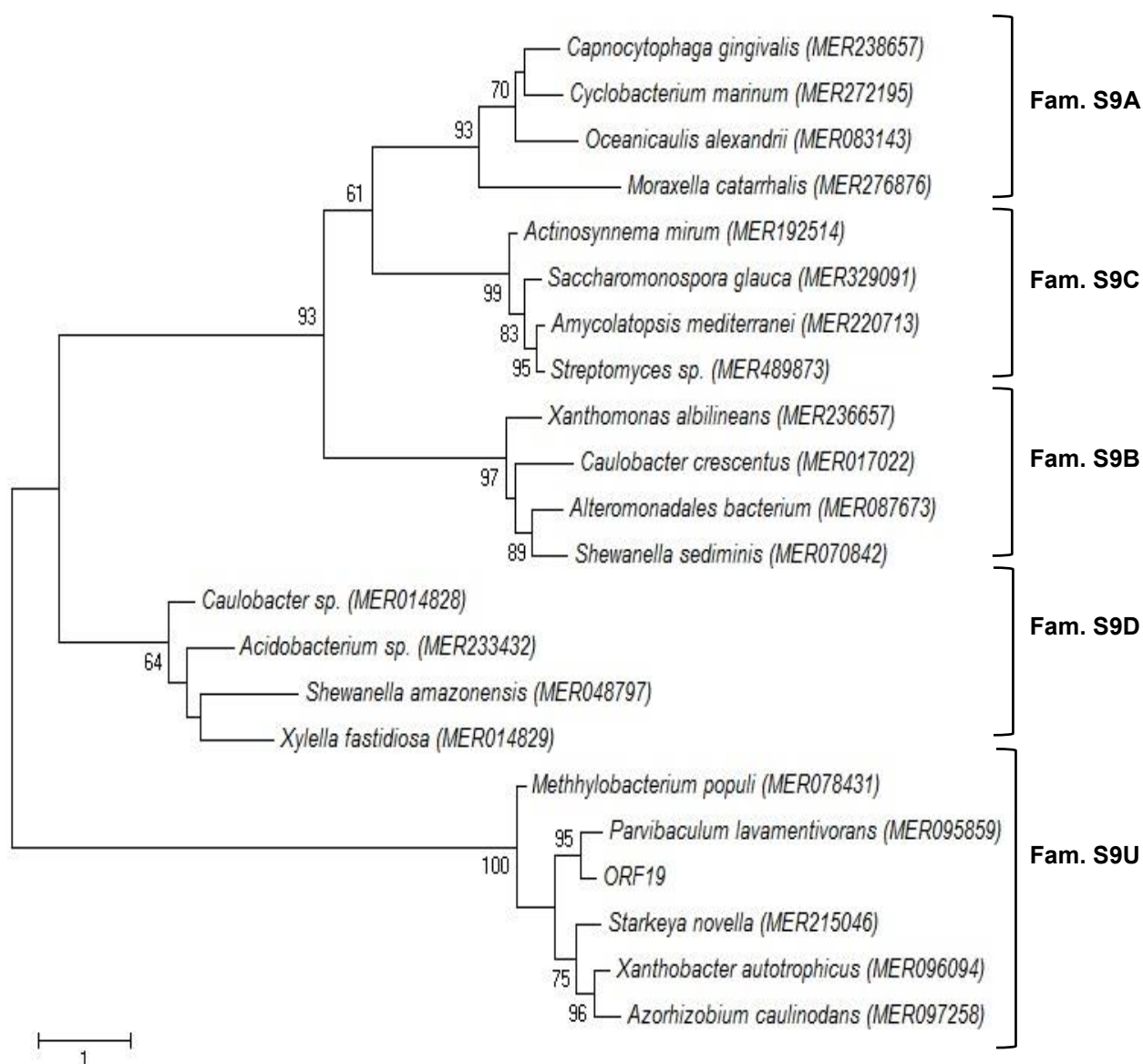


Figura 13. Árvore filogenética baseado em sequências de aminoácidos dos membros da família proteolítica S9 obtidos a partir do banco de dados MEROPS (números de acesso indicado), apresentando a relação filogenética de proteína ORF19, encontrado no contig do clone PL3C3. O valor estimado do parâmetro de forma para a distribuição gama discreta foi 4,9302. Padrão de substituição e as taxas foram estimados sob o modelo WAG (+ Gamma + Freq), com 5 categorias.

Os blocos conservados são localizados em torno do resíduo Serina, e é característico para cada subfamília: GG S XGGLL, em que X é tipicamente Asparagina ou Alanina (S9A), GW S YGGY (S9B), GG S YGG (S9C) e GGH S YGAFMT (S9D) (RAWLINGS et al., 2014). O alinhamento foi realizado com as sequências dos membros da família S9U, *Parvibaculum lavamentivorans* (número de acesso MER095859), *Starkeya novella* (número de acesso MER215046), *Methylobacterium populi* (número de acesso MER078431), *Xanthobacter autotrophicus* (número de acesso MER096094) e *Methylobacterium chloromethanicum* (número de acesso MER150404) mostrado na Figura 14. Através desta análise, foi possível identificar os aminoácidos que constituem o sítio ativo da proteína ORF19: Ser170, Asp265, His295. A partir deste alinhamento foram verificadas algumas modificações nos resíduos de aminoácidos. O resíduo de aminoácido S (serina) do motivo conservado GGSXGGLL foi alterado pelo glutamato (E) estando posicionado posteriormente. Outra modificação observada foi a substituição da leucina (L) pela histidina (H), sendo a sequência deste bloco GGESAGGHL.

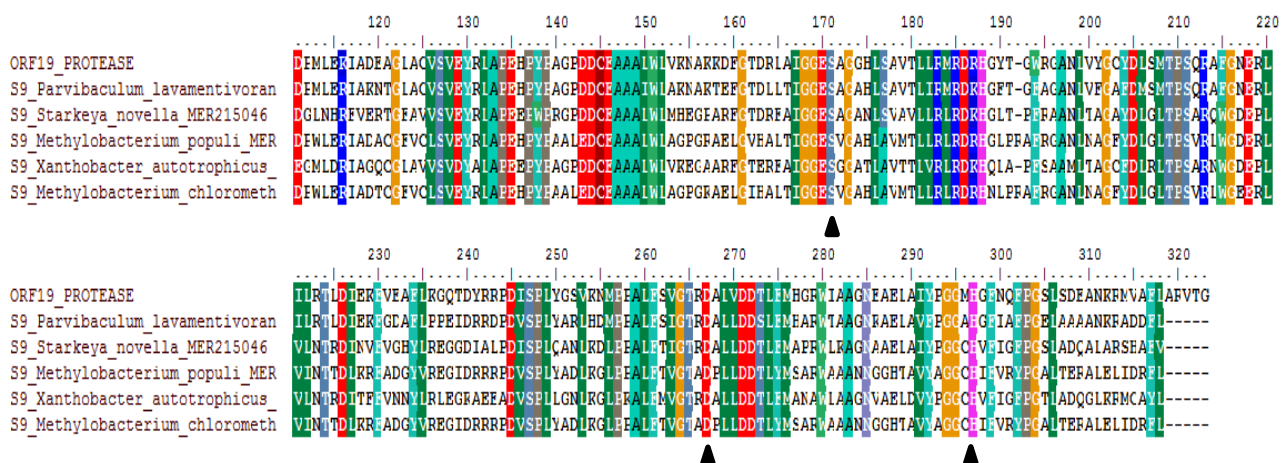


Figura 14. Alinhamento múltiplo de sequências de aminoácidos da ORF19 com cinco membros representativos da família S9U, representando as regiões conservadas de enzimas proteolíticas, a seta preta indica os resíduos de aminoácidos da tríade catalítica (▲).

