MARIANA RANGEL PEREIRA

**PROSPECÇÃO DE GENES CODIFICADORES DE ENZIMAS
LIPOLÍTICAS EM BIBLIOTECA METAGENÔMICA DE CONSÓRCIO
MICROBIANO DEGRADADOR DE ÓLEO DIESEL**

Dissertação apresentada ao Programa de Pós-Graduação Interunidades em Biotecnologia USP/Instituto Butantan/IPT, para obtenção do Título de Mestre em Biotecnologia.

**São Paulo
2011**

MARIANA RANGEL PEREIRA

**PROSPECÇÃO DE GENES CODIFICADORES DE ENZIMAS
LIPOLÍTICAS EM BIBLIOTECA METAGENÔMICA DE CONSÓRCIO
MICROBIANO DEGRADADOR DE ÓLEO DIESEL**

Dissertação apresentada ao Programa de Pós-Graduação Interunidades em Biotecnologia USP/Instituto Butantan/IPT, para obtenção do Título de Mestre em Biotecnologia.

Área de concentração: Biotecnologia

Orientadora:

Dra. Eliana Gertrudes de Macedo Lemos

**São Paulo
2011**

DADOS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
Serviço de Biblioteca e Informação Biomédica do
Instituto de Ciências Biomédicas da Universidade de São Paulo

reprodução não autorizada pelo autor

Pereira, Mariana Rangel.

Prospecção de genes codificadores de enzimas lipolíticas em biblioteca metagenômica de consórcio microbiano degradador de óleo diesel / Mariana Rangel Pereira. -- São Paulo, 2011.

Orientador: Eliana Gertrudes de Macedo Lemos.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação Interunidades em Biotecnologia USP/IPT/Instituto Butantan. Área de concentração: Biotecnologia. Linha de pesquisa: Prospecção de genes de interesse biotecnológico em biblioteca metagenômica.

Versão do título para o inglês: Screening for lipolytic enzyme codification genes in a metagenomic library of consortia specialized in diesel oil degradation.

Descritores: 1. Biblioteca metagenômica 2. Consórcio microbiano 3. Lipase 4. Esterase 5. Biotecnologia I. Lemos, Eliana Gertrudes de Macedo II. Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação em Biotecnologia III. Título.

ICB/SBIB007/2011

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS

Candidato(a): Mariana Pereira.

Título da Dissertação: Prospecção de genes codificadores de enzimas lipolíticas em biblioteca metagenômica de consórcio microbiano degradador de óleo diesel.

Orientador(a): Eliana Gertrudes de Macedo Lemos.

A Comissão Julgadora dos trabalhos de Defesa da **Dissertação de Mestrado**,
em sessão pública realizada a/...../....., considerou

() Aprovado(a)

() Reprovado(a)

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Presidente: Assinatura:
Nome:
Instituição:

Aos meus pais,
Fernanda Garcia Rangel e João César Aparecido
Pereira, pelo grande amor.

Ao meu irmão,
Tiago Rangel Pereira, por ser o pior e melhor irmão
do mundo.

Aos meus avôs,
Jacy Silva Garcia Rangel e Celso Carvalho Rangel,
pelas lembranças maravilhosas.

Às minhas tias,
Renata Garcia Rangel e Patrícia Garcia Rangel por
sempre estarem ao meu lado, meus anjos da
guarda.

AGRADECIMENTOS

A Profa. Dra. Eliana Gertrudes de Macedo Lemos pela orientação, acreditar nos meus sonhos e me guiar para que eles se tornassem reais;

À Profa. Dra. Lúcia Maria Carareto Alves e ao Prof. Dr. Jackson Marcondes Fernandes por revisarem e discutirem este trabalho, contribuindo grandemente na sua melhoria;

Ao Dr. João Carlos Campanharo pela amizade, dedicação e auxílio nos experimentos sempre que necessário;

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

Aos integrantes do Laboratório de Bioquímica de Micro-organismos e Plantas (LBMP) pelas conversas, amizade, troca de conhecimento científico e ensinamento. Com certeza, a presença deste grupo foi fundamental para o desenvolvimento deste trabalho;

À mestranda Érica Lopes por ser sempre tão prestativa e querer o meu bem;

À mestranda Thaís Carvalho Maester por ser minha amiga, colega de quarto e de trabalho, sou muito grata por você ter participado de tantos momentos importantes, como este. Espero que possamos caminhar juntas por muito tempo ainda, na busca incessante pelo conhecimento;

À aluna de Iniciação Científica Rosmeriana Afnis Marioto Garcia por ser a minha primeira co-orientada, aprendi e aprendo muito com você;

Ao Departamento de Tecnologia pela infra-estrutura na execução deste projeto;

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo financiamento e bolsa concedida;

Ao meu pai João César, pelo exemplo na determinação e no trabalho; e à minha mãe Fernanda, pelo maior e melhor coração do mundo;

Ao meu irmão Tiago por sentir tanto orgulho de mim;

Ao meu namorado, Diogo Pinetti Markezoni, pela amizade, zelo e amor. Sem você muita coisa teria sido diferente. Obrigada por caminhar ao meu lado;

Aos meus familiares da cidade de Ribeirão Preto e Rifaina pela constante torcida à distância;

Aos meus amigos Ana Raquel Monteiro e Vinícius Canato Santana por terem tornado São Paulo/SP muito mais divertida e muito menos cinza;

Às minhas amigas Livia, Thaís, Marina, Patrícia, Isabela e Carol pelo carinho desde a época do colegial;

À República Toca da Onça por ser aquele lugar que eu SEMPRE vou querer estar;

Às minhas amigas Onçona, Lesada, Sufrida, Gosminha, Pá-çoká, Pintada e Misskenta por tanta risada e felicidade, mesmo que fosse no final do dia, nas horas de cansaço. Sempre levarei vocês e a nossa república comigo, independente do lugar que eu esteja.

Este projeto foi realizado no Laboratório de Bioquímica de Microorganismos e Plantas, Campus de Jaboticabal, Departamento de Tecnologia, sendo financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), número do processo 2009/06691-0.

*“Não importa quantas vezes desatino,
nem quantas vezes a vida me espalma
Sou o mestre e senhor do meu destino:
Sou o capitão de minha alma.”*

William Ernest Henley

RESUMO

PEREIRA, M. R. **Prospecção de genes codificadores de enzimas lipolíticas em biblioteca metagenômica de consórcio microbiano degradador de óleo diesel**. 2011. 102 f. Dissertação (Mestrado em Biotecnologia) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2011.

Há uma longa tradição no uso de micro-organismos e enzimas microbianas para o processamento de materiais naturais. Atualmente na manufatura, industrialização alimentícia e produção de detergentes, as enzimas servem como “chaves” de ativação de componentes presentes nos mesmos. Dentro deste contexto, as enzimas lipolíticas vêm atraindo atenção no mercado global devido ao enorme potencial biotecnológico, como por exemplo: nas misturas enzimáticas para formulação de detergentes; na indústria de couro, podendo ser aplicada juntamente com outras hidrolases na remoção de pêlos e gorduras subcutâneas; na produção de cosméticos, fármacos, aromas, biodiesel, dentre outros. A maior parte das lipases é de origem microbiana, elas apresentam baixa toxicidade, são facilmente biodegradáveis e destacam-se pela quimiosseletividade, o que reduz os efeitos colaterais dos fármacos. Tendo em vista os estudos em diversidade microbiana, pode-se afirmar que a maioria dos micro-organismos que compõe a biosfera ainda não foi estudada, fato confirmado pelas estimativas da complexidade do DNA e descoberta de várias sequências únicas dos genes *16S rRNA* e *18S rRNA* de origem ambiental. A fim de acessar o recurso genético das espécies microbianas, a metagenômica tornou-se uma abordagem que oferece uma combinação quase ilimitada para encontrar novos genes codificadores de produtos relevantes como lipases e esterases. O objetivo deste trabalho foi o de prospectar genes para a codificação de enzimas lipolíticas em uma biblioteca metagenômica fosmidial de um consórcio microbiano degradador de óleo diesel, de aproximadamente 4500 clones. A seleção foi feita pela atividade lipolítica através do cultivo dos clones em placa de petri contendo o meio Luria-Bertani (LB) suplementado de 1% de tributirina (v/v), 1% de goma arábica (p/v), 0,00125% de cloranfenicol (v/v) e 0,001% arabinose (v/v). As células ficaram em cultivo a 37 °C por 48 horas e depois foram transferidas a 4 °C. A avaliação foi realizada pela observação de um de halo ao redor da colônia, sendo positiva para 30 clones dentre os quais dois se destacaram. Estes dois clones que apresentaram halos maiores foram selecionados, e tiveram o seu DNA subclonado em vetor puC19. Os DNAs das sub-bibliotecas foram sequenciados no ABI 3100 (Applied Biosystems), gerando um *contig* completo para cada clone, que foram comparados com as sequências do banco Nacional Center for Biotechnology Information (NCBI), através do ORF Finder. Uma ORF de 322 aminoácidos com 76% identidade para uma possível esterase/lipase foi identificada em um dos *contigs* analisados. Por outro lado, quatro ORFs relacionadas a esterase/lipase puderam ser identificadas no segundo clone analisado, dentre as quais uma alcançou 58% de identidade com uma possível esterase/lipase de bactéria não cultivável. Foram feitos alinhamentos com o programa Clustal W e construção de árvores filogenéticas pelo MEGA, para comparação entre as ORFs encontradas e sequências depositadas. As árvores indicam que o primeiro clone apresenta a ORF similar à família IV das enzimas lipolíticas, enquanto o segundo possui três ORFs pertencentes à família V e uma à família IV. Através dos alinhamentos foi possível identificar os sítios ativos representativos de cada família, confirmando o resultado das árvores filogenéticas. Quando a mesma análise foi realizada com sequências já patenteadas, algumas ORFs mostraram similaridade com sequências pertencentes à Bayer Chemicals e Monsanto. Duas ORFs foram selecionadas para clonagem em *Escherichia coli*

BL21, usando-se o vetor de expressão pET28a, e em ambos casos os clones obtidos apresentaram atividade lipolítica em placa de petri contendo tributirina.

Palavras-chave: Biblioteca metagenômica. Consórcio Microbiano. Lipase. Esterase. Biotecnologia.

ABSTRACT

PEREIRA, M. R. **Screening for lipolytic enzyme codification genes in a metagenomic library of consortia specialized in diesel oil degradation.** 2011. 102 f. Master thesis (Biotechnology) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2011.

Microorganisms and microbial enzymes have long been used for processing natural materials. Currently, these enzymes are used as “keys” for activating components in the process of food manufacturing and industrialization and the production of detergents. In this context, lipolytic enzymes have been attracting global market attention because they show enormous biotechnological potential. This includes uses such as in enzyme mixtures for the production of detergents, the processing of leather where the enzymes are applied with other hydrolytic enzymes for the removal of hair and subcutaneous fat, the production of cosmetics and other pharmaceuticals, perfumes and biodiesel. Most lipolytic enzymes are derived from microbes, present low toxicity, are easily biodegraded and are notably selective of chemicals, a very positive trait which permits pharmaceutical products made with these enzymes to have few collateral effects. Research done on microbe diversity shows that most of the microorganisms which compose our biosphere have not yet been studied, a fact confirmed by DNA complexity estimates and the discovery of many unique environmental *16S rRNA* and *18S rRNA* gene sequences. When accessing the genetic resources of microbe species, metagenomic science was found to be an approach which offers an almost unlimited combination for the discovery of new genes that codify relevant products such as lipases and esterases. The present work was done as an attempt to find genes which codify lipolytic enzymes in a fosmid metagenomic library composed of diesel oil degradation microbe consortia, with approximately 4500 clones. Clones were selected according to lipolytic activity and were assessed by cultivation in Petri dishes containing Luria-Bertani (LB) medium supplemented with 1% tributyrin (v/v), 1% gum arabic (w/v), 0.00125% chloramphenicol (v/v) and 0.001% arabinose (v/v). All cultures were maintained at 37°C for 48 hours and then transferred at 4°C. Assessment was done by observation of halos formed around the colonies, with 30 clones producing halo formations and identified as positives. Of these, two showed prominent positive results, producing larger halos compared to the others. These were selected and had their DNA sub cloned in pUC19 vectors. DNA from the sub libraries was sequenced by ABI 3100 (Applied Biosystems), generating a complete *contig* for each clone. These were compared to sequences from the National Center for Biotechnology Information (NCBI) bank by the ORF Finder. An ORF of 322 amino acids with 76% of identity for a possible esterase/lipase was identified in one of the assessed *contigs*. Notwithstanding, four ORFs related to esterase/lipase could be identified in the second assessed clone, with one of these attaining 58% of identity to a possible esterase/lipase of non cultivatable bacteria. For comparison of the found ORFs and the deposited sequences, alignments were done using the Clustal W program and cladograms were constructed using MEGA. Assessment of the cladograms showed that the first clone presented an ORF similar to that of lipolytic enzyme family IV, while the second had three ORFs pertaining to family V and one to family IV. The alignments made possible the identification of active sites which represent each family, confirming the results obtained with the construction of the cladograms. When the same evaluation was done using patented sequences, some ORFs showed similarities to sequences owned by Bayer Chemicals and Monsanto. Two ORFs were selected for cloning in *Escherichia*

coli BL21, using the pET28a expression vector. In both cases the obtained clones presented lipolytic activity when cultivated in Petri dishes containing tributyrin.

Key words: Metagenomic library. Microbe consortium. Lipase. Esterase. Biotechnology.

LISTA DE FIGURAS

Figura 1. Estrutura secundária do dobramento das alfa/beta hidrolases	32
Figura 2. Fluxograma metodológico identificando as etapas desenvolvidas neste projeto.....	35
Figura 3. Hidrólise da tributirina pela lipase e a formação dos produtos glicerol e ácido butírico	37
Figura 4. Esquema ilustrativo apresentando características do vetor pUC19.....	42
Figura 5. Esquema ilustrativo apresentando características do vetor pET28a.....	50
Figura 6. Análise <i>in vitro</i> da célula EPI300™-T1R <i>Escherichia coli</i> , não transformada, em placa de petri 90 mm X 15 mm.....	53
Figura 7. Análise <i>in vitro</i> nas 44 placas, 4224 clones, que constituem a biblioteca metagenômica com o meio de cultura LB suplementado com 1% de tributirina (v/v); 1% de goma arábica (p/v); 0,001% de arabinose (v/v); e 0,00125% de cloranfenicol (v/v)	54
Figura 8. Análise <i>in vitro</i> das placas 17 e 32 da biblioteca metagenômica com o meio de cultura LB suplementado com 1% de tributirina (v/v); 1% de goma arábica (p/v); 0,001% de arabinose (v/v); e 0,00125% de cloranfenicol (v/v).....	55
Figura 9. Análise <i>in vitro</i> em placa de petri dos 30 clones da biblioteca metagenômica, previamente selecionados, em meio LB suplementado com 1% de tributirina (v/v); 1% de goma arábica (p/v); 0,001% de arabinose (v/v); 0,00125% de cloranfenicol (v/v); e 0,001% de Rhodamine-B (v/v).	56
Figura 10. Perfil eletroforético em gel de agarose 0,8% representando a restrição do DNA fosmidial dos clones P117.E10 e P132.D09 com as enzimas: <i>EcoRI</i> ; <i>BamHI</i> e <i>HindIII</i>	57