5.6 Clonagem do gene em vetor de expressão pET28a

Mediante as análises filogenéticas, a ORF19 constituída de 321 aminoácidos, 966 nucleotídeos, foi clonada em vetor de expressão pET28a visando futura caracterização enzimática.

5.6.1 Amplificação do gene de interesse

A ORF19 foi submetida ao termociclador de gradiente (Biometra) com temperaturas variando de 54°C até 66°C.

A partir da análise eletroforética foi possível selecionar a temperatura de 65°C (Figura 15). Esta temperatura foi escolhida por propiciar a amplificação da região de interesse, aproximadamente 966 pb, e não apresentaram produtos de PCR inespecíficos.

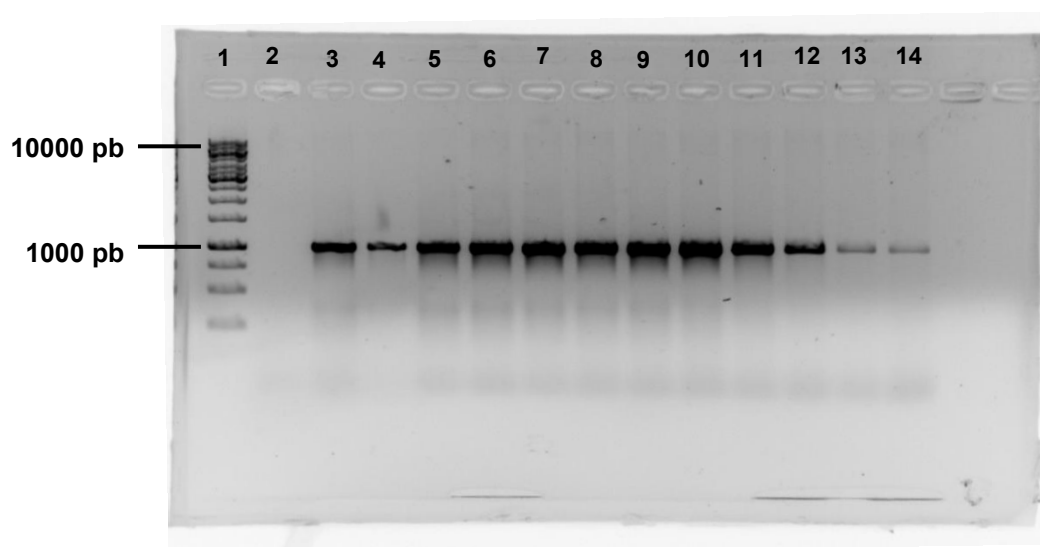


Figura 15. Perfil eletroforético em gel de agarose 1% mostrando padronização da temperatura de pareamento das reações de PCR em termociclador de gradiente (Biometra). 1 - Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 3 a 14 - Variação de temperatura de pareamento (54°C a 66°C).

Posterior à definição da melhor temperatura de pareamento, foi realizado a PCR com a enzima *Pfu* DNA Polimerase (Figura 16). Os produtos desta reação foram purificados com o kit Wizard Plus SV Gel and PCR Clean – Up System (Promega) e o resultado em gel de agarose confirmou que o produto de PCR estavam adequados para a clonagem em pET28a e sequenciamento.

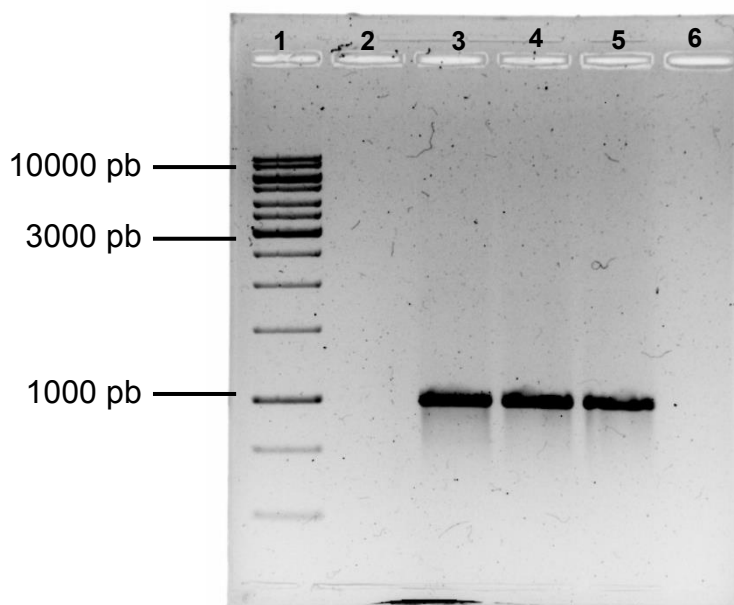


Figura 16. Perfil eletroforético em gel de agarose 1% da reação de PCR purificada com o Kit Wizard SV Plus Gel and PCR Clean-Up System (Promega). 1 - Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2 a 5 – Reação de PCR da ORF19 com a enzima *Pfu* DNA polimerase.

5.6.2 Construção pET28a-orf19 e transformação em células de *Escherichia coli* BL21(D3)

O produto de PCR purificado da ORF19 foi submetido à digestão enzimática com as enzimas de restrição XhoI e NdeI. Em seguida, as reações foram purificadas para eliminar possíveis restos de reação sendo que a quantificação no NanoDropTM 1000 Spectrophotometer permitiu verificar a concentração de 24,5 ng/μL e relação 260/280 de 1,53.

Após a purificação e digestão da ORF19, o produto foi inserido em vetor de expressão pET28a desfosforilado e digerido com XhoI e NdeI. Em seguida, os plasmídeos recombinantes foram transformados com o uso da célula competente *Escherichia coli* BL21(D3). Foram coletados dez clones que após o cultivo foram estocados a -80°C.

5.6.3 Confirmação da clonagem

Foi realizada PCR em colônias com os oligonucleotídeos específicos (ver item 3.7.1) dos dez clones coletados da construção pET28a-orf19 (Figura17).

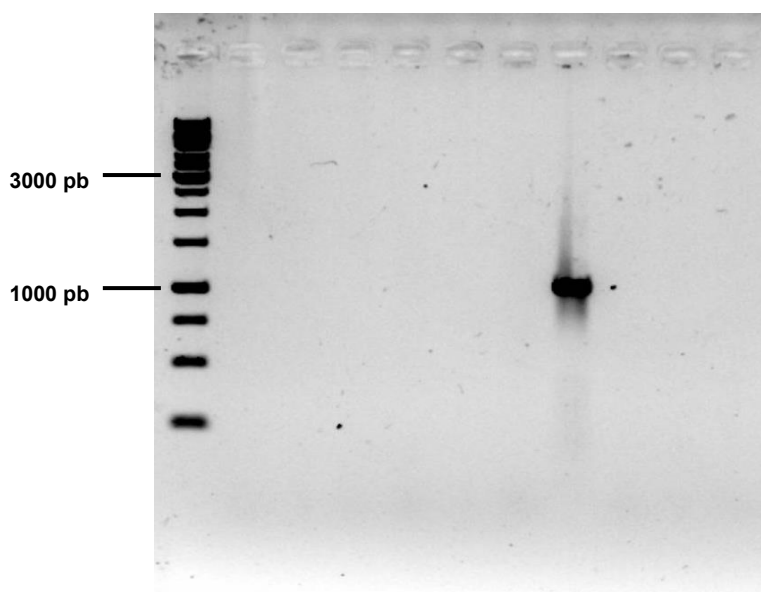


Figura 17. Perfil eletroforético em gel de agarose 1% para confirmar a presença do inserto referente à ORF19. 1 - Marcador molecular 1kb DNA Ladder (Fermentas); 2-10 - PCRs referentes aos clones coletados da transformação, ênfase para a canaleta 7 onde houve amplificação, confirmando a clonagem.

Através do perfil eletroforético em gel de agarose, dos dez clones coletados da construção pET28a – ORF19 apenas um apresentou amplificação da região de interesse (Figura 17). Os clones que não

amplificaram foram descartados. A confirmação da construção da ORF19 foi submetida ao sequenciamento e as sequências foram comparadas com o banco de dados do NCBI, sendo que a mesma apresentou 69% de identidade com alfa/beta hidrolase de *Parvibaculum lavamentivorans* (número de acesso ABS62095.1), permitindo concluir que a clonagem foi bem sucedida.

5.6.4 Expressão da proteína ORF19

Inicialmente foram observadas as características da ORF19 através da ferramenta ProtParam (web.expasy.org/protparam) (GASTEIGER et. al., 2005), sendo possível estimar o tamanho molecular e o ponto isoelétrico (pI), conforme os dados representados na Tabela 9.

Tabela 9. Características físico- químicas da proteína recombinante ORF19 obtidas a partir da ferramenta ProtParam

Análise ProtParam	ORF19
Número de aminoácidos	321
KDa	35.0
Coefficiente de extinção Abs de 0,1% (=1 g/ L)	1324
Ponto isoelétrico (pI)	6,67

A fim de verificar a expressão da proteína, testes foram realizados em diferentes temperaturas e concentrações de IPTG. Para a expressão recombinante de proteínas não existem fórmulas, sendo indispensável buscar o melhor protocolo para a proteína alvo. O teste inicial foi conduzido a 37°C, 0,1 mM de IPTG e agitação a 200 r.p.m. os testes de expressão, alíquotas de 1mL de extrato celular antes e após a indução (T0, T2 e T4) foram coletadas e centrifugadas a 10000 r.p.m. para serem aplicadas em gel de poliacrilamida 10%. Na figura 18, é possível analisar a expressão ao comparar as bandas do extrato celular antes e após a indução.

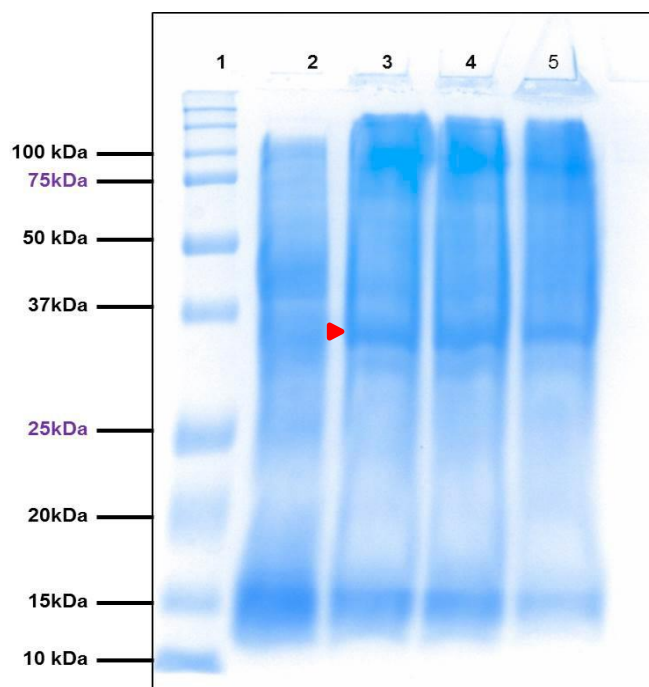


Figura 18. Teste de expressão. Eletroforese em gel de poliacrilamida 10%, corado com "Coomassie Brilliant Blue". 1 – Marcador Molecular Biorad Marcador de Peso Molecular Precision Plus Protein Standards Dual Color (Bio-Rad); 2 - Extrato total de células não induzidos (T0); 3 - Extrato total de células induzidas expressando a proteína ORF19 após 2 horas (T2); 4 - Extrato total de células induzidas expressando a proteína ORF19 após 4 horas (T4); 5 - Extrato total de células induzidas expressando a proteína ORF19 após 6 horas.

A partir dos resultados apresentados acima, houve a necessidade na variação da concentração de IPTG (0,4; 0,8 e 1.0 mM), temperatura (25 °C e 30°C) e agitação (150 r.p.m – 230 r.p.m) (Figura 19).

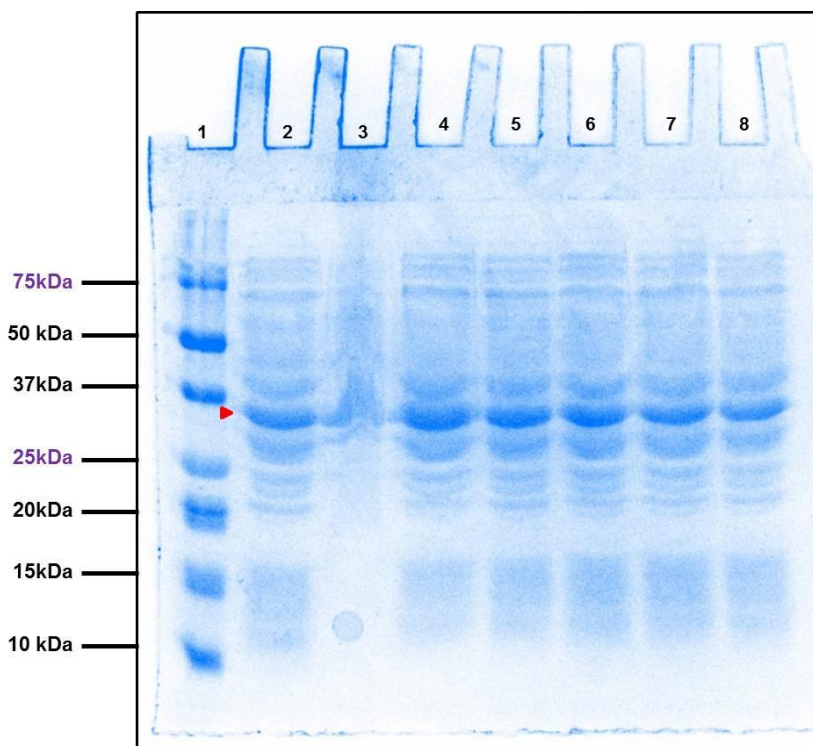


Figura 19. Teste de expressão. Eletroforese em gel de poliacrilamida 10%, corado com "Coomassie Brilliant Blue". 1 – Marcador Molecular Biorad Marcador de Peso Molecular Precision Plus Protein Standards Dual Color (Bio-Rad); 2 - Extrato total de células induzidas expressando a proteína ORF19 após 4 horas (T4 – 0,4mM IPTG, 25°C); 3 - Extrato total de células não induzidos (T0); 4 Extrato total de células induzidas expressando a proteína ORF19 após 4 horas (T4 – 0,4mM IPTG, 30°C); 5 - Extrato total de células induzidas expressando a proteína ORF19 após 4 horas (T4 – 0,8mM IPTG, 25°C); 6 - Extrato total de células induzidas expressando a proteína ORF19 após 4 horas (T4 – 0,8mM IPTG, 30°C); 7 - Extrato total de células induzidas expressando a proteína ORF19 após 4 horas (T4 – 1M IPTG, 25°C); Extrato total de células induzidas expressando a proteína ORF19 após 4 horas (T4 – 1M IPTG, 30°C).

Através dos testes de expressão, a proteína foi expressa em 1 litro de meio de cultura a 30°C, 0,4 mM de IPTG e agitação a 230 rpm utilizando o tempo de indução de 15 horas. Posteriormente, conduziu-se a extração da proteína com o intuito de tê-la na fração solúvel. O tamanho molecular da ORF19 recombinante equivalente a um polipeptídeo foi estimado através da análise de regressão linear obtida para os valores equivalentes ao log do tamanho molecular de proteínas do padrão correspondente à mobilidade relativa dos mesmos (SDS-PAGE), permitindo calcular o tamanho da proteína

através da equação da reta obtida. A proteína apresentou tamanho molecular de 35,19 kDa (Figura 20, canaleta 5), o que corresponde com a previsão da análise ProtParam.