Figura 11. Perfil eletroforético em gel de agarose 0,8% da extração do DNA fosmidial dos dois clones selecionados	58
Figura 12. Perfil eletroforético da nebulização a 3 kgf por 16 segundos do clone P117.E10 e P132.D09 em gel de agarose 0,8%	59
Figura 13. Perfil eletroforético do resultado da eluição dos clones P117.E10 e P132.D09, através do Kit Geneclean III, em gel de agarose 0,8%	59
Figura 14. Perfil eletroforético em gel de agarose 0,8% de DNA plasmidial extraído de 48 subclones originados da construção da sub-biblioteca do clone P117.E10	60
Figura 15. Perfil eletroforético em gel de agarose 0,8% de DNA plasmidial extraído de 48 subclones originados da construção da sub-biblioteca do clone P132.D09	61
Figura 16. Mapa físico gerado pelo programa Pages'09, versão 4.04, do clone P117.E10 que expressou atividade lipolítica.....	65
Figura 17. Dendrograma do agrupamento hierárquico baseado nas sequências de membros das famílias lipolíticas, mostrando as relações filogenéticas das ORFs de alpha/beta hydrolase encontradas no <i>contig</i> do clone P117.E10 em relação as sequências utilizadas por Arpigny e Jaeger (1999).....	68
Figura 18. Dendrograma do agrupamento hierárquico baseado nas sequências de membros das famílias lipolíticas, mostrando as relações filogenéticas das ORFs de alfa/beta hidrolase encontradas no <i>contig</i> do clone P117.E10 em relação as sequências utilizadas por Wu e Sun (2009).....	70
Figura 19. Dendrograma do agrupamento hierárquico baseado nas sequências de membros das famílias lipolíticas, mostrando as relações filogenéticas das ORFs de alfa/beta hidrolase encontradas no <i>contig</i> do clone P117.E10 em relação as sequências utilizadas por Couto et al. (2010)	72

Figura 20. Alinhamento das sequências da ORF8 com as sequências de aminoácidos de <i>Moxarella</i> sp. (número de acesso X53868), <i>Pseudomonas</i> sp. (número de acesso AF034088) e <i>Cupriavidus necator</i> (número de acesso L36817), que são representantes da família IV	74
Figura 21. Alinhamento das sequências das ORFs 16, 17 e 21 com as sequências de aminoácidos de <i>Sulfolobus acidocaldarius</i> (número de acesso AF071233), <i>Moxarella</i> sp. (número de acesso X53869) e <i>Psychrobacter immobilis</i> (número de acesso X66712), que são representantes da família V	75
Figura 22. Mapa físico gerado pelo programa Pages'09, versão 4.04, do clone PI32.D09 que apresentou atividade lipolítica	80
Figura 23. Dendrograma do agrupamento hierárquico baseado nas sequências de membros das famílias lipolíticas, mostrando a relação filogenética da ORF30 de alfa/beta hidrolase, encontrada no <i>contig</i> do clone PI32.D09, em relação as sequências utilizadas por Arpigny e Jaeger (1999).....	82
Figura 24. Alinhamento da sequência da ORF30 com as sequências de aminoácidos de <i>Moxarella</i> sp. (número de acesso X53868), <i>Pseudomonas</i> sp. (número de acesso AF034088) e <i>Cupriavidus necator</i> (número de acesso L36817), que são representantes da família IV	84
Figura 25. Dendrograma do agrupamento hierárquico baseado nas sequências de 32 micro-organismos extraídas do NCBI com a ORF30, encontrada no <i>contig</i> do clone PI32.D09	86
Figura 26. Alinhamento da sequência da ORF30 com as 32 sequências extraídas do NCBI.....	88
Figura 27. Dendrograma do agrupamento hierárquico baseado nas 11 sequências patenteadas e extraídas do NCBI, com as ORFs encontradas nos clones PI17.E10 e PI32.D09	89

Figura 28. Perfis eletroforéticos em gel de agarose 1% mostrando padronização da temperatura de pareamento das reações de PCR em termociclador de gradiente.....	90
Figura 29. Perfil eletroforético em gel de agarose 1% da reação de PCR purificada com o Kit Abstract for Wizard SV Gel and PCR Clean-Up System	91
Figura 30. Perfil eletroforético em gel de agarose 1% da reação de digestão com o DNA dos clones das construções pET28a + ORF16 e pET28a + ORF17, com as enzimas de restrição <i>EcoRI</i> e <i>XhoI</i>	92
Figura 31. Análise <i>in vitro</i> em placa de petri dos clones das construções pET28a + ORF16 e pET28a + ORF17 e da <i>Escherichia coli</i> BL21 contendo o vetor pET28a fechado (controle negativo)	93

LISTA DE TABELAS

Tabela 1 - Síntese dos pares de oligonucleotídeos iniciadores de síntese para as ORFs 16 e 17 presentes no clone P117.E10, com sítios de restrição para as enzimas <i>EcoRI</i> e <i>XhoI</i>	46
Tabela 2 - “ <i>Blast</i> ” do <i>contig</i> completo do clone P117.E10 com as sequências do NCBI, através da ferramenta BLASTN.....	62
Tabela 3 - Predição das 25 ORFs identificadas no <i>contig</i> do clone P117.E10 através ORF Finder (NCBI).....	63
Tabela 4 - Análise, através do BLASTP (Non-redundant protein sequence - nr), das quatro ORFs (8, 16, 17 e 21) que são possíveis codificadoras de enzimas lipolíticas encontradas no <i>contig</i> P117.E10.....	66
Tabela 5 - “ <i>Blast</i> ” do <i>contig</i> completo do clone P132.D09 com as sequências do NCBI, através da ferramenta BLASTN.....	76
Tabela 6 - Predição das 32 ORFs identificadas no <i>contig</i> do clone P132.D09 através ORF Finder (NCBI).....	76
Tabela 7 - Análise, através do BLASTP (Non-redundant protein sequence - nr), da ORF30 que é uma possível codificadora de enzima lipolítica, encontrada no <i>contig</i> P132.D09	80

LISTA DE ABREVIATURAS

BSA – albumina de soro bovino

DNA – ácido desoxiribonucleico

dNTP – desoxirribonucleotídeos fosfatados

EDTA – ácido etilenodiaminotetracético

IPTG – isopropyl β -D-1-thiogalactopyranoside

p/v – peso por volume

PCR – reação em cadeia da polimerase

pH – $\log [H^+]$

RNA – ácido ribonucleico

RNAse – ribonuclease

TEB – solução tampão Tris-ácido bórico-EDTA

Tris – Tris[hidroximetil]aminometano

UV – luz ultravioleta

v/v – volume por volume

X-GAL – 5-bromo-4-cloro-3-indolil- β -D-galactopiranosideo

LISTA DE UNIDADES

g – aceleração da gravidade

g – grama

kb – kilobase

L – litro

M – molar

mg – miligrama

mL – mililitro

μg – micrograma

μL – microlitro

μM – micromolar

ng – nanogramas

pb – pares de bases

U – unidades

V - volt

SUMÁRIO

1 INTRODUÇÃO	24
2 REVISÃO DE LITERATURA.....	27
2.1 Diversidade microbiana do solo.....	27
2.2 Abordagem metagenômica na diversidade de micro-organismos e prospecção de genes	28
2.3 Enzimas lipolíticas	29
2.4 A família com dobramento alfa/beta hidrolase	31
3 OBJETIVOS	33
4 MATERIAL E MÉTODOS	34
4.1 Fluxograma metodológico.....	34
4.2 Análise <i>in vitro</i> da produção de enzimas lipolíticas	36
4.3 Extração de DNA fosmidial e análise em gel de agarose	37
4.4 Digestão do DNA fosmidial dos clones selecionados.....	38
4.5 Construção das sub-bibliotecas	38
4.5.1 Reação de nebulização	39
4.5.2 Precipitação dos fragmentos e quantificação da nebulização	39
4.5.3 Reação de reparo	39
4.5.4 Reação de fosforilação	40
4.5.5 Gel Preparativo e eluição da banda	40
4.5.6 Reação de ligação do inserto ao vetor.....	41
4.5.7 Transformação bacteriana da célula competente <i>Escherichia coli</i> DH5a.....	42
4.5.8 Coleta e estoque dos subclones.....	43
4.5.9 Extração de DNA plasmidial em placa.....	43
4.6 Reação de sequenciamento.....	44
4.7 Análise das sequências.....	45
4.8 Clonagem das ORFs 16 e 17 em vetor pET28a	46
4.8.1 Reações de PCR	47
4.8.2 Purificação dos produtos de PCR	48
4.8.3 Sequenciamento dos produtos de PCR purificados	48

4.8.4 Digestão das extremidades	48
4.8.5 Ligação dos insertos em vetor de expressão pET28a	49
4.8.6 Transformação da célula competente <i>Escherichia coli</i> BL21	50
4.8.7 Coleta e estoque dos clones	51
4.8.8 Extração do DNA plasmidial pET28a e quantificação	51
4.8.9 Confirmação da clonagem	51
4.9 Análise <i>in vitro</i> dos clones pET28a + ORF16 e pET28a + ORF17	52
5 RESULTADOS E DISCUSSÃO.....	53
5.1 Análise <i>in vitro</i> da produção de enzimas lipolíticas	53
5.2 Digestão do DNA dos clones PL17.E10 e PI32.D09.....	56
5.3 Construção das sub-bibliotecas	57
5.4 Clone PI17.E10	61
5.4.1 Análise da sequência do contig	61
5.4.2 Análise das relações filogenéticas	67
5.4.3 Identificação dos motifs.....	73
5.5 Clone PI32.D09.....	75
5.5.1 Análise da sequência do contig	75
5.5.2 Análise das relações filogenéticas	81
5.5.3 Identificação dos motifs.....	83
5.6 Análise das ORFs encontradas nos clones PI17.E10 e PI32.D09 com sequências patenteadas.....	88
5.7 Clonagem das ORFs 16 e 17 em vetor pET28a	90
5.7.1 Reações de PCR e sequenciamento.....	90
5.7.2 Digestão dos produtos de PCR, ligação em vetor de expressão pET28a e transformação em célula competente <i>Escherichia coli</i> BL21	91
5.7.3 Extração do pET28a e quantificação.....	91
5.7.4 Confirmação da clonagem	92
5.7.5 Análise <i>in vitro</i> dos clones pET28a + ORF16 e pET28a + ORF17.....	92
6 CONCLUSÃO.....	94

REFERÊNCIAS	95
ANEXO A – Comparação do <i>contig</i> completo do clone PI17.E10, de 27485 pb, com as sequências do NCBI, usando a ferramenta BLASTN (nucleotide collection nr/nt)	101
ANEXO B – Comparação do <i>contig</i> completo do clone PI32.D09, de 35631 pb, com as sequências do NCBI, usando a ferramenta BLASTN (nucleotide collection nr/nt)	102

1 INTRODUÇÃO

O uso de micro-organismos e enzimas microbianas no processamento de materiais naturais tem uma longa tradição. Esta capacidade de exploração de biossintéticos microbianos, de qualquer modo, foi inconsciente e facilitada pela ubiquidade dos micro-organismos. Consequentemente, hoje em dia, as enzimas são usadas como “chaves” de ativação de componentes em detergentes, em vários alimentos industrializados e na manufatura de papel. A preferência pela química orgânica sintética para catálise enzimática e não pela clássica catálise química está baseada em enzimas de L-aminoácido, que apresentam características de quimio, régio e enantioseletividade. Adicionalmente, os produtos do metabolismo secundário microbiano são compatíveis tanto em relação à eficiência quanto com a sua interferência ambiental (LORENZ et al., 2002).

A biosfera é dominada por micro-organismos, mas a maioria ainda não foi estudada. Estima-se que 99% dos organismos procariontes não foram estudados em função dos métodos tradicionais de cultivo, que limitam o desenvolvimento em condições de laboratório (HENNE et al., 2000; HARDEMAN e SJOLING, 2007; ROH e VILLATTE, 2008). Os recentes avanços de pesquisas em ecologia molecular microbiana fornecem fortes evidências da existência de novas espécies de micro-organismos no meio ambiente, tanto em número quanto em diversidade (STEELE e STREIT, 2005), principalmente quando comparado com aqueles que são pouco cultiváveis em laboratório. A confirmação vem de estimativas da complexidade do DNA e da descoberta de várias sequências únicas dos genes *16S rRNA* e *18S rRNA* de várias fontes do ambiente (RONDON et al., 2000; SILVEIRA et al., 2006).

Para acessar o recurso genético da vasta maioria das espécies microbianas que há muito tempo têm escapado dos exames minuciosos científicos, uma nova abordagem denominada metagenoma começou a ser usada (HANDELSMAN et al., 1998). Esta estratégia abrange a extração direta do DNA genômico de amostras do ambiente e clonagem do material obtido (RONDON et al., 2000). Os resultados são bibliotecas metagenômicas, que servem como uma base para examinar vias metabólicas; analisar a diversidade microbiana; e identificar genes codificadores de proteínas de interesse biotecnológico. Estes genes podem ser prospectados através de sondas específicas, (hibridização de DNA-DNA); testes *in vitro* utilizando substratos

específicos; e protocolos de PCR, tendo como parâmetro sequências similares ou regiões conservadas.

Prévios estudos têm demonstrado que a análise metagenômica oferece uma combinação quase ilimitada para encontrar novos genes codificadores de produtos gênicos relevantes como lipases e esterases (HENNE et al., 2000; CIESLINSKI et al., 2009; WU e SUN, 2009; COUTO et al., 2010; HU et al., 2010). Por exemplo, nos últimos dez anos vários genes codificadores de enzimas lipolíticas puderam ser identificados em bibliotecas metagenômicas advindas de várias amostras ambientais como solo (HENNE et al., 2000; RONDON et al., 2000; LEE et al., 2004), água de lagoa e lago (REES et al., 2003; RANJAN et al., 2005), de mar (CHU et al., 2008), de rio (WU e SUN, 2009), sedimento de manguezais (COUTO et al., 2010) e campos termais (RHEE et al., 2005). Entre as descobertas, Hu et al. (2010) encontraram no sedimento marinho do mar do sul da China um gene codificador de esterase, o qual é importante candidato à aplicações industriais.

Já há seis anos, o comércio global das indústrias de enzimas está estimado em 2,3 bilhões de dólares, cujo principal lucro está dividido em detergentes (US\$ 789 milhões), aplicações alimentícias (US\$ 634 milhões), agricultura (US\$ 237 milhões), e outras aplicações incluindo enzimas para produção de tecidos e produtos químicos (LORENZ e ECK, 2005).

Dentro deste comércio, as enzimas lipolíticas estão atraindo enorme atenção por causa dos seus potenciais biotecnológicos (VAKHLU e KOUR, 2006). A maioria das lipases usadas nas indústrias são enzimas microbianas, de origem fúngica (*Candida* sp., *Geotrichium* sp., *Rhizopus* sp., etc) ou bacteriana (*Pseudomonas* sp., *Bacillus* sp., *Staphylococcus* sp., etc) (JAEGER; DIJKSTRA; REETZ, 1999). Por serem enzimas versáteis e amplamente utilizadas a expectativa é que as lipases futuramente sejam tão importantes como catalisadores industriais, o quanto são atualmente as proteases e carboidrases (TREVISAN, 2004). Um exemplo disto é a grande atenção destinada a esta enzima para a produção de biodiesel, através da transesterificação de gorduras, uma alternativa do combustível petróleo para os problemas ambientais (GAO et al., 2009).

As lipases de origem microbiana são economicamente atrativas por serem: facilmente biodegradáveis, causando menor impacto ambiental; atuarem geralmente em condições amenas (pH próximo à neutralidade e temperaturas de 20 a 40 °C); e serem quimio-seletivas, ou seja, atuam em um único tipo de grupo funcional, proporcionando para a indústria farmacêutica, por

exemplo, a obtenção de fármacos com efeitos colaterais reduzidos (JAEGER; DIJKSTRA; REETZ, 1999; VAKHLU e KOUR, 2006).

As enzimas derivadas de micro-organismos disponíveis para uso industrial são fontes de síntese de novos e interessantes produtos, permitindo assim, diversas aplicações com alto ganho nos processos biotecnológicos; maior eficiência, com consequentemente redução de gastos; menor necessidade de utilização de produtos químicos cáusticos e perigosos, o que gera vantagem ao meio ambiente (LORENZ e ECK, 2005). No entanto, quanto aos processos industriais de lipases, o maior problema está relacionado ao custo das etapas necessárias para a purificação que, em geral, provocam perdas na atividade das enzimas e aumentam o custo final, entretanto, devido aos recentes avanços na tecnologia de engenharia genética e de modificação e imobilização de lipases há perspectivas na mudança deste quadro (SAXENA et al., 2003; KOBLITZ e PASTORE, 2004).