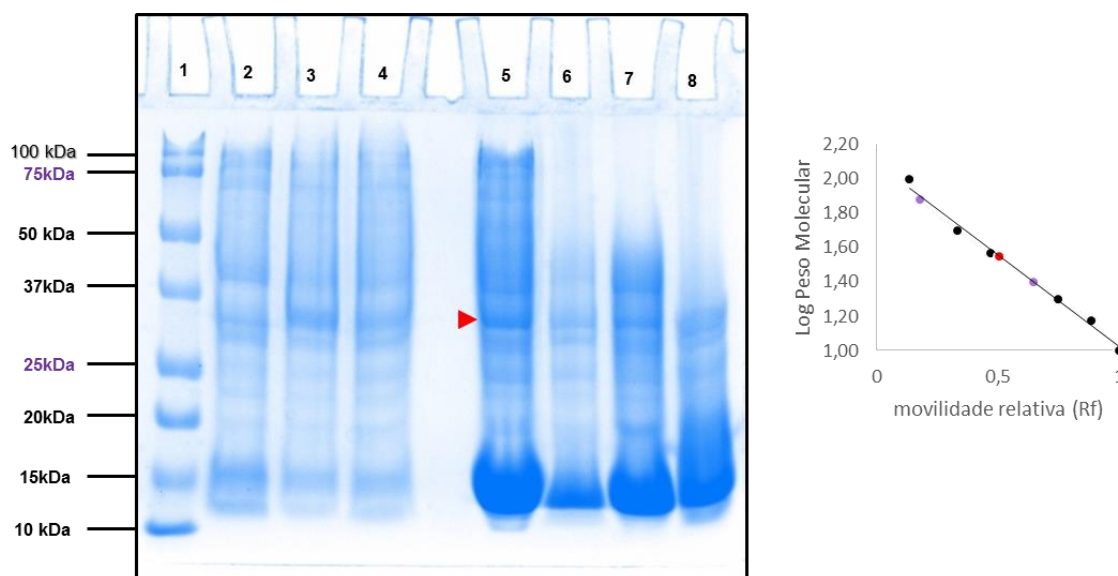


Figura 20. Análise de SDS-PAGE 10% da ORF19 expressa em *E.coli* BL21 (DE3): 1- marcador de peso molecular Precision Plus Protein Dual Color(Bio-Rad); 2 - extrato de células totais não induzidas (T0); 3 - extrato de células totais induzida com IPTG após 2 horas (T2); 4 - extrato de células totais induzida com IPTG após 4 horas (T4); 5 - extratos celulares solúveis (30°C – 15h); 6 - extratos de células insolúveis (sedimento – 30°C, 15h); 7 - extratos celulares solúveis (25°C – 15h); 8 - extratos de células insolúveis (sedimento – 25°C – 15h).

Após a indução com variância nas concentrações de IPTG (0,1mM a 1M) e tempo (2, 4, 6, 15 e 20 horas), pode-se observar que não houve influência significativa na expressão da proteína recombinante da ORF19 clonada em *Escherichia coli* BL21(D3). A *Escherichia coli* é uma bactéria que tem um sistema de expressão altamente variável possuindo rápida multiplicação celular e elevada velocidade para a obtenção de proteínas. Contudo, cada proteína a ser expressa é um desafio independente do organismo hospedeiro, e em cada situação por meio de variações em parâmetros da cultura são necessários para elevar o alto nível de síntese

(SWARTZ, 2001). Em *E. coli*, a expressão recombinante decorre naturalmente de diversos fatores para garantir a alta produtividade tais como a estabilidade do plasmídeo – sendo influenciada por suas propriedades de acordo com tamanho e sequência de nucleotídeos, e origem do hospedeiro (SCHMIDT, 2005); e do mRNA que muitas vezes é comprometida por distintas formas de RNases (ribonucleases) celulares que participam da degradação do RNA em *E.coli*, a frequência dos códons do gene de interesse, as cepas bacterianas (JANA & DEB, 2005; SORENSEN & MORTENSEN, 2005). Outro fator que interfere na expressão de proteínas em *E. coli* são os códons raros ou códons principais, assim denominados devido à combinação de códons mais frequentes que os outros. Os códons CCC, AUA, CGA, CUA e AGG/AGA comumente ocasionam problemas de erros de tradução durante a fase de produção, gerando a redução na qualidade e quantidade de proteína recombinante, sendo que alguns destes erros abrangem a terminação antecipada da tradução, a substituição de aminoácidos, ou a troca de fase de leitura em que alguns pares de nucleotídeos são excluídos ou inseridos (SORENSEN & MORTENSEN, 2005).

Para a sequência de aminoácidos da ORF19, não foi constatado a presença de peptídeo sinal (Figura 21).

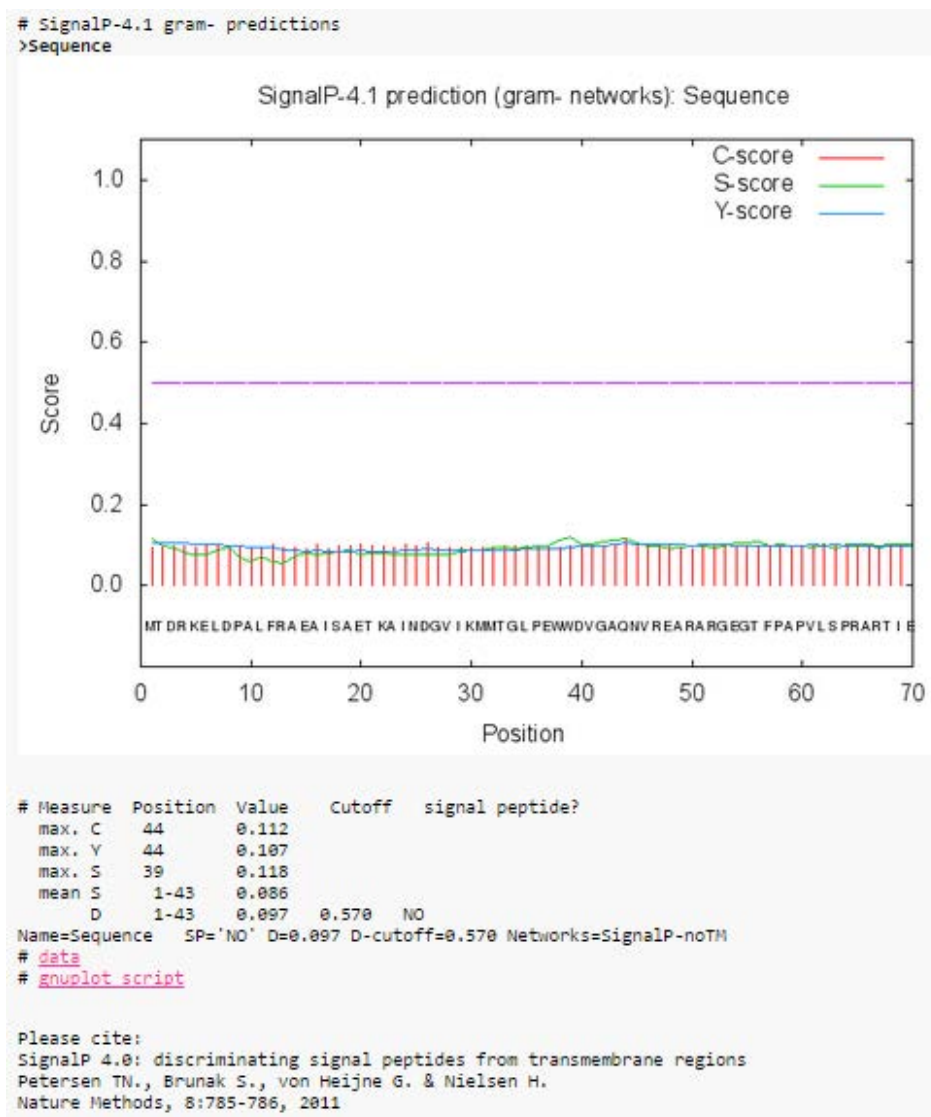


Figura 21. Análise de presença de peptídeo sinal na sequência da proteína ORF 19. (SignalP 4.1 Server- CBS (<http://www.cbs.dtu.dk/services/SignalP/>))

Uma solução para tal problema de expressão é uma nova clonagem em diferentes vetores de expressão e hospedeiros.