A fim de bioprospectar genes de interesse biotecnológico, dentre outras aplicações, foi construída no Laboratório de Bioquímica de Micro-organismos e Plantas (LBMP), na UNESP – Campus de Jaboticabal, uma biblioteca fosmidial de DNA metagenômico a partir de um consórcio microbiano degradador de óleo diesel. Esta biblioteca foi confeccionada por Paixão et al. (2010) a partir de um consórcio microbiano extraído do solo com contaminação permanente, por um período de 15 anos, de uma antiga fábrica de lubrificantes localizada na cidade de Ribeirão Preto - SP. A biblioteca possui aproximadamente 4500 clones e vários genes de importância biotecnológica estão sendo prospectados, entre eles, genes codificadores de enzimas lipolíticas. O consórcio microbiano, do qual foi construída a biblioteca, apresenta uma gama de micro-organismos degradadores de óleo diesel (PAIXÃO et al., 2010), e nela almeja-se obter enzimas lipolíticas com características distintas das já descritas e disponíveis na literatura.

2 REVISÃO DE LITERATURA

2.1 Diversidade microbiana do solo

O solo constitui-se, indubitavelmente, num dos maiores reservatórios de carbono orgânico da Terra e num dos mais importantes habitats para micro-organismos, principalmente procariotos. A abundância do carbono procariótico e outros elementos sugere que cerca de metade do protoplasma vivo da Terra seja de origem microbiana. O número de micro-organismos e suas biomassas coletivas variam dentro e entre os diferentes tipos e condições dos solos, sendo o grupo das bactérias o mais numeroso (WHITMAN; COLEMAN; WIEBE, 1998).

Antes, com o desenvolvimento de técnicas de cultivo puro, os micro-organismos puderam ser estudados individualmente e caracterizados, principalmente, por critérios nutricionais. Entretanto, a abordagem do cultivo limitou seriamente a avaliação taxonômica e filogenética, como estimativa da diversidade microbiana, devido à falha de cultivo da maioria dos micro-organismos pelos métodos convencionais (HENNE et al., 2000; HARDEMAN e SJOLING, 2007).

Em função destas limitações dos métodos tradicionais de cultivo e isolamento, técnicas moleculares baseadas na amplificação por PCR do gene *16S rRNA* têm aumentado o conhecimento sobre a diversidade microbiana dos solos (SILVEIRA et al., 2006; PEDRINHO et al., 2009; PAIXÃO et al., 2010), assim como, outras ferramentas moleculares que vêm reduzindo essas limitações e revelando novas perspectivas sobre a diversidade microbiana (RONDON et al., 2000).

Dentro deste contexto, de acordo com estudos realizados, a utilização de ferramentas moleculares possibilitou um melhor entendimento da comunidade microbiana do petróleo (HAMME; SINGH; WARD, 2003; PAIXÃO et al., 2010), e vários micro-organismos estão sendo identificados através de métodos independentes de cultivo, dentre os quais pode-se citar a *Pseudomonas putida*, que apresenta vias para a degradação de petróleo nos solos contaminados. Este dado é de grande importância na procura de produtos biotecnológicos, produtos metabólicos e entendimento da diversidade microbiana, especialmente de bactérias ainda não cultivadas.

2.2 Abordagem metagenômica na diversidade de micro-organismos e prospecção de genes

A fim de acessar os micro-organismos não cultiváveis em laboratório, a abordagem metagenômica reúne várias técnicas moleculares significantes desenvolvidas no último século, tornando possível aos cientistas investigarem melhor a ecologia do reino procarioto e desvendar o vasto potencial biológico que representa a população procariótica (STEELE e STREIT, 2005). Os avanços tecnológicos no isolamento do DNA que abriram o caminho para o estudo metagenômico, fizeram com que as sequências quase completas de micro-organismos não cultiváveis tornassem acessíveis através da construção de complexas bibliotecas (ROH e VILLATTE, 2008). Com o desenvolvimento de bibliotecas metagenômicas, a diversidade genética de micro-organismos em determinado meio pôde ser explorada e novos genes que produzem produtos biotecnológicos puderam ser encontrados (HENNE et al., 2000; RONDON et al., 2000; HONG et al., 2007).

Recentes estudos estão sendo realizados para a criação de bibliotecas metagenômicas advindas de ambientes que estão sujeitos a condições extremas, onde há probabilidade de se encontrar enzimas que funcionam sob tais condições (ELEND et al., 2006; HARDEMAN e SJOLING, 2007; ROH e VILLATTE, 2008; CIESLINSKI et al., 2009; WU e SUN, 2009). Um exemplo é o isolamento de enzimas termofílicas de uma biblioteca metagenômica construída com DNA do solo da Tailândia, onde a temperatura do solo a 10-20 cm da superfície é 70 °C. Nesta biblioteca foram isoladas uma nova fosfolipase e uma esterase, que mantiveram a atividade em temperaturas que ficavam entre 50-75 °C, com temperatura ótima de 70 °C. Além disto, estas enzimas demonstraram atividade com substrato com 10 átomos de carbono, fato não comum entre as fosfolipases e esterases. Os resultados sugerem que estas são novas enzimas e com grande potencial em aplicações biotecnológicas (TIRAWONGSAROJ et al., 2008).

Da mesma forma, também foram construídas bibliotecas metagenômicas advindas de ambientes de baixa temperatura (HARDEMAN e SJOLING, 2007; ROH e VILLATTE, 2008; CIESLINSKI et al., 2009; HU et al., 2010), a fim de procurar lipases com atividade ótima nestas condições. As aplicações para lipases que atuam em baixas temperaturas são diversas, como: aditivo de detergentes, diminuição de custos energéticos envolvidos na catálise e biorremediação em solos ou águas de ecossistemas frios (HARDEMAN e SJOLING, 2007).

2.3 Enzimas lipolíticas

Os procariotos produzem diferentes classes de enzimas lipolíticas, que podem ser classificadas em oito famílias, baseado nas sequências conservadas, *motifs*, e propriedades biológicas, denominadas: verdadeiras lipases, GDSL, lipase hormônio-sensitiva (HSL) e as famílias III, V-VIII (AIRPIGNY e JAEGER, 1999).

De acordo com estes autores, as verdadeiras lipases (família I) apresentam seis subfamílias, sendo que, na maioria há a presença do pentapeptídeo conservado <Gly-Xaa-Ser-Xaa-Gly>. Dentro desta família, as subfamílias I.1 e I.2 compartilham homologia na posição de dois resíduos de aspartato e dois de cisteína, que estão envolvidos na formação do sítio de ligação ao Ca^{+2} e pontes dissulfeto, respectivamente. Acredita-se que por estes resíduos estarem perto dos sítios catalíticos da histidina e aspartato, eles estejam envolvidos na estabilização do centro ativo das enzimas. A família II não apresenta o pentapeptídeo conservado da família I, ao contrário apresenta o *motif* <Gly-Asp-Ser-(Leu)> sendo, por isto, conhecida como família GDSL. Os membros da família III dispõem do dobramento regular alfa/beta hidrolase e possuem a típica tríade catalítica, além de exibirem 20% de identidade com sequências de aminoácidos do PAF-AH humano (Acetil-hidrolase do fator ativador de plaquetas).

Já a família IV exibe similaridade com lipases hormônio-sensitiva de mamíferos, por isto são conhecidas como HLS. Os membros desta família segundo Jaeger, Dijkstra e Reetz (1999) apresentam um sítio conservado HGGG e logo após, o pentapeptídeo GDSAG localizado no N-terminal da proteína. A família V possui membros com estruturas similares à dehalogenases, haloperoxidases e epoxide hidrolases, exibindo folha alfa/beta hidrolase como característica da estrutura terciária, e de acordo com Arpigny e Jaeger (1999), os blocos conservados GXSMGG, PTLV e GH são característicos dos membros desta família.

Com massa molecular entre 23-26 kDa, as enzimas presentes na família VI estão entre as menores esterases conhecidas. O sítio ativo destas enzimas é um dímero, sendo que, a subunidade possui dobramento alfa/beta hidrolase e a clássica tríade catalítica <Ser-Asp-His>. Quanto à família VII, grande parte das enzimas bacterianas desta família divide homologia (30% de identidade, 40% de similaridade) com sequências de aminoácidos de acetilcolina esterases de eucariotos e com carboxilesterases de intestino e fígado. Segundo Arpigny e Jaeger (1999), a oitava família é constituída por enzimas com 380 resíduos de aminoácidos, as quais

compartilham similaridade com várias β -lactamases da classe C. As esterases da família VIII possuem o sítio ativo semelhante ao das β -lactamases da classe C, o qual é constituído do *motif* <Ser-Xaa-Xaa-Gly>, no entanto, para outros autores o *motif* que as caracteriza é outro, sendo necessária mais informações sobre esta família.

Apesar da classificação abrangente proposta por Arpigny e Jaeger (1999), novas lipases e esterases estão sendo identificadas em diferentes amostras ambientais devido à baixa identidade entre as existentes e as que estão sendo descobertas, como por exemplo: a LipG (LEE et al., 2006); EstA (CHU et al., 2008); LipEH166 (KIM et al., 2009); e EstY (WU e SUN, 2009). De acordo com Ranjan et al. (2005), a busca incessante por lipases e esterases irá aumentar a diversidade das enzimas lipolíticas e consequentemente o número de famílias.

As enzimas lipolíticas podem ser classificadas em três principais grupos de acordo com especificidade do substrato: lipases (EC 3.1.1.3), esterases/carboxiesterases (EC 3.1.1.1) e fosfolipases (JAEGER et al., 1999). As lipases (EC 3.1.1.3) catalisam a hidrólise dos triacilglicerídeos com ácidos graxos de cadeia longa (≥ 10 átomos de carbono), e têm máxima atividade na interface aquosa/não aquosa, sendo que a atividade está associada com mudanças conformacionais no sítio do substrato. As esterases (EC 3.1.1.1) agem em substratos solúveis em água e catalisam a hidrólise de estergliceróis com cadeia acil curta (≤ 10 átomos de carbono) em glicerídeos parciais e ácidos graxos. As fosfatolipases catalisam a hidrólise de vários tipos de fosfolipídeos (LEE et al., 2004; TIRAWONGSAROJ et al., 2008).

Entre as bactérias produtoras de lipases deve-se dar especial atenção às bactérias do gênero *Pseudomonas*, que são especialmente interessantes para a biotecnologia (JAEGER; DIJKSTRA; REETZ, 1999). As lipases decorrentes de *Pseudomonas* apresentam uma característica especial e não comum entre as lipases produzidas por micro-organismos, elas são termoresistentes e apresentam atividade em pH alcalino (MARTINEZ e SOBERÓN-CHÁVEZ, 2001). Várias espécies do gênero *Bacillus* também estão entre as bactérias produtoras de lipase e apresentam um sistema lipolítico bastante estudado para aplicações biotecnológicas. Em função do aumento no interesse por lipases bacterianas e novas lipases oriundas de *Bacillus*, cientistas estão investigando o mecanismo destas para a produção de lipase extracelular (RUIZ et al., 2002).

A atividade lipolítica decorrente da expressão de um gene pode ser avaliada através de análises colorimétricas. Métodos de detecção usam emulsões de tributirina ou trioleína com o

meio de cultura em placas de petri (TREVISAN, 2004; ROH e VILLATTE, 2008; WU e SUN, 2009; COUTO et al., 2010; HU et al., 2010). A produção de lipase é indicada pela formação de halos límpidos ao redor das colônias cultivadas em placas de petri com meio de cultura sólido contendo tributirina como um substrato indicador (VOGET et al., 2003). Para buscar atividade lipolítica em bibliotecas metagenômicas, os meios de culturas são emulsificados com 1% de tributirina e os clones que apresentarem atividade hidrolítica com a emulsificação de tributirina são selecionados depois de dois dias de cultivo a 37 °C (LEE et al., 2004).

Como exemplo, a biblioteca metagenômica de solo de uma floresta da Coreia, com 31.000 clones, obteve um total de sete clones com atividade lipolítica, que foi selecionado através da hidrólise de tributirina (HONG et al., 2007). Recentemente, foi realizado outro ensaio bioquímico com tributirina em biblioteca metagenômica de sedimento marinho do sul da China, de 40.000 clones, e 19 tiveram atividade, sendo que, um clone em especial apresenta características interessantes para aplicação industrial (HU et al., 2010).

2.4 A família com dobramento alfa/beta hidrolase

As lipases e esterases são representantes da família com dobramento alfa/beta hidrolase, assim como as proteases, dehalogenases, peroxidase e epóxido hidrolase, o que faz desta família uma das mais versáteis e com dobramentos protéicos mais conhecidos.

As enzimas que constituem esta família não compartilham similaridade de sequências e substratos, no entanto, apresentam uma preservação na disposição dos sítios catalíticos, homologia estrutural, sugerindo um possível ancestral comum (OLLIS et al., 1992). O dobramento das alfa/beta hidrolases é caracterizado por um núcleo de oito β -folhas conectadas por α -hélices, para dar a disposição $\alpha/\beta/\alpha$ (Figura 1). Na maioria dos membros da família as β -folhas são paralelas, mas alguns mostram uma inversão na ordem das β -folhas, tendo por resultado a orientação antiparalela. Para manter a maquinaria catalítica conservada grandes inserções foram toleradas, incluindo alguns resíduos de aminoácidos ou até mesmo um domínio extra completo, dando aos membros desta família uma habilidade para a adaptação e evolução (NARDINI e DIJKSTRA, 1999).

O dobramento das alfa/beta hidrolases pode fornecer um arcabouço estável para os sítios ativos de uma grande variedade de enzimas. Os resíduos catalíticos são sempre constituídos de

uma tríade altamente conservada: o nucleofílico (serina, cisteína ou ácido aspártico) posicionado após a folha $\beta 5$; um resíduo ácido, posicionado quase sempre após a folha $\beta 7$; e uma histidina, que é um resíduo absolutamente conservado e está situado após a última β folha (NARDINI e DIJSTRA, 1999). O membro nucleofílico está sempre localizado em uma curva acentuada, chamada “cotovelo nucleofílico”, onde facilita a aproximação com o substrato. Esta geometria do “cotovelo nucleofílico” contribui na formação do sítio de ligação do oxi ânion, que é necessário para estabilizar o estado transitório durante a hidrólise. Este sítio é identificado por uma sequência Sm-X-Nu-X-Sm (Sm=pequeno resíduo, X=qualquer resíduo, e Nu=nucleofílico). O membro ácido da tríade catalítica está localizado em uma curva reversa, na folha $\beta 7$. Já a histidina é o único resíduo da tríade que é absolutamente conservado, entretanto, a forma e o comprimento do local que contém a histidina pode diferir consideravelmente entre os vários membros da família.

Figura 1. Estrutura secundária do dobramento das alfa/beta hidrolases. α -hélices e β -folhas estão representadas através dos cilindros brancos e setas cinzas, respectivamente. A localização da tríade catalítica está indicada pelos círculos pretos.

Fonte: Nardini e Dijkstra (1999).

3 OBJETIVOS

Objetivo geral

Detectar clones produtores de enzimas lipolíticas em biblioteca metagenômica de DNA de um consórcio microbiano degradador de óleo diesel e caracterizar os genes codificadores de enzimas lipolíticas.