5.6.5 Zimograma

O zimograma foi usado para verificar a atividade da enzima no gel (Figura 22). Foi utilizado extrato bruto, deste modo a intensa atividade verificada na canaleta 7 caracterizada por proteínas entre 150kDa e 30 kDa, provavelmente reflita a mais de uma enzima. Não foi observada nenhuma banda nítida no gel, indicando que as bandas correspondem a enzimas com características de serino- protease, o que corrobora com os resultados obtidos a partir das análises da sequência de aminoácidos da ORF19.

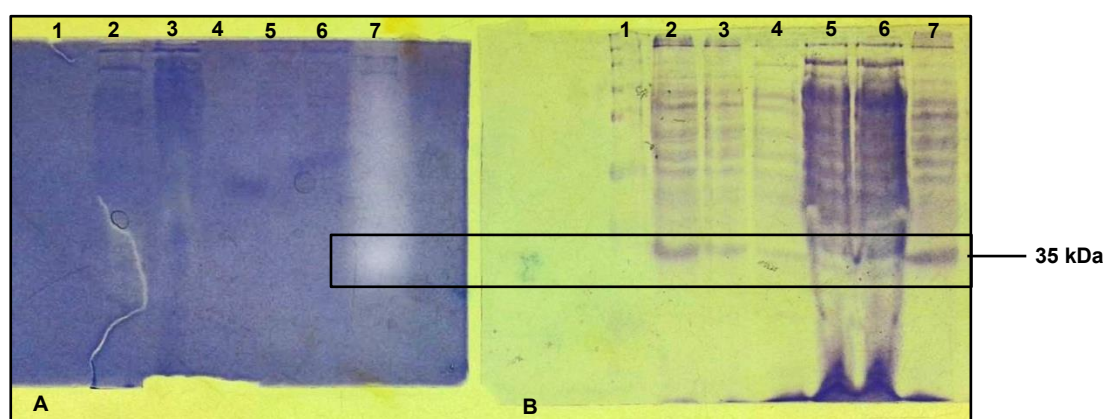


Figura 22. Comparação do perfil proteico em gel SDS-PAGE e análise da atividade proteolítica em Zimograma.

(A) Zimograma. 1- Marcador de Peso Molecular Precision Plus Protein Standards Dual Color (Bio-Rad); 2- Extrato Solúvel induzido 0,8mM IPTG – 25°C; 3 - Extrato Solúvel induzido 0,8mM IPTG – 30°C, 4- Extrato Solúvel induzido 0,6mM IPTG – 25°C; 5- Extrato Solúvel induzido 0,6mM IPTG – 30°C; 6- Extrato Solúvel induzido 0,4mM IPTG – 25°C; 7 - Extrato Solúvel induzido 0,4mM IPTG – 30°C. SDS-PAGE.

(B) SDS-PAGE. 1- Marcador de Peso Molecular Precision Plus Protein Standards Dual Color (Bio-Rad); 2- Extrato Solúvel induzido 0,8mM IPTG – 25°C; 3 - Extrato Solúvel induzido 0,8mM IPTG – 30°C, 4- Extrato Solúvel induzido 0,6mM IPTG – 25°C; 5- Extrato Solúvel induzido 0,6mM IPTG – 30°C; 6- Extrato Solúvel induzido 0,4mM IPTG – 25°C; 7 - Extrato Solúvel induzido 0,4mM IPTG – 30°C. SDS-PAGE.

6. CONCLUSÕES

- Através da abordagem metagenômica utilizando a biblioteca do consórcio degradador de óleo diesel, foi possível identificar uma protease putativa da família Serino-peptidase.
- A partir de análise bioquímica e análises filogenéticas, é possível afirmar que a ORF19 é uma nova proteína da família S9U de enzimas proteolítica bacteriana.
- Não foi possível obter uma superexpressão da proteína em *E.coli* BL21(DE3), porém a mesma está clonada e apresentou atividade mediante ao zimograma.
- O estudo preliminar desta proteína indica uma forma eficiente para descobrir novas enzimas com potencial biotecnológico, através de técnicas moleculares para a obtenção de proteases extracelulares.
- Preveem-se mais estudos sobre as condições ótimas de produção da protease extracelular da ORF19, com intuito de obter resultados para a caracterização funcional desta enzima, uma vez que esta *orf* é similar a uma classe de protease pouco estudada, e que as futuras análises de sua caracterização podem ser eficientes para se obter uma nova enzima para o uso industrial.

REFERÊNCIAS

- ACEVEDO, J.P.; REYES, F.; PARRA, L.P.; SALAZAR, O.; ANDREWS, B.A.; ASENJO, J.Á. Cloning of complete genes for novel hydrolytic enzymes from Antarctic sea water bacteria by use of an improved genome walking technique. **Journal Biotechnologic**, v. 133, p. 277–286, 2008.
- ADRIO, J.; e DEMAIN, E. Microbial Enzymes: Tools for Biotechnological Processes. **Biomolecules**, v.4, n.1, p.117-139, 2014.
- ALLISON, S. D.; VITOUSEKI, P. M.; Responses of extracellular enzymes to simple and complex nutrient inputs: **Soil Biology & Biochemistry**, v.37, p. 937-944, 2004.
- ALTENHOFF, A. M.; STUDER, R. A.; RECHAVI-R M.; DESSIMOZ C. Resolving the ortholog conjecture: orthologs tend to be weakly, but significantly, more similar in function than paralogs. **PLoS Computational Biology**, 8, e1002514, 2012.
- AMANN, R.I.; LUDWIG, W.; SCHLEIFER; K.H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. **Microbiological Reviews**, v. 59, p. 143-169, 1995.
- ANWAR, A.; SALEEMUDDIN, M. Alkaline proteases: a review. **Bioresource Technology**, v. 64, p. 175-183, 1998.
- ASTSCHUL, S.F.; MADDEN, T.L.; SCHAFFER, A.A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. **Nucleic Acids Res.**, v. 25,n. 17, p. 3389-3402, 1997
- AZEREDO, L. A. I.; FREIRE, D. M. G.; SOARES, R. M. A.; LEITE, S. G. F.; COELHO, R. R. R. Production and partial characterization of thermophilic proteases from *Streptomyces sp* isolated from Brazilian cerrado soil. **Enzyme and Microbial Technology**, v. 34, n. 3/4, p. 354-358, 2004.
- BAIRD, I.R.; MOSI, R.; OLSEN, M.; CAMERON, B.R.; FRICKER, S. P.; SKERLJ, R.T. '3 + 1' mixed – ligand oxorhenium (V) complexes and their inhibition of the cysteine proteases cathepsin B an cathepsin K. **Inorganica Chimica Acta**, v. 359, p. 2736 – 2750, 2006.

BARRETT, A. J. Classification of peptidases. In: BARRETT, A. J. (ed). **Methods in Enzymology**. New York: Academic Press. v. 244, p. 1-59, 1994.