Objetivos específicos

- Realizar análise *in vitro*, em meio de cultivo sólido com tributirina, nos clones da biblioteca metagenômica para detecção de enzimas lipolíticas;
- Realizar a subclonagem do(s) fragmento(s) inserido(s) no(s) clone(s) positivo(s) selecionado(s);
- Sequenciar a(s) sub-biblioteca(s);
- Analisar as sequências da(s) sub-biblioteca(s) e obter a sequência completa do fragmento metagenômico de cada clone positivo;
- Identificar e selecionar as sequências codificadoras de enzimas lipolíticas;
- Construir árvores filogenéticas e analisar as sequências de aminoácidos;
- Identificar a família lipolítica a qual o(s) gene(s) encontrado(s) pertence;
- Selecionar ORFs para subclonagem em pET28a.

4 MATERIAL E MÉTODOS

4.1 Fluxograma metodológico

O presente trabalho foi realizado no Laboratório de Bioquímica de Micro-organismos e Plantas (LBMP), localizado na Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de Jaboticabal. As etapas desenvolvidas deste projeto podem ser visualizadas no fluxograma metodológico abaixo (Figura 2).

Figura 2. Fluxograma metodológico identificando as etapas desenvolvidas neste projeto.

4.2 Análise *in vitro* da produção de enzimas lipolíticas

Todos os clones da biblioteca metagenômica do consórcio microbiano degradador de óleo diesel foram submetidos à análise *in vitro* com tributirina em placa de petri, a fim de realizar a seleção por atividade lipolítica. No entanto, como ensaio inicial foi utilizado a célula EPI300™-T1R *Escherichia coli* não transformada do Copy Control™ HTP Fosmid Library Production Kit (EPICENTRE Biotechnologies, Madison, Wisconsin, Estados Unidos da América), a mesma que foi utilizada na confecção da biblioteca metagenômica, a fim verificar a viabilidade como controle negativo. Para isto, utilizou-se o meio cultura Luria-Bertani (LB) [1% de triptona (p/v), 0,5% de extrato de levedura (p/v), 1% de NaCl (p/v) e 1,5% de ágar (p/v) (pH 7,0)], com adição de 1% de tributirina (v/v) e 1% de goma arábica (p/v). O meio foi autoclavado à 121 °C por 20 minutos a 1 atm e resfriado, para inversão em placa de petri 90 mm X 15 mm. Este ensaio foi realizado com e sem a adição de 0,001% de Rhodamine-B (v/v).

Para as análises *in vitro* dos clones da biblioteca, estocados à -80 °C, foi utilizado o meio de cultura LB para recuperação das células viáveis. O meio foi autoclavado a 121 °C por 20 minutos a 1 atm e logo após o resfriamento, acrescentou-se 0,00125% de cloranfenicol (v/v), sendo 70 mL do mesmo adicionado em placa de petri 150 mm x 15 mm. Posteriormente, as células recuperadas foram usadas para a inoculação em meio de cultura LB suplementado com: 1% de tributirina (v/v); 1% de goma arábica (p/v); 0,00125% de cloranfenicol (v/v); e 0,001% de arabinose (v/v), que foi adicionada ao meio para promover o alto número de cópias do fosmídeo através do sistema *oriV/trfA* (WILD; HRADECNA; SZYBALSKI, 2002). Também foi realizado outro ensaio bioquímico com a mesma composição do meio de cultura descrita acima, mas com a adição de 0,001% de Rhodamine-B (v/v) para facilitar a visualização da formação dos halos.

Durante o preparo dos meios de cultura, a tributirina foi adicionada e emulsionada (LEE et al., 2004) em um aparelho de ultrassom Branson Sonifier 250, duty cycle 30% em 3 ciclos de 15 segundos. O meio foi autoclavado, nas mesmas condições já descritas, e adicionado nas placas de petri 150 mm x 15 mm. Os clones foram cultivados a 37 °C, por 3 dias na estufa.

A análise dos ensaios foi feita diariamente através da observação do desenvolvimento dos halos lípidos ao redor das colônias, que caracteriza a hidrólise do triacilglicerídeo em glicerol e ácido butírico (Figura 3).

4.4 Digestão do DNA fosmidial dos clones selecionados

Uma alíquota do DNA dos clones selecionados foi digerida com três enzimas de restrição, a fim de verificar se os clones são diferentes entre eles e estipular o tamanho do inserto clonado no fosmideo. As enzimas utilizadas foram: *EcoRI*, *HindIII* e *BamHI*.

Para a primeira enzima foi feita de acordo com a reação: 1 µL de DNA [3540,8 ng/µL]; 1 µL *EcoRI* (10U/µL, Fermentas) e 1 µL Tampão 10X *EcoRI* [50 mM Tris-HCl (pH 7,5), 10 mM MgCl₂, 100 mM NaCl, 0,02% Triton X-100 e 0,1 mg/mL BSA]; e água Milli-Q ultra-pura para completar 10 µL. A reação foi incubada a 37 °C por 2 horas e depois a 65 °C por 20 minutos para inativação da enzima. Para a enzima *HindIII* FastDigest foi feita a reação: 1 µL DNA [3540,8 ng/µL]; 1 µL *HindIII* FastDigest (Fermentas) e 1 µL Tampão 10X FastDigest (cat. no. FD0504, Fermentas); e água Milli-Q filtrada para completar 10 µL. A reação foi submetida a 37 °C por 10 minutos, seguida a 80 °C por 10 minutos, para inativação da enzima. Com a enzima *BamHI* FastDisget (cat. no. FD0054, Fermentas) foi feita à mesma reação descrita para a *HindIII*, no entanto, a enzima foi submetida a 37 °C por 5 minutos e depois foi inativada a 80 °C por 5 minutos.

O volume total de cada digestão foi aplicado em gel de agarose 1% contendo brometo de etídeo (0,5 mg/µL), conforme descrito no item 4.3, a fim de visualizar e comparar o perfil de restrição entre os clones.

4.5 Construção das sub-bibliotecas

Os clones selecionados para o estudo tiveram seu DNA fosmidial extraído conforme descrito no item 4.3. O DNA fosmidial foi utilizado para a construção de sub-biblioteca *shotgun* em vetor pUC19 desfosforilado e digerido com *SmaI*, para possibilitar o acesso à informação genética por meio de sequenciamento do DNA.

4.5.1 Reação de nebulização

O DNA dos clones foi nebulizado para gerar fragmentos de tamanho entre 1 a 3 kb para serem sub-clonados. Para cada subclonagem foi utilizado o mesmo procedimento, descrito abaixo.

Foram utilizados 140 μL (aproximadamente 300 μg) de DNA; 100 μL Tris-HCl 1M pH 8,0, filtrado em filtro Millipore 0,22 μm (Billerica, Massachusetts, Estados Unidos da América); 30 μL MgCl_2 1M, filtrado em filtro Millipore 0,22 μm ; 1 mL glicerol 50%; e 730 μL de água milli-Q (reação para 2 mL). A nebulização foi feita com pressão de nitrogênio por 16 segundos a 3 KgF/cm^2 , e após, a amostra foi centrifugada a 4 °C a 162 xg durante 2 minutos para precipitar o material. O volume total recuperado da nebulização de cada clone foi dividido em alíquotas de 700 μL para tubos tipo “eppendorfs” de 1,5 mL.

4.5.2 Precipitação dos fragmentos e quantificação da nebulização

À cada alíquota do DNA nebulizado foram adicionados 70 μL de acetato de sódio 3M pH 5,2 e 700 μL de isopropanol 100%, as amostras foram homogeneizadas e incubadas por 22 horas a -20 °C. Após a incubação a suspensão foi centrifugada a 4 °C, 15000 xg, por 25 minutos, e o sedimento obtido foi lavado com 1 mL de etanol 80% e centrifugado novamente nas mesmas condições já descritas, por 15 minutos. O sedimento foi deixado a temperatura ambiente para secar e em seguida ressuspensionado em 20 μL de água ultra-pura. Depois de ressuspensionado, o DNA foi incubado em banho-maria a 37 °C por 1 hora e 30 minutos. O volume das alíquotas foi reunido em um único tubo de 1,5 mL para cada clone. Uma alíquota do DNA nebulizado de cada clone foi quantificado em gel de agarose 0,8% a 80 V por 1 hora e 30 minutos, conforme descrito no item 4. 3.

4.5.3 Reação de reparo

Após a obtenção dos fragmentos de DNA estes foram submetidos ao reparo das extremidades. Para isto, as enzimas T4 DNA polimerase e Klenow DNA polimerase foram utilizadas para a incorporação de nucleotídeos complementares livres, nos terminais 3' da dupla

fita de DNA. Além disso, a T4 DNA polimerase apresenta atividade exonucleásica 3' → 5', que degrada extremidades que se sobressaem nos terminais 3' (SAMBROOK et al., 1989).

Na reação de reparo de cada clone foram utilizados: 55 µL de DNA [10ng/µL]; 10 µL de Tampão NEB2 10X [50 mM NaCl, 10 mM Tris-HCl, 10 mM MgCl₂ e 1 mM Dithiothreitol, pH 7,9, (New England Biolabs, Ipswich, Massachusetts, Estados Unidos da América)] ; 1 µL BSA 100X (20 mM KPO₄, 50 mM NaCl, 0,1 mM EDTA e 5% Glicerol, pH 7,0, New England Biolabs); 1 µL dNTP (10 mM, Fermentas); e 4 µL de T4 Polimerase (3 U/µL, New England Biolabs), o volume foi completado para 100 µL com água ultra-pura. Em seguida a reação foi incubada a 12 °C por 20 minutos. Após, foi adicionado 0,5 µL dNTP (10 mM, Fermentas) e 1 µL de Klenow (5 U/µL, New England Biolabs) e a reação foi submetida a 37 °C por 30 minutos, seguido de 75 °C por 20 minutos, para a inativação da enzima.

4.5.4 Reação de fosforilação

Depois de reparado, o DNA foi fosforilado através da transferência do gama fosfato do ATP para a extremidade 5'-OH do DNA (SAMBROOK et al., 1989).

Para isto, foi adicionado ao material reparado de cada clone: 12 µL de Tampão T4 DNA Ligase (50 mM Tris-HCl, 10 mM MgCl₂, 10 mM Dithiothreitol e 1 mM ATP, pH 7,5, New England Biolabs); 1 µL de T4 Polinucleotídeo Kinase (10 U/µL, New England Biolabs); e água ultra-pura para completar o volume da reação para 120 µL. A reação foi incubada a 37 °C por 30 minutos e logo após, a 65 °C por 20 minutos, para a inativação da enzima.

4.5.5 Gel Preparativo e eluição da banda

Todo o material reparado e fosforilado dos clones foi submetido à eletroforese preparativa em gel de agarose de baixo ponto de fusão 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.

Ao volume total de cada amostra (120 µL) foram adicionados 14 µL de azul de bromofenol 6X (3,3,5,5-tetrabromofenolsulfonftaleína). Durante a aplicação no gel de agarose foram aplicados: 6 µL do padrão molecular 1 kb DNA Ladder (Fermentas) na primeira canaleta; uma alíquota de 6 µL da amostra na próxima; e três canaletas foram unidas para a aplicação do

restante do material. Este procedimento foi realizado para cada amostra.

A eletroforese foi conduzida em tampão TBE 1X isento de brometo de etídio, a voltagem constante de 75 V durante 2 horas. Após, foi realizado, para cada amostra, um corte vertical no gel de agarose resultando duas porções de gel, uma contendo os insertos a serem recuperados e outra com o marcador e a amostra comparativa para corar. A porção do gel que continha o marcador foi corada com brometo de etídio e observada sob luz UV, sendo que, o marcador serviu para guiar na obtenção do inserto na porção não corada. As duas partes do gel foram colocadas lado a lado e a região contendo o inserto do tamanho desejado foi selecionada, cortada e retirada do gel restante. Esta parte retirada foi dividida em tubos de 1,5 mL com 400 mg de gel/tubo, aproximadamente, o restante do gel foi corado para observar se foi feito com êxito todo o processo de seleção de banda. A visualização foi feita em luz UV e documentado em fotodocumentador Gel Doc 1000 (Bio-Rad).

A eluição dos fragmentos de DNA metagenômico do gel de agarose foi feita com o Kit Geneclean III (BIO 101, Inc., San Diego, Califórnia, Estados Unidos da América), seguindo as instruções do fabricante. Uma alíquota de 3 µL de cada amostra eluída foi quantificada em gel de agarose 1% contendo brometo de etídeo (0,5 mg/µL), 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

A reação de ligação foi feita seguindo a razão: 300 ng de insertos para 100 ng de vetor pUC19 desfosforilado e digerido com *Sma*I, que apresenta aproximadamente 2.686 pares de bases (Figura 4). O cálculo da concentração inserto:vetor foi realizado de acordo com o Technical Manual pGEM-T and pGEM-T Easy Vector Systems (Promega), conforme demonstrado abaixo.

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

Bactérias do Departamento de Biologia Aplicada à Agropecuária da Faculdade de Ciências Agrárias e Veterinárias, Campus de Jaboticabal, UNESP.

Na reação foram utilizados 5 µL do DNA ligado e 200 µL da célula competente DH5α. A transformação foi feita através de choque térmico, assim, a reação ficou 20 minutos no gelo, logo em seguida foi colocada a 42 °C por 90 segundos e recolocada no gelo por 2 minutos. Depois da transformação bacteriana foram adicionados 970 µL de meio SOC (Tryptona, Extrato de levedura, NaCl 1M, KCl 1M, Mg²⁺ 2M e Glicose 2M) a fim de propiciar o cultivo da *Escherichia coli*, colocada sob agitação de 230 rpm a 37 °C, por 1 hora e 30 minutos. Após este período, uma alíquota de 85 µL da cultura foi distribuída, com alça de Drigalski esterelizada, em placas de petri contendo o meio LB. Durante a preparação deste meio foram adicionados 70 µg/mL de ampicilina e, após a solidificação, 100 µL de IPTG (isopropyl β-D-1-thiogalactopyranoside) [2,4% (p/v) em água] e 20 µL de X-GAL (5-bromo-4-cloro-3-indolil-β-D-galactopiranosídeo) [5% (p/v) em N, N'-dimetil formamida] foram espalhados sobre a superfície, tornando o meio seletivo para o cultivo de células transformadas. As placas ficaram em incubadas por 16 horas a 37 °C.

4.5.8 Coleta e estoque dos subclones

A coleta foi feita com palitos de madeira esterilizados, um para cada subclone, e logo após a coleta foram colocados em placas estéreis de poliestireno para 96 clones, que continham 120 µL de meio LB (suplementado com ampicilina para uma concentração final de 70 µg/mL). As placas foram incubadas a 37 °C por 22 horas em estufa. Depois do período de cultivo foram adicionados 110 µL glicerol 40% (v/v) em cada subclone e as placas foram estocadas a -80 °C.

4.5.9 Extração de DNA plasmidial em placa

Os subclones obtidos nas subclonagens foram cultivados em placas de cultivo de bactérias (96 poços), contendo 1,1 mL de meio LB suplementado com 70 µg/mL de ampicilina a 230 rpm, por 22 horas a 37 °C.

Após a incubação, as placas foram centrifugadas por 6 minutos, a 3220 xg, a 20 °C. O sedimento foi ressuspensionado em 140 µL de solução GET [Glicose 50 mM; Tris-HCl 23 mM (pH

8,0); EDTA 10 mM (pH 8,0)], através de agitação vigorosa por 2 minutos e repetiu-se a centrifugação anterior. Em seguida, as placas foram invertidas e o excesso de líquido escoado em papel toalha. Foram adicionados 60 µl da solução GET e o material foi agitado vigorosamente, à suspensão celular foram adicionados 2,5 µl RNase (10mg/mL) e o material ficou 5 minutos a temperatura ambiente. Após, 120 µl da solução de lise [NaOH 0,2N/ SDS 1% (p/v)] foram adicionados e as placas ficaram 10 minutos a temperatura ambiente. Posteriormente, foram adicionados 90 µL de acetato de potássio 3M e as placas foram incubadas a 90 °C por 30 minutos. Após incubação, o material foi resfriado em gelo por 7 minutos e centrifugado a 3220 xg, por 10 minutos, a 20 °C. O sobrenadante foi aspirado gentilmente e transferido para novas placas de polipropileno. Ao material foram adicionados 110 µL de isopropanol absoluto gelado e as placas foram seladas com adesivos resistentes a álcool, e invertidas 20 vezes. O material foi centrifugado a 3220 xg, por 45 minutos a 20 °C, e o sobrenadante descartado. O DNA precipitado foi lavado com 150 µL de etanol 70% (v/v) gelado e as placas foram centrifugadas por 5 minutos, a 3220 xg. O sobrenadante foi novamente descartado e o material ficou secando por 1 hora à temperatura ambiente. O DNA foi ressuscitado em 30 µL de água Milli-Q esterilizada e visualizado em gel de agarose 0,8% contendo brometo de etídeo (0,5mg/µL), conforme descrito no item 4.3.

4.6 Reação de sequenciamento

O DNA plasmidial extraído dos subclones das sub-bibliotecas foi sequenciado utilizando os oligonucleotídeos iniciadores de síntese M13 Forward e M13 Reverse, e as reações foram realizadas seguindo as condições: 1 µL DYEnamic ET Terminator Cycle Sequencing Kit (GE Healthcare, Londres, Inglaterra); 3 µL Tampão 5X (400 mM Tris-HCl pH 9,0; 10 mM MgCl₂); 10 pmoles do oligonucleotídeo (M13 forward ou M13 reverse); 100-120 ng de DNA e água Milli-Q esterilizada para completar o volume de 10 µL. Após, as reações de sequenciamento foram levadas ao termociclador, seguindo o programa: 1 passo de 96 °C por 2 minutos, 40 ciclos a 96 °C por 10 segundos, 52 °C por 20 segundos e 60 °C por 4 minutos, e um passo final a 10 °C.

As pontas dos fósforos foram sequenciadas de acordo com a reação: 2 µL de BigDye Terminator (Applied Biosystems, Carlsbad, Califórnia, Estados Unidos da América); 6 µL

Tampão 5X (400 mM Tris-HCl pH 9,0; 10 Mm MgCl₂); 10 pmoles do oligonucleotídeo [pCC2TM Forward (5' – GTACAACGACACCTAGAC – 3') ou pCC2TM Reverse (5' – CAGGAAACAGCCTAGGAA – 3')]; 1 µg de DNA fosmidial e água Milli-Q esterilizada para completar o volume de 20 µL. As reações foram levadas ao termociclador, seguindo o programa: 40 ciclos a 95 °C por 20 segundos, 50 °C por 15 segundos e 60 °C por 2 minutos, e um passo final a 4°C.

Após as reações, os fragmentos de DNA foram precipitados e os dNTPs, marcados por fluorescência, que não foram incorporados, foram eliminados por sucessivas lavagens. A precipitação do DNA amplificado, marcado pela PCR de sequenciamento, foi realizada através da adição de 80 µL de isopropanol 75% (v/v) às amostras, posteriormente, agitação cuidadosa em vortex e incubação à temperatura ambiente por 15 minutos, seguido de centrifugação a 20 °C por 45 minutos, a 3220 xg. Após, os sobrenadantes foram descartados e 150 µL de etanol 70% (v/v) foram adicionados. As amostras foram centrifugadas novamente por 10 minutos, seguindo as condições já descritas, sendo os sobrenadantes novamente descartados. O material foi secado por 30 minutos no fluxo laminar na ausência de luz.

Para a aplicação no sequenciador, as amostras foram ressuspensas em 9 µL de Hi-Di Formamide ABI PRISM (Applied Biosystems) e submetidas à desnaturação por 5 minutos a 95 °C. O sequenciamento foi realizado no aparelho ABI 3100 (Applied Biosystems).

4.7 Análise das sequências

Após o sequenciamento, os eletroferogramas das sub-bibliotecas foram analisados com o auxílio do programa PhredPhrap (EWING e GREEN, 1998; EWING et al., 1998), para a limpeza e remoção de vetores, e agrupamento das sequências *forward* e *reverse*. Quando as sequências são agrupadas baseadas na sua similaridade, elas geram sequências consenso conhecidas como *contigs*. Nos casos onde não há o agrupamento, sequências *forward* e *reverse* permanecem separadas gerando os *singlets*.

Com o auxílio de outro programa, ContGEN, foram selecionadas somente aquelas sequências que apresentaram mais de 400 bases com qualidade Phred ≥ 20 . Regiões iniciais contendo sequências de vetores foram extraídas com o auxílio do módulo de análise de sequências da ferramenta OC Identifier (CANTÃO; FERREIRA; LEMOS, 2007).

Primeiramente, o *contig* foi comparado as sequências depositadas no banco de nucleotídeos *National Center for Biotechnology Information* (NCBI), usando a ferramenta de alinhamento local BLASTN (ASTSCHUL et al., 1997). Isto foi realizado com o objetivo de identificar a similaridade com as sequências já depositadas dos diferentes tipos de micro-organismos, verificando se o *contig* sequenciado já foi em algum momento estudado, ou se é algo inédito.

Para a anotação gênica foi utilizado o programa Open Reading Frame Finder (ORF Finder) do NCBI, localizado no endereço eletrônico <http://www.ncbi.nlm.nih.gov/projects/gorf/> e as ORFs encontradas foram comparadas com sequências do banco de genes do próprio NCBI. A predição da função foi realizada através do BLASTX, no qual a sequência de aminoácidos, em formato FASTA, é comparado com sequências similares de proteínas depositadas no banco de proteína não redundante (NR).

Para as análises das relações filogenéticas, as sequências distintas de aminoácidos foram alinhadas usando ClustalW, através do programa BioEdit Sequence Alignment Editor. As árvores filogenéticas foram construídas com o auxílio do programa Mega 4. 1 (TAMURA et al., 2007) usando o algoritmo do vizinho mais próximo neighbor-joining (SAITO e NEI, 1987).

4.8 Clonagem das ORFs 16 e 17 em vetor pET28a

As ORFs 16 e 17 identificadas no clone p117.E10 foram selecionadas para clonagem em vetor de expressão pET28a, para isto, foram sintetizados pares de oligonucleotídeos iniciadores de síntese para ambas as ORFs com sítios de restrição para as enzimas *EcoRI* e *XhoI*, sendo que, no *forward* está o sítio para *EcoRI* e no *reverse* o sítio para *XhoI* (tabela 1).

Tabela 1 - Síntese dos pares de oligonucleotídeos iniciadores de síntese para as ORFs 16 e 17 presentes no clone P117.E10, com sítios de restrição para as enzimas *EcoRI* e *XhoI*.

Clone	ORF	Oligonucleotídeo iniciador de síntese	Enzimas
P117.E10	16	Forward: CTGAATTCTCCATGCCGCAGGTTTCAG Reverse: GGGCTCGAGCGTCACTCCGCC	Forward: <i>EcoRI</i> Reverse: <i>XhoI</i>
	17	Forward: GACGAATTCCCATGACCCAGCAGCAAGC Reverse: TATACTCGAGTGCTATTTCGGCGGCGGCCTTG	Forward: <i>EcoRI</i> Reverse: <i>XhoI</i>

4.8.1 Reações de PCR

Foram feitas reações de PCR com ambos os pares de oligonucleotídeos a fim de definir a temperatura de pareamento, assim, foram utilizadas as seguintes condições: 2 µL de Tampão PCR 10X [20 mM Tris-HCl (pH 8,4); 50 mM KCl, (cat. no. Y02028, Gibco, Carlsbad, Califórnia, Estados Unidos da América)]; 0,4 µL de dNTP (10mM, Fermentas); 0,32 µL de MgCl₂ (50mM, Gibco); 50 ng de DNA fosmidial do clone Pl17.E10; 12,5 pmoles dos oligonucleotídeos forward e reverse; 1U de Taq Polimerase; 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 2 minutos, 30 ciclos a 95 °C por 45 segundos, 35 °C – 61 °C por 45 segundos, 72 °C por 2 minutos, e um passo final de 72 °C por 5 minutos. A fim de abranger todo o gradiente do termociclador foram feitas doze repetições, que tiveram uma diferença de 3 °C de uma amostra para outra na temperatura de pareamento.

Após definir a temperatura de pareamento, foram realizadas reações de PCR com a *Pfu* DNA Polimerase, uma vez que esta enzima apresenta menor taxa de erro conhecida nas amplificações. Isto deve-se a sua atividade exonucleásica 3'→ 5', que lhe permite corrigir os erros inseridos durante a amplificação, garantindo confiabilidade do produto de PCR para clonagens em vetor de expressão.

Assim, para as reações com a enzima *Pfu* DNA Polimerase foram utilizadas as seguintes condições: 2 µL de Tampão 10X com MgSO₄ [200 mM Tris-HCl (pH 8,8), 100 mM (NH₄)₂SO₄, 100 mM KCl, 1 mg/mL BSA, 1% Triton X-100 (v/v), 20 mM MgSO₄ (Fermentas)]; 0,4 µL de dNTP (10mM, Fermentas); 50 ng de DNA fosmidial do clone Pl17.E10; 12,5 pmoles dos oligonucleotídeos forward e reverse; 1 µL da *Pfu* DNA Polimerase (2,5 U/µL, Fermentas); 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 35 °C para a ORF16 e 50 °C para a ORF17.

Uma alíquota de 3 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.

4.8.2 Purificação dos produtos de PCR

Após a confirmação da amplificação em gel de agarose, os produtos de PCR das ORFs 16 e 17 foram purificados com o Kit Abstract for Wizard SV Gel and PCR Clean-Up System (Promega), seguindo instruções do fabricante. Um alíquota de 3 µL de cada purificação foi analisada em gel de agarose 1% contendo brometo de etídeo (0,5 mg/µL), conforme descrito no item 4.3.

4.8.3 Sequenciamento dos produtos de PCR purificados

O produto purificado da ORF 16 e 17 foi sequenciado utilizando os seus respectivos oligonucleotídeos iniciadores de síntese, de acordo com a reação: 1 µL DYEnamic ET Terminator Cycle Sequencing Kit (GE Healthcare); 3 µL Tampão 5X (400 mM Tris-HCl pH 9,0; 10 mM MgCl₂); 10 pmoles do oligonucleotídeo (M13 forward ou M13 reverse); 50-70 ng de DNA e água Milli-Q esterilizada para completar o volume de 10 µL. As condições usadas foram as mesmas descritas no item 4.6.

4.8.4 Digestão das extremidades

Os produtos de PCR purificados foram digeridos com as enzimas de restrição *EcoRI* e *XhoI*, que flanqueiam suas extremidades.

Para o produto de PCR da ORF16 foi feita a seguinte reação: 1,7 µL do produto de PCR [128,9 ng/µL]; 1 µL da *EcoRI* FastDigest (Fermentas); 2 µL de Tampão 10X FastDigest (cat. no. FD0274, Fermentas); e água Milli-Q filtrada para completar 20 µL. Já para o produto de PCR da ORF17 foi feita a mesma reação, no entanto, foram utilizados 3,5 µL produto de PCR [63 ng/µL]. As reações de digestão foram submetidas a 37 °C por 30 minutos e após, a 80 °C por 20 minutos para inativação da enzima. Um 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 das duas digestões foi purificado conforme descrito no item 4.8.2 e os produtos de PCR digeridos das ORFs 16 e 17,

insertos, foram quantificados no NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific).

4.8.5 Ligação dos insertos em vetor de expressão pET28a

As reações de ligação foram feitas utilizando o vetor de expressão pET28a (Figura 5), visto que, o sistema pET foi desenvolvido para a clonagem e expressão de proteínas recombinantes em *Escherichia coli*. Este vetor apresenta um promotor T7 onde a T7 RNA polimerase atua quando induzida, sendo que, por esta enzima ser seletiva e ativa, faz com que quase todos os recursos da célula sejam convertidos para a expressão do gene alvo. Além disto, o vetor pET28 possui: duas sequências codificadoras de cauda de histidina, tanto no N-terminal, quanto no C-terminal da proteína; a sequência referente ao *operon* da lactose, *lacI*; o gene que confere resistência à kanamicina; e uma região com múltiplos sítios de clonagem.

Assim, a ligação do inserto ao vetor foi feito seguindo as condições: 2 µL do Tampão T4 DNA Ligase (50 mM Tris-HCl, 10 mM MgCl₂, 10 mM Dithiothreitol e 1 mM ATP, pH 7,5, New England Biolabs); 1,5 µL da enzima Ligase (400 U/µL, New England BioLabs); 1 µL do vetor pET28a desfosforilado e digerido com *EcoRI* e *XhoI* [55 ng/µL (Novagen, Gibbstown, New Jersey, Estados Unidos da América)]; 2 µL do inserto [aproximadamente 17,5 ng/µL]; e água Milli-Q ultra-pura, para completar a reação para 20 µL. A reação foi incubada a 16 °C, por 14 horas, aproximadamente.

Figura 5. Esquema ilustrativo apresentando características do vetor pET28a (Novagen).

4.8.6 Transformação da célula competente *Escherichia coli* BL21

As transformações foram feitas com a célula competente *Escherichia coli* BL21 também cedida pelo Prof. Dr. Manoel Victor Franco Lemos. Nas reações foram utilizados 15 µL do DNA ligado e 200 µL da célula competente BL21. A transformação foi feita através de choque térmico, assim, a reação ficou 5 minutos no gelo, logo em seguida foi colocada a 42 °C por 30 segundos e

recolocada no gelo por 2 minutos. Depois da transformação bacteriana foram adicionados 970 μL de meio SOC (Tryptona, Extrato de levedura, NaCl 1M, KCl 1M, Mg^{2+} 2M e Glicose 2M) e o cultivo foi colocado sob agitação de 250 rpm a 37 °C, por 1 hora e 30 minutos. Após este período, uma alíquota de 165 μL da cultura foi distribuída, com alça de Drigalski esterelizada, em placas de petri contendo meio LB com 50 $\mu\text{g/mL}$ de kanamicina. As placas ficaram em incubadas por 16 horas a 37 °C.

4.8.7 Coleta e estoque dos clones

Foram coletados clones de cada clonagem (ORF16 e ORF17). Os clones foram coletados com palitos de madeira e colocados em tubo do tipo falcon contendo 5 mL de meio LB com 50 $\mu\text{L/mL}$ kanamicina. Os tubos foram incubados sob agitação de 230 rpm, a 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.

4.8.8 Extração do DNA plasmidial do pET28a e quantificação

Um clone de cada construção pET28a + ORF16 e outro da construção pET28a + ORF17 teve seu DNA plasmidial extraído e quantificado, conforme descrito no item 4.3.

4.8.9 Confirmação da clonagem

Após a análise e quantificação dos DNAs plasmidiais extraídos foi feito a confirmação da clonagem através da digestão com as enzimas de restrição *EcoRI* e *XhoI*. As reações foram realizadas de acordo com o descrito no item 4.8.4 e a visualização foi feita em gel de agarose 1%, conforme mencionado no item 4.3.

4.9. Análise *in vitro* dos clones pET28a + ORF16 e pET28a + ORF17

Os clones pET28a + ORF16 e pET28a + ORF17 foram analisados *in vitro* com tributirina em placa de petri, a fim de verificar a formação de halos límpidos que representa a hidrólise do triacilglicerídeo em glicerol e ácido butírico.

Primeiramente, os clones pET28a + ORF16 e pET28a + ORF17 e a *Escherichia coli* BL21 contendo o vetor pET28a fechado (controle negativo), estocados à -80 °C, tiveram suas células recuperadas; em placa de petri contendo meio de cultura LB com 50 µL/mL de kanamicina, em estufa a 37 °C por 20 horas.