BARRETT, S. L.; RAWLINGS, N.D. Evolutionary lines of cysteine peptidases. **Journal of Biological Chemistry**, v. 382, p. 727 – 733, 2001.

BEARE, M. H. et al. A hierarchical approach to evaluating the significance of soil biodiversity to biogeochemical cycling. **Plant Soil**. v.170, p. 5-22, 1995.

BEG, Q. K.; GUPTA, R. Purification and characterization of na oxidationstable, tiol-dependent serine alkaline protease form *Bacillus mojavensis*. **Enzyme and Microbial Technology**, v. 32, n. 2, p. 294-304, 2003.

BELITZ, H. D.; GROSCH, W. **Food chemistry**. 2. ed. Berlin: Springer-Verlag, 1999. 992 p.

BERG, J. M; STRYER, L.; TYMOCZKO, J. L. **Bioquímica**. 5 ed. Rio de Janeiro: Guanabara Koogan. 1104 p., 2004.

BEYNON, R.J.; BOND, J.S. **Proteolytic enzymes: a practical approach**. Oxford England, 1989. p.87-89.

BLUNDELL, T.L. Metalloproteinase superfamilies and drug desing. **Nature Struct Biol**, v. 1, p. 73-75, 1994.

BRAGA, A. A.; MORAIS, P. B.; LINARDI, V. R. Screening of yeast from Brazilian Amazon Rain Forest for extracellular proteinases production. **Systematic and Applied Microbiology**, v. 21, n. 3, p. 353-359, 1998.

COTTRELL, M.T.; MOORE, J.A.; KIRCHMAN, D.L. Chitinases from uncultured marine microorganisms. **Appl Environ Microbiol**, v.65, p. 2553–2557, 1999.

COTTRELL, M. T., WAIDNER, L.A., YU, L. E KIRCHMAN, D.L. Bacterial diversity of metagenomic and PCR libraries from Delaware River. **Environmental Microbiology**. v.7, n.12, p.1883-1895, 2005.

DANIEL, R. The metagenomics of soil. **Nature**. v.3, p. 470-478, 2005.

DURAND-POUSSEREAU, N.; FEVRE, M. Characterization of a protease deficient strain of *Penicillium roqueforti* generated by heterologous plasmid integration: potential use for protein production. **Journal of Biotechnology**, Amsterdam, v. 51, n. 1, p. 97-105, 1996.

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

FERRERO, M. A.; CARTRO, G. R.; ABATE, C. M.; BAIGORI, M. D.; SINERIZ, F. Thermostable alkaline protease of *Bacillus licheniformis* MIR 29: isolation, production and characterization, **Appl. Microbiol. Biotechnol**, v.45, p. 327 – 332, 1996.

FERRER, M., GOLYSHINA, O.V., CHERNIKOVA, T.N., KHACHANE, A.N., MARTINS, D.S.V., YAKIMOV, M.M., TIMMIS, K.N., GOLYSHIN, P.N.. Microbial enzymes mined from the Urania deep-sea hypersaline anoxic basin. **Chem Biol**, v.12, p. 895–904, 2005.

FREDERIKS W.M., MOOK O.R.F. Metabolic mapping of proteinase activity with emphasis on in situ zymography of gelatinases: Review and protocols. **J Histochem & Citochem** v.52, p. 711-722, 2004.

GARCIA, R.A.M.; PEREIRA, M. R.; MAESTER, T.C. Investigation, expression, and molecular modeling of ORF2, a metagenomic lipolytic enzyme. **Appl Biochem Biotechnol**. v.175, p. 3875 – 3887., 2015.

GASTEIGER, E., HOOGLAND, C., GATTIKER, A., DUVAUD, S. E., WILKINS, M. R., APPEL, R. D., BAIROCH, A. Protein identification and analysis tools on the ExPASyServer. In: Walker, J. M. **The Proteomics Protocols Handbook: Humana Press**, 571-607, 2005.

GENCKAL,H.;TARI,C. Alkaline protease production from alkalophilic *Bacillus* sp. isolated from natural habitats. **Enzyme and Microbial Technology**, v. 39, p. 703–710, 2006.

GHORBEL,B.;KAMOUN,A.S.;NASRI,M. Stability studies of protease from *Bacillus cereus* BG1.**Enzyme and Microbial Technology**, v.32, p. 513-518,2003.

GODFREY, T.; WEST, S. Industrial enzymology. New York: **Macmillan Publishers Inc.**, v.2 p.609, 1996.

GORMAM, M.J.; ANDREEVA, O.V.; PASKEWITZ, S.M. Molecular characterization of five serine protease genes cloned from *Anopheles gambiae* hemolymph. **Insect Biochemistry and Molecular Biology**.v. 30, p. 35-46, 2000a.

GORMAM, M.J.; ANDREEVA, O.V.; PASKEWITZ, S.M. Sp22D: a multidomain serine protease with a putative role in insect immunity. **Gene**. v. 251, p. 9-17, 2000b.

GUPTA, R.; BEG, Q.K.; LARENZ, P. Bacterial alkaline proteases: molecular approaches and industrial applications. **Applied Microbiology and Biotechnology**, v.59, p. 15-32, 2002a.

GUPTA, R., BEG, K.Q., KHAN, S., CHAUHAN. An overview on fermentation,downstream processing and properties of microbial alkaline proteases. **Applied Microbiology Biotechnology**. v.60, p.381-395, 2002b.

HAKI, G.D.; RAKSHIT, S.K. Developments in industrially important thermostable enzymes: a review. **Bioresource Technology**, v. 89, p. 17-34, 2003.

HALL, T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. **Nucleic Acids Symposium Series**, v.41, p.95-98, 1999.

HAMEED, A., KESHAVARZ, T. EVANS, C. S. Effect of dissolved oxygen tension and pH on the production of extracellular protease from a new isolate of *Bacillus subtilis* K2, for use in leather processing, **J. Chem. Technol. Biotechnol**, v.74, p.5 – 8, 1999.

HANDELSMAN, J.; RONDON M.R.; BRADY S.F.; CLARDY, J.; GOODMAN, R.M. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. **Chemistry & Biology**, v. 5, p. 245-249, 1998.

HANDELSMAN J. Metagenomics: Application of Genomics to Uncultured Microorganisms. **Microbiology and Molecular Biology Reviews**, v.68, p. 669-685, 2004.

HARDEMAN, F., SJOLING, S. Metagenomic approach for the isolation of a novel low-temperature-active lipase from uncultured bacteria of marine sediment. **Fems Microbiol Ecol**, v. 59, p. 524–534, 2007.

HEUSSEN, C. DOWDLE, EB. Electrophoretic analysis of plasminogen activators in polyacrylamide gels containing sodium dodecyl sulfate and copolymerized substrates. **Analytical Biochemistry**. 102, p.196-202, 1980.

HIGASHIOKA, Y., KOJIMA, H., SATO, S., FUKUI, M. Microbial community analysis at crude oil-contaminated soils targeting the 16S ribosomal RNA, xylM, C23O, and bcr genes. **J. Appl. Microbiol**, v.105, p.1364-5072, 2009.

HUGENHOLTZ, P.; PACE, N. R. Identifying microbial diversity in the natural environment: a molecular phylogenetic approach. **Trends Biotechnol**. V.14, p. 190-197, 1996.

HUSON, D. et al. Methods for comparative metagenomics. **BMC Bioinformatics**, v. 10, p. 1-12, 2009.

JANA, S.; DEB, J.K.; Strategies for eficiente production of heterologous proteins in *Escherichia coli*. **Appl Microbiol Biotechnol**, v.67, p.289 – 298, 2005.

JEON, J.H., KIM, J.T., KANG, S.G., LEE, J.H., KIM, S.J. Characterization and its potential application of two esterases derived from the arctic sediment metagenome. **Mar Biotechnol** (NY), v. 11, p.307–316, 2009.