Após, uma alçada de cada cultura foi transferida para tubos de 2 mL do tipo eppendorf contendo 1,5 mL de meio LB até a cultura atingir a densidade ótica a 600 nm de 0,8. Uma alíquota de 30 µL desta cultura foi distribuída em placa de petri contendo o meio de cultura LB suplementado com: 1% de tributirina (v/v); 1% de goma arábica (p/v); 0,001% de Rhodamine-B (v/v); e 50 µL/mL de kanamicina, sendo adicionado 100 µL de IPTG (isopropyl β-D-1-thiogalactopyranoside) [2,4% (p/v) em água] após a solidificação do meio. O material foi cultivado a 37 °C por 3 dias na estufa e depois foi transferido a 4 °C.

5 RESULTADOS E DISCUSSÃO

5.1 Análise *in vitro* da produção de enzimas lipolíticas

As análises com a célula EPI300™-T1R *Escherichia coli* não transformada do Copy Control™ HTP Fosmid Library Production Kit (EPICENTRE Biotechnologies) possibilitaram verificar a ausência da formação de halo ao redor das colônias (Figura 6). Isto confirma a viabilidade desta célula como controle negativo nos ensaios com os clones da biblioteca e diminui a possibilidade de “falsos positivos”, que já foram encontrados em trabalhos de seleção por atividade funcional com outros genes de interesse biotecnológico em biblioteca metagenômica (JONES; SUN; MARCHESI, 2007).

Figura 6. Análise *in vitro* da célula EPI300™-T1R *Escherichia coli*, não transformada, em placa de petri 90 mm X 15 mm. Foi utilizado o meio de cultura LB suplementado com 1% de tributirina (v/v) e 1% de goma arábica (p/v). A- Sem adição de Rhodamine-B ao meio; B- Com adição de Rhodamine-B ao meio.

Os ensaios com tributirina foram realizados nos clones da biblioteca, 44 placas, e o cultivo de todos os clones foi acompanhado diariamente (Figura 7). Dos 4224 clones, 30 apresentaram formação de halo nítido, o que demonstra atividade lipolítica na hidrólise da tributirina.

A identificação de um alto número de clones com atividade lipolítica na biblioteca pode ser devido à cuidadosa análise em selecionar os clones, até mesmo quando apresentarem baixa atividade lipolítica. Este raciocínio deve-se ao fato de que alguns clones podem não ter boa expressão em *E. coli* devido: a dificuldade de expressão dos genes lipolíticos em hospedeiro heterólogo, em função do não reconhecimento do sistema regulador; ao requerimento de dobramento de proteínas e transportadores; e a toxicidade da proteína

lipolítica expressada (RANJAN et al., 2005). O alto número de clones com atividade lipolítica, comparado com demais trabalhos que também utilizaram a *E. coli* como hospedeiro, pode também estar relacionado a outros fatores, tais como: 1) a amostra ambiental (consórcio microbiano extraído de um solo contaminado com lubrificantes) da qual foi construída a biblioteca; 2) aos micro-organismos que tiveram efetivamente o seu DNA extraído e representaram a totalidade na construção da biblioteca; 3) a eficiência da ligação e transformação; e 4) ao método para selecionar os clones positivos.

Figura 7. Análise *in vitro* nas 44 placas, 4224 clones, que constituem a biblioteca metagenômica com o meio de cultura LB suplementado com 1% de tributirina (v/v); 1% de goma arábica (p/v); 0,001% de arabinose (v/v) e 0,00125% de cloranfenicol (v/v). O cultivo foi analisado diariamente, sendo possível verificar formação de halo, atividade lipolítica, em alguns clones.

Durante as análises de seleção por enzimas lipolíticas, os halos formaram-se no decorrer dos três dias de cultivo a 37 °C, mas ficaram ainda mais visíveis quando ficaram estocados a 4 °C, após o cultivo (Figura 8). Através deste ensaio foi possível verificar que os clones P17.E10 e P132.D09 tiveram nítida formação de halo.

Figura 8. Análise *in vitro* das placas 17 e 32 da biblioteca metagenômica com o meio de cultura LB suplementado com 1% de tributirina (v/v); 1% de goma arábica (p/v); 0,001% de arabinose (v/v) e 0,00125% de cloranfenicol (v/v). A e B- Clone PI17.E10 no terceiro dia de cultivo e estocado por 7 dias a 4 °C, respectivamente; C e D- Clone PI32.D09 no terceiro dia de cultivo e estocado por 7 dias a 4 °C, respectivamente.

A análise *in vitro* foi repetida com os 30 clones, no entanto, com a adição de Rhodamine-B ao meio de cultura, que facilitou a visualização e confirmação do resultado obtido no ensaio anterior (KOUKER e JAEGER, 1987) (Figura 9). Assim, com a análise de ambos os ensaios, foi possível selecionar os clones PL17.E10 e PI32.D09, dentre os 4224 clones, para a confecção das sub-bibliotecas e busca dos genes codificadores de enzimas lipolíticas. Os outros 28 clones, que também apresentaram formação de halo, apesar de não terem sido estudados neste trabalho, poderão futuramente contribuir na busca por novas enzimas lipolíticas.

Figura 9. Análise *in vitro* em placa de petri dos 30 clones da biblioteca metagenômica, previamente selecionados, em meio LB suplementado com 1% de tributirina (v/v); 1% de goma arábica (p/v); 0,001% de arabinose (v/v); 0,00125% de cloranfenicol (v/v); e 0,001% de Rhodamine-B (v/v). A- Clone PI17.E10; B- Clone PI32.D09; C- Controle negativo.

5.2 Digestão do DNA dos clones PL17.E10 e PI32.D09

A digestão do DNA fosmidial dos clones PL17.E10 e PI32.D09 com as enzimas *EcoRI*, *HindIII* e *BamHI* permitiu afirmar que eles são diferentes entre si, e possibilitou calcular o valor aproximado do inserto clonado em cada vetor: 27 kb para o clone PI17.E10 e 35 kb para o PI32.D09 (Figura 10).

Figura 10. Perfil eletroforético em gel de agarose 0,8% representando a restrição do DNA fosmidial dos clones P117.E10 e P132.D09 com as enzimas: *EcoRI* (Fermentas); *BamHI* (Fermentas) e *HindIII* (Fermentas). 1, 5, 9,13- Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2, 6, 10, 14- Marcador DNA Lambda *EcoRI/HindIII* (Fermentas); 3 e 4- DNA fosmidial não digerido dos clones P117.E10 e P132.D09, respectivamente; 7 e 8- DNA dos clones P117.E10 e P132.D09, respectivamente, digeridos com *EcoRI*; 11 e 12- DNA dos clones P117.E10 e P132.D09, respectivamente, digeridos com *BamHI*; 15 e 16- DNA dos clones P117.E10 e P132.D09, respectivamente, digeridos com *HindIII*.

5.3 Construção das sub-bibliotecas

O DNA fosmidial dos clones P117.E10 e P132.D09 extraído foi visualizado em gel de agarose 0,8% (Figura 11), e quantificado em NanoDropTM 1000 Spectrophotometer. A análise permitiu verificar a não fragmentação do DNA e a concentração suficiente para a realização da subclonagem, respectivamente. Para o clone P117.E10 a concentração foi de 2747,1 ng/μL e relação 260/280 de 1,92; e o clone P132.D10 apresentou concentração de 2235,7 ng/μL e relação de 1,90, que indica ausência de contaminação com proteínas.

Figura 11. Perfil eletroforético em gel de agarose 0,8% da extração do DNA fosmidial dos dois clones selecionados. 1- Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2- DNA do clone PI17.E10; 3- DNA do Clone PI32. D09.

As sub-bibliotecas dos dois clones PL17.E10 e PL32.D09 foram construídas iniciando-se através da seleção dos fragmentos, de 1 a 3 kb pela nebulização. Através da análise, em gel de agarose, de uma alíquota de cada material nebulizado, foi possível verificar que a nebulização de ambos foi bem sucedida, pois a concentração dos fragmentos gerados estava maior na região de interesse (Figura 12). A concentração do material nebulizado do clone PI17.E10 foi de 50 ng/μL e do clone PI32.D09 foi de 100 ng/μL, aproximadamente, de acordo com a relação ng/μL do marcador de tamanho molecular 1kb DNA Ladder (Fermentas).

A seleção da região no gel de agarose de baixo ponto de fusão foi bem realizada (Figura 13) e os fragmentos de 1 a 3 kb, já reparados e fosforilados, puderam ser selecionados. No entanto, a eluição não foi eficiente para a recuperação dos fragmentos dos dois clones, sendo a concentração dos materiais eluídos de aproximadamente 20 ng/μL, de acordo com a relação ng/μL descrita acima. A perda de fragmentos de DNA pode ser verificada através da comparação dos perfis eletroforéticos entre a Figura 12 (canaletas 2 e 4) e 13 (canaletas 2 e 4). Esta perda de concentração é comum nesta etapa durante a construção de bibliotecas, e está diretamente relacionada com a metodologia ou eficiência do Kit utilizado.

Figura 12. Perfil eletroforético da nebulização a 3 kgf por 16 segundos do clone P117.E10 e P132.D09 em gel de agarose 0,8%. 1 e 3- Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2- DNA fosmidial nebulizado do clone P117.E10; 4- DNA fosmidial nebulizado do clone P132.D09.

Figura 13. Perfil eletroforético do resultado da eluição dos clones P117.E10 e P132.D09, através do Kit GeneClean III (BIO 101, Inc.), em gel de agarose 0,8%. 1 e 3- Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2- DNA fosmidial eluído do clone P117.E10; 4- DNA fosmidial eluído do clone P132.D09.

Os fragmentos de DNA fosmidial, dos clones PL17.E10 e P132.D09, recuperados da eluição, foram inseridos em vetor pUC19 desfosforilado e digerido com *Sma*I. Posteriormente os plasmídeos recombinantes foram transformados na célula competente *Escherichia coli* DH5 α . Para cada uma das transformações foram coletados 672 subclones, ou seja, sete placas

de poliestireno (microplacas de 96 poços). Os subclones das sete placas de cada sub-biblioteca tiveram seu DNA plasmidial extraído e quantificado em gel de agarose, que possibilitou analisar a qualidade dos mesmos. A extração plasmidial da maioria dos subclones foi de qualidade suficiente para o sequenciamento, conforme pode ser verificado na Figura 14 e 15.

Figura 14. Perfil eletroforético em gel de agarose 0,8% de DNA plasmidial extraído de 48 subclones originados da construção da sub-biblioteca do clone P117.E10. 1 e 26- Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2 a 25 e 27 a 50- DNA plasmidial dos subclones.

Figura 15. Perfil eletroforético em gel de agarose 0,8% de DNA plasmidial extraído de 48 subclones originados da construção da sub-biblioteca do clone PI32.D09. 1 e 26- Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2 a 25 e 27 a 50- DNA plasmidial dos subclones.

5. 4 Clone PI17.E10

5.4.1 Análise da sequência do contig

Após o sequenciamento dos subclones do clone PI17.E10, os eletroferogramas foram analisados com o auxílio do programa PhredPhrap e através do programa ContGEN foi possível construir um *contig* de 27485 pb.

O *contig* completo foi comparado com as sequências do NCBI, usando a ferramenta BLASTN (nucleotide collection nr/nt), e o resultado foi de 83% de identidade com o micro-organismo *Parvibaculum lavamentivorans* (número de acesso CP000774.1), 52% de *Query coverage*, e 0.0 de E-value, seguido de *Rhodopseudomonas palustris* (número de acesso BX572608.1) com 75% de identidade, 2% de *Query Coverage* e 1e-85 de E-value (Tabela 2), conforme pode ser visualizado no anexo I. A interpretação do resultado indica que o *contig* obtido do clone PL17.E10 ainda não foi estudado, pois apenas pequenas regiões tiveram

similaridade. Isto sugere ainda que a sequência pertence ao gênero de *Parvibaculum*, no entanto, não é possível afirmar se a mesma espécie. A alta similaridade com o micro-organismo *Parvibaculum lavamentivorans* concorda com os dados obtidos por Paixão et al. (2010), que identificaram o gênero *Parvibaculum* como o segundo de maior frequência na mesma biblioteca de consórcio microbiano degradador de óleo diesel, devido à possibilidade destes micro-organismos terem se beneficiado pela presença de biosurfactantes produzidos pelos membros do consórcio. Segundo Schleheck et al. (2000), o surfactante LAS (alquilbenzeno linear sulfonado) disponibilizado como única fonte de carbono e energia é facilmente degradado e assimilado pelo isolado *Parvibaculum lavamentivorans* DS-1, o mesmo micro-organismo que apresentou alta similaridade com o *contig* do clone PI17.E10.

Tabela 2 - “Blast” do *contig* completo do clone PI17.E10 com as sequências do NCBI, através da ferramenta BLASTN. Similaridade com a sequência do micro-organismo *Parvibaculum lavamentivorans* (número de acesso CP000774.1), 83% de identidade e 52% de *Query coverage*.

Organismo	Descrição	Identidade (%)	Query coverage	E-value
<i>Parvibaculum lavamentivorans</i>	CP000774.1	83	52	0.0
<i>Rhodopseudomonas palustris</i>	BX57608.1	75	2	1e-85
<i>Ralstonia eutropha</i>	CP000091.1	72	4	3e-76
<i>Bradyrhizobium japonicum</i>	BA000040.2	83	6	2e-73

Após a análise do possível micro-organismo ao qual pertence o *contig* do clone PI17.E10, foi feita a predição das ORFs com o programa Open Reading Frame Finder (ORF Finder) do NCBI. O ORF Finder gerou 147 ORFs e todas foram analisadas, principalmente quanto a *Max score*, *Query coverage* e *E-value*. Assim, através desta análise foi possível selecionar 25 ORFs e identificar a frame de leitura onde cada uma está localizada; determinar o início e término da ORF; aferir o tamanho em nucleotídeos e aminoácidos; e prever qual o seu papel biológico, conforme demonstrado na Tabela 3. Todas as ORFs, encontradas foram representadas através de um mapa físico ilustrado na Figura 16.

Tabela 3 - Predição das 25 ORFs identificadas no *contig* do clone P117.E10 através ORF Finder (NCBI). Descoberta de possíveis enzimas lipolíticas, identificadas nas ORFs 8, 16, 17 e 21. (continua)

ORFs	Início e término da ORF	ORF (bp)	ORF (aa)	Descrição (número de acesso)	Organismo	Identidade (%)	E-value
1	148-1497	1350	449	AMP dependet sytnthetase and ligase (YP_001412936.1)	<i>Parvibaculum lavamentivorans</i> DS-1	78	0.0
2	1636-3501	1866	621	Acyl-CoA synthetase (YP_001412937.1)	<i>Parvibaculum lavamentivorans</i> DS-1	76	0.0
3	3523-4065	543	180	Hipothetical protein (YP_001412938.1)	<i>Parvibaculum lavamentivorans</i> DS-1	71	2e-70
4	4277-4651	375	124	Acetyltransferase (ZP_03228112.1)	<i>Bacillus coahuilensis</i>	37	2e-05
5	4882-5613	732	243	Short chain dehydrogenase reductase (YP_001412939.1)	<i>Parvibaculum lavamentivorans</i> DS-1	67	2e-93
6	5627-7801	2175	726	Methyl acepcting chemotaxis sensory transducer (ZP_02191221.1)	Alpha Proteobacterium	41	1e-124
7	8008-8292	285	94	Hypothetical protein (YP_001412940.1)	<i>Parvibaculum lavamentivorans</i> DS-1	77	5e-23
8	8537-9472	936	311	Alpha/beta hydrolase (YP_001412941.1)	<i>Parvibaculum lavamentivorans</i> DS-1	71	3e-112
9	9571-11070	1500	499	AMP dependent syhthetase and ligase (YP_001412943.1)	<i>Parvibaculum lavamentivorans</i> DS-1	76	0.0
10	11104-12810	1707	568	Sulphate transporter (YP_001412944.1)	<i>Parvibaculum lavamentivorans</i> DS-1	73	0.0
11	12819-14492	1674	557	Alkylglycerone phosphate synthetase (YP_001412945.1)	<i>Parvibaculum lavamentivorans</i> DS-1	78	0.0

Tabela 3 - Predição das 25 ORFs identificadas no *contig* do clone PI17.E10 através ORF Finder (NCBI). Descoberta de possíveis enzimas lipolíticas, identificadas nas ORFs 8, 16, 17 e 21.
(continuação)

12	14512-15126	615	204	Diacylglycerol kinase catalytic region (YP_001412946.1)	<i>Parvibaculum lavamentivorans</i> DS-1	61	3e-71
13	15468-17204	1737	578	FAD dependent oxidoreductase (YP_001412947.1)	<i>Parvibaculum lavamentivorans</i> DS-1	73	0.0
14	17273-17980	708	235	DSBA oxidoreductase (YP_001412949.1)	<i>Parvibaculum lavamentivorans</i> DS-1	79	6e-94
15	18022-19122	1101	366	Hypothetical protein (ZP_01102591.1)	<i>Congregibacter litoralis KT71</i>	49	8e-93
16	19250-20155	906	301	Alpha/beta hydrolase fold (YP_001412950.1)	<i>Parvibaculum lavamentivorans</i> DS-1	82	3e-142
17	20161-21066	906	301	Alpha/beta hydrolase fold (YP_001412951.1)	<i>Parvibaculum lavamentivorans</i> DS-1	77	5e-129
18	21157-21381	225	74	Hypothetical protein (YP_379372.1)	<i>Chlorobium chlorochromatii</i>	45	1e-11
19	21363-21899	537	178	Hypothetical protein (YP_003448028.1)	<i>Azospirillum sp</i>	32	2e-13
20	21832-22044	213	70	Hypothetical protein (ZP_0185455.1)	<i>Planctomyces maris</i>	53	0.77
21	22004-23155	1151	383	Alpha/beta hydrolase fold (YP_001412953.1)	<i>Parvibaculum lavamentivorans</i> DS-1	76	2e-137
22	23183-23989	807	268	Phosphoesterase PA phosphatase related (ZP_03266448.1)	<i>Burkholderia sp</i> H160	36	1e-27
23	24006-24773	768	768	Short chain dehydrogenase reductase SDR (YP_001412954.1)	<i>Parvibaculum lavamentivorans</i> DS-1	81	9e-119
24	24837-26483	1647	548	Glutathione regulated potassium efflux system protein (YP_002548643.1)	<i>Agrobacterium vitis S4</i>	53	4e-157

Tabela 3 - Predição das 25 ORFs identificadas no *contig* do clone P117.E10 através ORF Finder (NCBI). Descoberta de possíveis enzimas lipolíticas, identificadas nas ORFs 8, 16, 17 e 21. (conclusão)

25	26890-27447	558	185	Acyl-CoA dehydrogenase domain containing protein (YP_001412955.1)	<i>Parvibaculum lavamentivorans</i> DS-1	82	2e-87
----	-------------	-----	-----	---	---	----	-------

Figura 16. Mapa físico gerado pelo programa Pages'09, versão 4.04 (Apple Inc., Cupertino, Califórnia, Estados Unidos da América), do clone P117.E10 que expressou atividade lipolítica. Foram encontradas 25 ORFs no *contig*, sendo as ORFs em destaque os possíveis genes codificadores de enzimas lipolíticas.

A análise apurada de todas as ORFs possibilitou encontrar através do BLASTP (Non-redundant protein sequence - nr) quatro ORFs, uma congruente a outra, codificadoras de alfa/beta hidrolase. A primeira (ORF8) é constituída de 311 aminoácidos e apresentou 71% de identidade com alfa/beta hidrolase (número de acesso YP_001412941.1) de *Parvibaculum lavamentivorans*, seguido de duas enzimas lipolíticas (número de acesso ACL67843.1 e ABQ11271.1) de organismos não cultiváveis, com 58% e 54% de identidade, respectivamente. A sequência do primeiro organismo não cultivável corresponde ao clone lip9A, identificado por Hu et al. (2010) em biblioteca metagenômica do sedimento marinho do sul da China, enquanto que, a sequência do segundo organismo é um gene codificador de enzima lipolítica selecionado em biblioteca metagenômica do solo da floresta Gwangneung, na Coreia.

A segunda (ORF16) possui 301 aminoácidos e teve 82% de identidade com alfa/beta hidrolase (número de acesso YP_001412950.1) de *Parvibaculum lavamentivorans*, e depois 61% de identidade com alfa/beta hidrolase (número de acesso YP_001412953.1) do mesmo micro-organismo. A terceira (ORF17) apresenta 301 aminoácidos e teve 77% de identidade também com alfa/beta hidrolase (número de acesso YP_001412951.1) de *Parvibaculum lavamentivorans*, seguido de uma possível hidrolase (número de acesso ZP_01863361.1) de *Erythrobacter* sp. A quarta (ORF21) apresentou as duas primeiras identidades com alfa/beta

hidrolase (número de acesso YP_001412953.1 e YP_001412950.1), 76% e 63%, respectivamente, do micro-organismo *Parvibaculum lavamentivorans* (Tabela 4).

Tabela 4 - Análise, através do BLASTP (Non-redundant protein sequence - nr), das quatro ORFs (8, 16, 17 e 21) que são possíveis codificadoras de enzimas lipolíticas encontradas no *contig* P117.E10. Para cada ORF encontrada foram colocados os quatro primeiros resultados fornecidos pelo NCBI. (continua)

ORFs	Descrição	Organismo	Identidade (%)	E-value
8	Alpha/beta hydrolase domain containing protein (YP_001412941.1)	<i>Parvibaculum lavamentivorans</i> DS-1	71	3e-112
	Enzyme lipolytic (ACL67843.1)	Uncultured bacterium	58	9e-97
	Lipase/Esterase (ABQ11271.1)	Uncultured bacterium	54	1e-91
	Estersase (BAA82510.1)	<i>Oleomonas sagarensis</i>	55	1e-90
16	Alpha/beta hydrolase fold (YP_001412950.1)	<i>Parvibaculum lavamentivorans</i> DS-1	82	3e-142
	Alpha/beta hydrolase fold (YP_001412953.1)	<i>Parvibaculum lavamentivorans</i> DS-1	61	9e-100
	Alpha/beta hydrolase fold (YP_617264.1)	<i>Parvibaculum lavamentivorans</i> DS-1	58	5e-89
	Alpha/beta hydrolase fold (YP_001685157.1)	<i>Caulobacter sp</i>	56	6e-88
17	Alpha/beta hydrolase fold (YP_001412951.1)	<i>Parvibaculum lavamentivorans</i> DS-1	77	5e-129
	Probable hydrolase (ZP_01863361.1)	<i>Erythrobacter sp</i>	50	6e-75
	Alpha/beta hydrolase fold (YP_617264.1)	<i>Shingopyxis alaskensis</i>	51	8e-75
	Alpha/beta hydrolase fold (YP_001685157.1)	<i>Caulobacter sp</i>	49	7e-72
21	Alpha/beta hydrolase fold (YP_001412953.1)	<i>Parvibaculum lavamentivorans</i> DS-1	76	2e-137

Tabela 4 - Análise, através do BLASTP (Non-redundant protein sequence - nr), das quatro ORFs (8, 16, 17 e 21) que são possíveis codificadoras de enzimas lipolíticas encontradas no *contig* P117.E10. Para cada ORF encontrada foram colocados os quatro primeiros resultados fornecidos pelo NCBI. (conclusão)

	Alpha/beta hydrolase fold (YP_001412950.1)	<i>Parvibaculum lavamentivorans</i> DS-1	63	1e-109
	Alpha/beta hydrolase fold (YP_003167682.1)	<i>Candidatus accumulibacter</i>	56	4e-89
	Alpha/beta hydrolase fold (YP_617264.1)	<i>Shingopyxis alaskensis</i>	55	3e-83

As sequências das quatro ORFs também foram analisadas com as sequências depositadas de amostras ambientais, através do BLASTP, no banco de dados de amostras ambientais do NCBI. Todas as ORFs (8, 16, 17 e 21) tiveram o resultado para proteínas hipotéticas de Metagenoma Marinho que não estão classificadas, no entanto, todas possuem o resultado para o domínio conservado da superfamília de Esterase/Lipase. A ORF8 apresentou 49% de identidade com a sequência de número de acesso ECW36000.1; a ORF16 apresentou 52% de identidade com a sequência de número de acesso EDJ20548.1; a ORF17 apresentou 41% de identidade com a sequência de número de acesso ECZ95719.1; e a ORF21 apresentou 53% de identidade com a sequência de número de acesso EDJ44489.1.

5.4.2 Análise das relações filogenéticas

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

As relações filogenéticas foram construídas com o auxílio do programa Mega 4. 1 usando distintas sequências que representassem as famílias lipolíticas. Foram feitos diferentes alinhamentos e construções filogenéticas, sendo as sequências utilizadas extraídas dos trabalhos de Arpigny e Jaeger (1999), Wu e Sun (2009), Couto et al. (2010).

De acordo com o trabalho de Arpigny e Jaeger (1999) existem oito famílias lipolíticas, sendo que a primeira família apresenta seis subfamílias. Com as sequências utilizadas por este autor (50 sequências de enzimas lipolíticas) e as sequências das quatro ORFs obtidas do *contig* do clone P117.E10 para alfa/beta hidrolase foi possível analisar as relações filogenéticas entre elas, usando o algoritmo do vizinho mais próximo (neighbor-joining). Com esta construção foi possível observar que a ORF8 está localizada na família IV, enquanto que as ORFs 16, 17 e 21 estão todas juntas dentro da família V (Figura 17).

Figura 17. Dendrograma do agrupamento hierárquico baseado nas sequências de membros das famílias lipolíticas, mostrando as relações filogenéticas das ORFs de alfa/beta hidrolase encontradas no *contig* do clone P117.E10 em relação as sequências extraídas do trabalho de Arpigny e Jaeger (1999).

Já no trabalho de Wu e Sun (2009), as oito famílias também foram propostas, incluindo as enzimas lipolíticas não definidas, que provavelmente irão constituir nova(s) família(s). Foi feita uma construção da relação filogenética utilizando as sequências utilizadas por Wu e Sun (2009), 36 sequências de enzimas lipolíticas, mais as sequências das quatro ORFs, usando o mesmo parâmetro descrito acima. Foi possível observar novamente que a ORF8 está localizada dentro da família IV e as ORFs 16, 17 e 21 na família V (Figura 18).

Figura 18. Dendrograma do agrupamento hierárquico baseado nas sequências de membros das famílias lipolíticas, mostrando as relações filogenéticas das ORFs de alfa/beta hidrolase encontradas no *contig* do clone P117.E10 em relação as sequências utilizadas por Wu e Sun (2009).

Um agrupamento diferente foi observado com as sequências de Couto et al. (2010), que usaram sequências das oito famílias lipolíticas e acrescentaram sequências do “lip G”, que é uma nova família sugerida por Lee et al., (2006). Alinhando-se as 33 sequências utilizadas por Couto et al. (2010) e as sequências das quatro ORFs, usando os parâmetros já descritos, foi visto que a ORF8 está localizada próxima da família VIII e as ORFs 16, 17 e 21 estão próximas do micro-organismo *Psychobacter immobilis* (número de acesso X66712), que segundo este autor é pertencente à família na família IV, contrariando Arpigny e Jaeger (1999) que afirmam que a mesma pertence à família V (Figura 19).

Figura 19. Dendrograma do agrupamento hierárquico baseado nas sequências de membros das famílias lipolíticas, mostrando as relações filogenéticas das ORFs de alfa/beta hidrolase encontradas no *contig* do clone P117.E10 em relação as sequências utilizadas por Couto et al. (2010).

5.4.3 Identificação dos *motifs*

Através da comparação entre todas as relações filogenéticas construídas é possível supor a identificação de novos genes codificadores de alfa/beta hidrolase: a ORF8 dentro da família IV e as ORFs 16, 17 e 21 dentro da família V. Segundo Arpigny e Jaeger (1999), as enzimas lipolíticas apresentam o típico *motif* GX SXG, no qual a serina está localizada na porção central. Este *motif* foi encontrado em todas as ORFs, em algumas foi visto modificações dos aminoácidos que o constituem, o que é característico da família a qual pertence.

Os membros da família IV pertencem ao grupo de lipases que atuam em baixas temperaturas e exibem similaridade com as Lipases Hormônio Sensitivas de mamíferos (HLS). Elas apresentam o sítio ativo com a serina como resíduo dentro do pentapeptídeo GDSAG, além de apresentarem um *motif* estritamente conservado HGGG, localizado *upstream* do sítio ativo (JAEGER; DIJKSTRA; REETZ, 1999), cuja função está relacionada na formação das pontes de hidrogênio para estabilização durante a catálise (WEI et al., 1999).

Com os alinhamentos, no ClustalW, realizados com a sequência da ORF8 e de alguns micro-organismos da família IV foi possível identificar as sequências de aminoácidos <Ser159-Asp191-His283> para esta ORF, que é da tríade catalítica conservada, presente nas enzimas lipolíticas da superfamília alfa/beta hidrolase.

A sequência de aminoácidos HGGG; a do primeiro sítio catalítico (GDSAG); a do segundo sítio (DPLXD); e do terceiro sítio (HGF) foram identificados na ORF8 (Figura 20). Estas sequências são encontradas nos representantes da família IV e têm sido identificados em clones advindos da abordagem metagenômica, como por exemplo, os clones identificados por Hu et al. (2010), que foram caracterizados como membros desta família. Assim, levando em consideração todas as análises com a ORF8, sugere-se que esta seja uma ORF representante da família IV.

Figura 20. Alinhamento das sequências da ORF8 com as sequências de aminoácidos de *Moxarella* sp. (número de acesso X53868), *Pseudomonas* sp. (número de acesso AF034088) e *Cupriavidus necator* (número de acesso L36817), que são representantes da família IV. As caixas representam as regiões conservadas das enzimas lipolíticas; • Possíveis resíduos de aminoácidos envolvidos na tríade catalítica.

De acordo com Jaeger, Dijkstra e Reetz (1999) ambas as famílias IV e V apresentam lipases pertencentes aos micro-organismos psicrófilos (*Moxarella* sp., *Pseudomonas* sp. e *Psychrobacter immobilis*) e termófilos (*Archaeoglobus fulgidus* e *Sulfolobus acidocalcaldirus*). Os membros da família V possuem estruturas similares à dehalogenases, haloperoxidases e epóxido hidrolases, exibindo folha alfa/beta hidrolase como característica de estrutura terciária. Segundo Arpigny e Jaeger (1999), os blocos conservados GXSMGG, PTLV e GH são característicos nos membros da família V, sendo que, é possível encontrar os sítios catalíticos nos blocos conservados.

Com os alinhamentos no ClustalW, realizados com a sequência das ORFs 16, 17 e 21, e de alguns micro-organismos da família V foi possível identificar as sequências de aminoácidos do sítio catalítico (Figura 21), nas quais para a ORF16 foi identificada a sequência <Ser118-Asp245-His273>; para a ORF17 a sequência <Ser118-Asp245-His273>; e para a ORF21 a sequência <Ser118-Asp245-His274>. Os alinhamentos permitiram confirmar os resultados obtidos na construção das árvores filogenéticas, ou seja, que todas estas ORFs

são pertencentes à família V, sendo possível identificar os blocos conservados (ARPIGNY e JAEGER, 1999) e as tríades catalíticas (ARPIGNY e JAEGER, 1999; HU et al., 2010).

Figura 21. Alinhamento das sequências das ORFs 16, 17 e 21 com as sequências de aminoácidos de *Sulfolobus acidocaldarius* (número de acesso AF071233), *Moxarella* sp. (número de acesso X53869) e *Psychrobacter immobilis* (número de acesso X66712), que são representantes da família V. As caixas representam as regiões conservadas das enzimas lipolíticas; • Possíveis resíduos de aminoácidos envolvidos na tríade catalítica.