JEON, J.H., KIM, J.T., KIM, Y.J., KIM, H.K., LEE, H.S., KANG, S.G., KIM, S.J., LEE, J.H. Cloning and characterization of a new cold-active lipase from a deep-sea sediment metagenome. **Appl Microbiol Biotechnol**, v. 81, p. 865–874, 2009.

JONES, B.; SUN, F.; MARCHESI, J. Using skimmed milk agar to functionally screen a gut metagenomic library for proteases may lead to false positives. **Letters in Applied Microbiology**, v. 45, p. 418–420, 2007.

JOO,H.S.;CHANG,C.S. Production of an oxidant and SDS-stable alkaline protease from an alkaophilic *Bacillus clausii* I-52 by submerged fermentation: Feasibility as a laundry detergentadditive. **Enzyme and Microbial Technology**, v. 38, p.176–183, 2006.

KAMPHUIS, I.G.; KALK, K.H.; SWARTE, M.B.A.; DRENT, J. Structure of papain refined at 1.65 Å resolution. **J. Mol. Biol.**, v.179, p. 233-256, 1984.

KISS, A.L., SZELTNER, Z., FULOP, V., POLGAR, L. His507 of acylaminoacyl peptidase stabilizes the active site conformation, not the catalytic intermediate. **FEBS Lett**, v. 571, p. 17-20, 2004.

KITANO, H.; KATAOKA, K.; FURUKAWA, K.; HARA, S. Specific expression and temperature-dependent expression of the acid protease-encoding gene (pep A) in *Aspergillus oryzae* in solid state culture (rice-koji). **Journal of Bioscience and Bioengineering**, v. 93, n. 6, p. 563-567, 2002.

KNIGHT, V.; SANGIER, J.J.; DITULLIO, D.; BRACCILI, S.; BONNER, P.; WATERS, J.; HUGHES, D.; ZHANG, L. Diversifying microbial natural products for drug discovery. **Applied and Environmental Microbiology**, v. 62, p. 446-458, 2003.

KUNAKBAEVA, Z.; CARRASCO, R.; ROZAS, I. Na approximation to the mechanism of inhibition of cysteine proteases: nucleophilic sulfur addition to Michael acceptors type compounds. **Journal of Molecular Structure**, v. 626, p. 209 – 216, 2003.

LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, v.227, n.5259, p.680–685, 1970.

LAMMLE, K., ZIPPER, H., BREUER, M., HAUER, B., BUTA, C., BRUNNER, H., RUPP, S. Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. **J Biotechnol**, v.127, p. 575–592, 2007.

LARKIN, M.A., BLACKSHIELDS, G., BROWN, N.P., CHENNA, R., MCGETTIGAN, P.A., MCWILLIAM, H., VALENTIN, F., WALLACE, I.M., WILM, A., LOPEZ, R., THOMPSON, J.D., GIBSON, T.J., HIGGINS, D.G. Clustal W and Clustal X version 2.0. **Bioinformatics**. v.1; n.23(21), p. 2947-2948, 2007.

LECLEIR, G.R., BUCHAN, A., HOLLIBAUGH, J.T. Chitinase gene sequences retrieved from diverse aquatic habitats reveal environment-specific distributions. **Appl Environ Microbiol**, v.70, p.6977–6983, 2004.

LEUNG-TOUNG, R.; LI, W.; TAM, T.F.; KARIMIAN, K. Thiol-dependent enzymes and their inhibitors: a Review. **Current Medicinal Chemistry**, v.9, n.9, p. 979-1002, 2002.

LEVEAU J.H.J. The magic and menace of metagenomics: prospects for the study of plant growth-promoting rhizobacteria. **Eur J Plant Pathol**, v. 119, p. 279 – 300, 2007.

LI, S., YANG, X., YANG, S., ZHU, M., WANG, X. Technology Prospecting on Enzymes: Application, Marketing and Engineering. **Computational and Structural Biotechnology Journal**. v.2 n.3, e201209017, 2012.

LIAW, R.B.; CHENG, M.P.; WU, M.C.; LEE, C.Y. Use of metagenomic approaches to isolate lipolytic genes from activated sludge. **Bioresource Technology**, v. 101, p. 8323 – 8329, 2010.

LODISH, H., A. BERK, P. MATSUDAIRA *et al.* **Biologia Celular e Molecular** (5a ed.) WH Freeman, p. 66-72, 2004.

LÓPEZ-OTÍN, C., BOND, J.S. Proteases: multifunctional enzymes in life and disease, **J. Biol. Chem.** v.283,p.30433–30437, 2008.

LUNA, J. M.; RUFINO, R. D.; ALVES DA SILVA, C. A.; CAMPOSTAKAKI, G. M.; SARUBBO, L. A. Detecção do potencial biotecnológico em bactérias e leveduras isoladas de sedimentos de mangue. In: **SEMINÁRIO BRASILEIRO DE BIOTECNOLOGIA ENZIMÁTICA**, 3 Brasília. **Resumos...**Brasília, DF, 2002.

MANFIO, P. G. Coleção de microrganismos. Sociedade Brasileira para o Progresso da Ciência. Print ISSN 009-6725. **Ciências Cultura**, v. 54, n01, p. 1-6, 2002.

MARTA, E.; ALIÓS, C. P.; HENRIKSEN, P. *et al.* Diversity of planktonic photoautotrophic microorganisms along a salinity gradient as depicted by microscopy, flow cytometry, pigment analysis and DNA-based methods. **FEMS Microbiology Ecology**, v. 49, p. 281, 2004

MAURER, K. H. Detergent proteases. **Current Opinion in Biotechnology**, v. 15, p. 330- 334, 2004.

MOREIRA, K. A.; SILVA, A. C. J.; BEZERRA, R. S.; CARVALHO-JÚNIOR, L. B.; TEIXEIRA, M. F. S.; CHAVES, A. C.; PORTO, A. L. F.; LIMA-FILHO, J. L. Estudos preliminares das propriedades cinéticas da protease alcalina de *Nocardia* sp In: **SEMINÁRIO BRASILEIRO DE BIOTECNOLOGIA ENZIMÁTICA**, 3., Brasília **Resumos...**Brasília, DF, 2002.

MORRIS, L.S., EVANS, J., MARCHESI, J.R. A robust plate assay for detection of extracellular microbial protease activity in metagenomic screens and pure cultures. **Journal of Microbiological Methods**, v. 9, p. 144 – 146, 2012.

MUKHERJEE, A. K.; ADHIKARI, H.; RAI, S.K. Production of alkaline protease by a thermophilic *Bacillus subtilis* under solid – state fermentation (SSF) condition using *Imperata cylindrica* grass and potato peel as low-cost médium: Characterization and application of enzyme in detergente formulation. **Biochemical Engineering Journal**, v.39, p. 353 – 361, 2008.

NEHETE, P. N., SHAH, V. D., KOTHARI, R. M. Profiles of alkaline protease production as a function of composition of the slant, age, transfer and isolate number and physiological state of culture, **Biotechnol.** v.7, p. 413 – 418, 1985.

NIKOLAKI, S.; TSIAMIS, G. Microbial diversity in the Era of Omic Technologies. **Biomed Research Corporation**, v.958719, 2013.

PETERSEN, T.N., BRUNAK, S., VON HEIJNE, G., NIELSEN, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, v. 8, p. 785 – 786, 2011.

PIRÔLLO, P.S., MARIANO, A.P., LOVAGLIO, R.B., COSTA, S.G.V.A.O., WALTER, V., HAUSMANN, R., CONTIERO, J. Biosurfactant synthesis by *Pseudomonas aeruginosa* LBI isolated from a hydrocarbon-contaminated site M. **Journal of Applied Microbiology**, v.105, p.1484–1490, 2008.

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

PAGE, M.J.; Di CERA, E. Serine peptidases: classification, structure and function. **Cellular and Molecular Life Sciences**, v.65, p.1220-1236, 2008.

PAIXÃO, D. A.; DIMITROV, M.R.; PEREIRA, R. M.; ACCORSINI, F.R.; VIDOTTI, M.B.; LEMOS, E. G. M. Molecular analysis of the bacterial diversity in a specialized consortium for diesel oil degradation. **Revista Brasileira de Ciência do Solo**, v.34, p. 773 – 781, 2010.

PATIL, U.; AMBALAL, C. Purification and characterization of solvent-tolerant, thermostable, alkaline metalloprotease from alkalophilic *Pseudomonas aeruginosa* MTCC 7926. **J Chem technol biotechnol**, v.84, p. 1255 – 1262, 2009.

POZA, M.; MIGUEL, T.; SIERO, C.; VILLA, T. G. Characterization of a broad pH range protease of *Candida caseinolytica*. **Journal of Applied Microbiology**, Oxford, v. 91, n. 5, p. 916-921, 2001.

RAO, M.B., TANKSALE, A.M., GHATGE, M.S., DESHPANDE, V.V., Molecular and biotechnological aspects of microbial proteases. **Microbiology and Molecular Biology Reviews**, v.62, n. 3, p. 597 - 635, 1998.

RAWLINGS, N. D.; BARRETT, A. J. MEROPS: the peptidase database. **Nucleic Acids Research**, v. 27, p.325 – 331, 1999.

RAWLINGS, N.D., POLGAR,L.,BARRETT, A.J. A new family of serine-type peptidases related to prolyl oligopeptidase. **Biochem J**, v. 279,p. 907-908, 1991.

RAWLINGS, N.D., WALLER, M., BARRETT, A.J. & BATEMAN, A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. **Nucleic Acids Research**, v.42, p.D503-D509, 2014.

RICHARDSON, T.H., TAN, X., FREY, G., CALLEN, W., CABELL, M., LAM, D., MACOMBER, J., SHORT, J.M., ROBERTSON, D.E., MILLER, C. A novel, high performance enzyme for starch liquefaction. Discovery and optimization of a low pH, thermostable alpha-amylase. **J Biol Chem**, v. 277, p. 26501–26507, 2002.

RIESENFELD, C.S.; SCHLOSS, P.D.; HANDELSMAN, J. Metagenomics: Genomic analysis of microbial communities. **Annu Rev Genet**, v. 38, p. 525-52, 2004.

RHEE, J.K., AHN, D.G., KIM, Y.G., OH, J.W. New thermophilic and thermostable esterase with sequence similarity to the hormone-sensitive lipase family, cloned from a metagenomic library. **Appl Environ Microbiol**, v.71, p. 817–825, 2005.

RONQUIST, F., M. TESLENKO, VAN DE MARK, P., AYRES, D., DARLING, A., HÖHNA, S., LARGET, B., LIU, L., SUCHARD, M. A., HUELSENBECK, J. P. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. **Systematic Biology**. v. 61, n. 3, p. 539 - 542, 2011.

ROBE, P.; NALIN, R.; CAPELLANO, C.; VOGEL TM. C.; SIMONET, P. Extraction of DNA from soil. **Eur J Soil Biol**, v.39, p. 183-190, 2003.

ROMERO, F.J.; GARCI'A, L.A.; SALAS, J.A.; DI'AZ, M.; QUIRO'S ,L.M. Production, purification and partial characterization of two extracellular proteases from *Serratia marcescens* grown in whey. **Process. Biochem.**v. 36, p. 507–515, 2001.

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

SAJID, M.; McKERROW, J.H. Cysteine proteases of parasitic organisms. **Molecular and Biochemical Parasitology**, v.120, p. 1-21, 2002.

SCHLEHECK, D., LECHNER, M., SCHONENBERGR, R., SUTER, M. J. F., COOK , A. M. Desulfonation and Degradation of the Disulfodiphenylethercarboxylates from Linear Alkyldiphenyletherdisulfonate Surfactants. **Applied and environmental microbiology**, v.69, p.938–944, 2003.

SCHMID, F.R. Optimization and scale up of industrial fermentation processes. **Appl Microbiol Biotechnol**, v. 68, p. 425-435, 2005.

SERVANT, F., BRU, C., CARRÉRE, S., COURCELLE, E., GOUZY, J., PEYRUC, D., KAHN, D. Prodom: Automated clustering of homologous domains. **Briefings in Bioinformatics**, v.3 n.3, 246 – 251, 2002.

SILVA NEVES, K.C.; PORTO, A.L.F.; TEIXEIRA, M.F.S. Seleção de leveduras da Região Amazônica para produção de protease extracelular. **Acta Amazônica**, v.36, p. 299-306, 2006.

SINGHAL, P.; NIGAM, V. K.; VIDYARTHI, A.S; Studies on production, characterization and applications of microbial alkaline proteases. **International Journal of advanced Biotechnology an Research**, v.3; p. 653 – 669, 2012.

SORENSEN, H. P.; MORTENSEN, K. K.; Advanced genetec strategies for recombinant protein expression in Escherichia coli. **Journal of biotechnology**, v.115. p. 113-128, 2005.

STEELE H, STREIT WR Metagenomics: Advances in ecology and Biotechnology. **FEMS Microbiology Letters**, v. 247, p. 105-111, 2005.

TAMURA, K., STECHER, G., PETERSON, D., FILIPSKI, A., KUMAR, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. **Molecular Biology and Evolution**, v. 30, p. 2725-2729, 2013.

TORSVIK, V.; OVREAS, L., THINGSTAD, T. F. Prokaryotic diversity – Magnitude, dynamics and controlling factors. **Science**, v. 296, p. 1064-1066, 2002a.

TORSVIK, V.; OVREAS, L. Microbial diversity and function in soil: from genes to ecosystems. **Current Opinion in Microbiology**, v. 5, p. 240 – 245, 2002b.

UCHIYAMA, T.; MIYAZAKI, K. Functional metagenomics for enzyme discovery: challenges to efficient screening. **Current Opinion in Biotechnology**, v. 20, p. 612 – 622, 2009.

VAN BEILEN, J.B.; LI, Z. Enzyme technology: an overview. **Current Opinion in Biotechnology**; v. 42, p.338 – 344, 2002.

VENTER, J. C. et al. Environmental genome shotgun sequencing of the Sargasso Sea. **Science**, v. 304, p. 66–74, 2004.

WHITMAN, W. B.; COLEMAN, D. C.; WIEBE, W. J. Prokaryotes: The unseen majority. **Proceedings of the National Academy of Sciences of the United States of America**, v.95, p. 6578 – 6583, 1998.

WILSON, I.; VOGEL, J.; SOMERVILLE, S. Signalling pathways: A common theme in plants and animals? **Current Opinion in Biology**.v. 7, n. 3, p. 175-178, 1997

WINDLE, H.; KELLEHER, D. Identification and Characterization of a Metalloprotease Activity from Helicobacter pylori. **Infection and Immunity**, v. 65, n. 8, p. 3132–3137, 1997.

WOESE C. R., Bacterial evolution. **Microbiol. Rev.**, v. 51, p. 221–271, 1987.

WUNSCH, L.; BRUGGEMANN, L.; BABEL, W. Determination of substrate utilization patterns of soil microbial communities: An approach to assess population changes after hydrocarbon pollution. **FEMS Microbiol. Ecol**, v.17, p. 295-305, 1995.

ZHANG, Y., ZHAO, J., ZENG, R. Expression. And characterization of a novel mesophilic protease from metagenomic library derived from Antarctic. **Extremophiles**. v. 15, p. 23 – 29, 2011.