5.5 Clone PI32.D09

5.5.1 Análise da sequência do contig

Já os subclones do clone PI32.D09 também foram sequenciados e as análises utilizadas foram as mesmas descritas para o clone PI17.E10, sendo possível construir um *contig* de 35631 pb através do programa ContGEN.

O *contig* completo foi comparado com as sequências do NCBI, usando a ferramenta BLASTN e a coleção de nucleotídeos, e o resultado foi de 100% de identidade, 43% de *Query coverage* e 0.0 de E-value com a sequência depositada de *Parvibaculum lavamentivorans* (número de acesso CP000774.1) (Tabela 5), conforme pode ser visualizado no anexo II. O resultado indica que o *contig* obtido do clone PL32.D10 também não foi estudado, já que

apenas pequenas regiões tiveram similaridade, sugerindo que a sequência pertence ao gênero de *Parvibaculum*, não sendo possível afirmar se é uma nova espécie de *Parvibaculum*.

Tabela 5 - “Blast” do *contig* completo do clone P132.D09 com as sequências do NCBI, através da ferramenta BLASTN. Similaridade com a sequência do micro-organismo *Parvibaculum lavamentivorans* (número de acesso CP000774.1), 100% de Identidade e 43% de *Query coverage*.

Organismo	Descrição	Identidade (%)	Query coverage	E-value
<i>Parvibaculum lavamentivorans</i>	CP00774.1	100	43	0.0
<i>Caulobacter crescentus</i>	CP001340.1	81	11	0.0
<i>Caulobacter crescentus</i>	AE005673.1	81	11	0.0
<i>Caulobacter segnis</i>	CP002008.1	86	11	0.0

Após a análise do possível micro-organismo que pertence o *contig* do clone P132.D09, foi feita a predição das ORFs com o programa Open Reading Frame Finder (ORF Finder) do NCBI, o qual gerou 186 ORFs e 32 foram selecionadas em função do *Max score*, *Query coverage* e *E-value*. As ORFs foram analisadas detalhadamente do mesmo modo que as ORFs encontradas no clone P117.E10 e o resultado está presente na Tabela 6. Todas as ORFs, encontradas foram representadas através de um mapa físico ilustrado na Figura 22.

Tabela 6. Predição das 32 ORFs identificadas no *contig* do clone P132.D09 através ORF Finder (NCBI). Descoberta de uma possível enzima lipolítica, ORF 30, e de Tioesterases, ORFs 25, 31 e 32. (continua)

ORFs	Início e término da ORF	ORF (bp)	ORF (aa)	Descrição (número de acesso)	Organismo	Identidade (%)	E-value
1	119-343	225	74	Prephenate dehydratase (YP_001411532.1)	<i>Parvibaculum lavamentivorans</i> DS-1	82	7e-29
2	990-1940	951	316	Nudix hydrolase (YP_001411534.1)	<i>Parvibaculum lavamentivorans</i> DS-1	76	1e-143

Tabela 6. Predição das 32 ORFs identificadas no *contig* do clone PI32.D09 através ORF Finder (NCBI). Descoberta de uma possível enzima lipolítica, ORF 30, e de Tioesterases, ORFs 25, 31 e 32. (continuação)

3	2185-4032	1848	615	DNA polymerase III, subunits gamma and tau (YP_001411535.1)	<i>Parvibaculum lavamentivorans</i> DS-1	74	0.0
4	4432-4818	387	128	Recombination protein RecR (YP_001411538.1)	<i>Parvibaculum lavamentivorans</i> DS-1	82	7e-33
5	5109-5762	654	217	Cytochrome B561 (YP_426331.1)	<i>Rhodospirillum rubrum</i>	44	5e-43
6	5883-6416	533	177	NADPH dependent FMN reductase (YP_825045.1)	<i>Solibacter usilatus</i>	53	9e-38
7	6423-7721	687	228	Hypothetical protein (YP_001411540.1)	<i>Parvibaculum lavamentivorans</i> DS-1	75	8e-145
8	7944-8468	525	174	Peptide diformylase (YP_001411541.1)	<i>Parvibaculum lavamentivorans</i> DS-1	80	1e-71
9	8497-9435	939	312	Methionyl-tRNA formyltransferase (YP_001411542.1)	<i>Parvibaculum lavamentivorans</i> DS-1	75	9e-130
10	9435-10214	780	259	tRNA Pseudouridine synthetase A (YP_001411543.1)	<i>Parvibaculum lavamentivorans</i> DS-1	79	6e-115
11	10803-11993	1191	396	Succinyl-diaminopimelate desuccinylase (YP_001411546.1)	<i>Parvibaculum lavamentivorans</i> DS-1	72	2e-169
12	12040-12885	846	281	2,3,4,5 tetrahydropyridine 2,6 dicarboxylate (YP_001411547.1)	<i>Parvibaculum lavamentivorans</i> DS-1	88	2e-142
13	12979-14238	1260	419	Diguanylate phosphodiesterase (YP_001411548.1)	<i>Parvibaculum lavamentivorans</i> DS-1	70	7e-165

Tabela 6. Predição das 32 ORFs identificadas no *contig* do clone PI32.D09 através ORF Finder (NCBI). Descoberta de uma possível enzima lipolítica, ORF 30, e de Tioesterases, ORFs 25, 31 e 32. (continuação)

14	14584-15063	480	159	Glutathione-dependent formaldehyde activating (YP_001411550.1)	<i>Parvibaculum lavamentivorans</i> DS-1	70	3e-64
15	15112-15843	732	243	OmpA/MotB domain containing protein (YP_001411552.1)	<i>Parvibaculum lavamentivorans</i> DS-1	74	1e-72
16	15981-16745	764	254	Short-chain dehydrogenase/reductase (YP_001411553.1)	<i>Parvibaculum lavamentivorans</i> DS-1	86	4e-128
17	16996-17679	684	227	GCNS-related N-acetyltransferase (YP_001411554.1)	<i>Parvibaculum lavamentivorans</i> DS-1	69	6e-45
18	17330-18823	1494	497	acyl- CoA dehydrogenase domain- containing protein (YP_001411779.1)	<i>Parvibaculum lavamentivorans</i> DS-1	89	6e-149
19	18795-20339	1544	514	Acyltransferase (YP_001413218.1)	<i>Parvibaculum lavamentivorans</i> DS-1	37	6e-79
20	21417-22367	951	316	Hypothetical protein (YP_001411777.1)	<i>Parvibaculum lavamentivorans</i> DS-1	65	2e-108
21	22427-23629	1203	400	Acyl-coA dehydrogenase domain containing protein (YP_001411776.1)	<i>Parvibaculum lavamentivorans</i> DS-1	79	0.0
22	23694-24410	717	238	Hypothetical protein (ZP_01089606.1)	<i>Blastopirellula marina</i>	49	1e-61
23	24794-25411	618	205	Dehydratase (YP_001411775.1)	<i>Parvibaculum lavamentivorans</i> DS-1	87	2e-70

Tabela 6. Predição das 32 ORFs identificadas no *contig* do clone PI32.D09 através ORF Finder (NCBI). Descoberta de uma possível enzima lipolítica, ORF 30, e de Tioesterases, ORFs 25, 31 e 32. (conclusão)

24	25425-26918	1493	497	Acetyl-CoA acetyltransferase (YP_001411774.1)	<i>Parvibaculum lavamentivorans</i> DS-1	83	0.0
25	26917-27345	429	142	Thioesterase superfamily protein (YP_001411772.1)	<i>Parvibaculum lavamentivorans</i> DS-1	66	6e-42
26	27677-28237	561	186	Glutathione- dependent formaldehyde activating (YP_001867409.1)	<i>Nostoc pinetiforme</i>	40	6e-28
27	28357-29646	1290	429	3alpha, 7alpha, 12alpha- trihydroxy 5- beta cholest 24 enoyl Coa hydratase (YP_001411770.1)	<i>Parvibaculum lavamentivorans</i> DS-1	75	1e-127
28	30190-31236	1047	348	AraC family transcriptional regulator (YP_001411748.1)	<i>Parvibaculum lavamentivorans</i> DS-1	67	3e-120
29	31233-31865	633	210	ThiJ/PfpI domain- containing protein (YP_001411750.1)	<i>Parvibaculum lavamentivorans</i> DS-1	75	3e-87
30	32697-33665	968	322	Alpha/beta hydrolase domain- containing protein (YP_001411752.1)	<i>Parvibaculum lavamentivorans</i> DS-1	76	1e-146
31	33670-34611	942	313	Thioesterase-like protein (YP_001411756.1)	<i>Parvibaculum lavamentivorans</i> DS-1	35	4e-51
32	34608-35501	894	297	Thioesterase-like protein (YP_001411756.1)	<i>Parvibaculum lavamentivorans</i> DS-1	72	3e-125

Figura 22. Mapa físico gerado pelo programa Pages'09, versão 4.04 (Apple Inc., Cupertino, Califórnia, Estados Unidos da América), do clone PI32.D09 que apresentou atividade lipolítica. Foram encontradas 32 ORFs no *contig*, sendo a ORF30 (rosa escuro) o possível gene codificador da enzima lipolítica; e as ORFs 25, 31 e 32 (rosa claro) genes relacionados com a codificação de Tioesterases.

A análise detalhada de todas as ORFs possibilitou encontrar através do BLASTP (Non-redundant protein sequence - nr) uma ORF codificadora de alfa/beta hidrolase. A ORF30 é constituída de 322 aminoácidos e apresentou 76% de identidade com alfa/beta hidrolase (número de acesso YP_001411752.1) de *Parvibaculum lavamentivorans*, seguido de 55% de identidade com alfa/beta hidrolase (número de acesso YP_003692034.1) de *Starkeya novella* (Tabela 7).

Tabela 7 - Análise, através do BLASTP (Non-redundant protein sequence - nr), da ORF30 que é uma possível codificadora de enzima lipolítica, encontrada no *contig* PI32.D09. Para esta ORF foram colocados os quatro primeiros resultados fornecidos pelo NCBI.

ORF	Descrição	Organismo	Identidade (%)	E-value
30	Alpha/beta hydrolase domain containing protein (YP_001411752.1)	<i>Parvibaculum lavamentivorans</i> DS-1	76	3e-147
	Alpha/beta hydrolase fold (YP_003692034.1)	<i>Starkeya novella</i>	55	7e-92
	Alpha/beta hydrolase fold (YP_001926035.1)	<i>Methylobacterium populi</i>	52	5e-86
	Alpha/beta hydrolase domain containing protein (YP_001640612.1)	<i>Methylobacterium extorquens</i>	52	7e-85

A sequência da ORF30 também foi analisada com as sequências depositadas de amostras ambientais, através do BLASTP, no banco de dados de amostras ambientais do NCBI. O resultado foi de 43% de identidade com proteína hipotética (número de acesso

EDI67145.1) de Metagenoma Marinho, no entanto, apesar de estar classificada como proteína hipotética, foi verificado o domínio conservado da Superfamília de Esterase/Lipase. A razão da maioria das ORFs serem mais próximas de bactérias cultivadas é devido ao fato de que poucas enzimas lipolíticas derivadas do metagenoma foram depositadas no banco de dados de proteína, isto diminui a probabilidade de obter resultado de bactéria não cultivável com o BLASTP (LIAW et al., 2010).

5.5.2 Análise das relações filogenéticas

A análise da relação filogenética do clone PL32.D09 foi feita da mesma forma que o clone PL17.E10. Assim, após a predição das ORFs e localização no mapa gênico, as análises dos alinhamentos foram realizadas através do ClustalW, usando o programa BioEdit Sequence Alignment Editor, e as relações filogenéticas foram construídas com o auxílio do programa Mega 4. 1. Foram utilizadas as mesmas sequências extraídas dos trabalhos de Arpigny e Jaeger (1999), de forma que, em todos os alinhamentos e construções filogenéticas as famílias lipolíticas foram representadas de forma fidedigna.

Com as sequências utilizadas por Arpigny e Jaeger (1999) e a sequência da ORF30 foi possível identificar as relações filogenéticas entre elas, usando o algoritmo do vizinho mais próximo (neighbor-joining). Através da análise desta construção foi observado que a ORF30 está próxima aos micro-organismos *Moxarella* sp. (X53868), *Cupriavidus necator* (L36817) e *Pseudomonas* sp. (AF034088), que são representantes da família IV (Figura 23).

Figura 23. Dendrograma do agrupamento hierárquico baseado nas sequências de membros das famílias lipolíticas, mostrando a relação filogenética da ORF30 de alfa/beta hidrolase, encontrada no *contig* do clone PI32.D09, em relação as sequências utilizadas por Arpigny e Jaeger (1999).

5.5.3 Identificação dos *motifs*

Através da análise da árvore filogenética construída é possível supor a identificação de um novo gene codificador de alfa/beta hidrolase dentro da família IV. Conforme já mencionado, as enzimas lipolíticas apresentam o típico *motif* GX SXG, no qual a serina está localizada na porção central (ARPIGNY e JAEGER, 1999), sendo que, para os membros da família IV este *motif* é representado pelo pentapeptídeo GDSAG. Outra característica desta família é o *motif* estritamente conservado HGGG, localizado *upstream* do sítio ativo (JAEGER; DIJKSTRA; REETZ, 1999).

Com os alinhamentos no ClustalW realizados com a sequência da ORF30 e de alguns micro-organismos da família IV foi possível identificar as sequências de aminoácidos <Ser170-Asp265-His295> para esta ORF, que corresponde à tríade catalítica conservada.

A sequência de aminoácidos HGGG e os três sítios catalíticos (GDSAG; DPLXD e HGG) foram identificados na ORF30 (Figura 24), no entanto, foram verificadas algumas modificações nos resíduos de aminoácidos. O resíduo de aminoácido fenilalanina (F) do *motif* conservado HGGGF foi substituído por triptofano (W), no sítio *upstream* ao sítio catalítico da serina, sendo a sequência deste bloco, HGGGW. Quanto ao *motif* GDSAG, característico desta família, o resíduo de aspartato (D) está substituído por glutamato (E), sendo esta substituição também verificada por Hu et al. (2010) em um clone (FLS9B) com atividade lipolítica, advindo de uma biblioteca metagenômica de sedimento marinho da China. Outra modificação foi vista no segundo sítio catalítico, do aspartato, no qual o resíduo de aminoácido prolina (P) está trocado por alanina (A), formando o bloco DALVD.

Figura 24. Alinhamento da sequência da ORF30 com as sequências de aminoácidos de *Moxarella* sp. (número de acesso X53868), *Pseudomonas* sp. (número de acesso AF034088) e *Cupriavidus necator* (número de acesso L36817), que são representantes da família IV. As caixas representam as regiões conservadas das enzimas lipolíticas; • Possíveis resíduos de aminoácidos envolvidos na tríade catalítica.

A fim de confirmar os resultados obtidos na árvore filogenética e no alinhamento com representantes da família IV, uma nova construção filogenética foi feita utilizando sequências extraídas do NCBI. A sequência da ORF30 foi comparada com o NCBI, BLASTP (Non-redundant protein sequence - nr), e os 32 primeiros resultados tiveram suas sequências extraídas para análise em árvore filogenética e alinhamento de aminoácidos.

Assim, as sequências extraídas foram alinhadas junto com a ORF30 no ClustalW, através do BioEdit Sequence Alignment Editor, e posteriormente a árvore filogenética foi construída no Mega 4.1, usando o algoritmo do vizinho mais próximo (neighbor-joining). A árvore construída possibilitou verificar a aproximação da ORF30 com o micro-organismo *Parvibaculum lavamentivorans* (YP00141175) (Figura 25), fato esperado devido ao resultado com o banco NCBI. No entanto, nenhum outro micro-organismo está localizado no ramo

destas duas sequências, apenas os micro-organismos *Caulobacter segnis* (YP03592835) e *Streptomyces scabiei* (YP003487189) estão no ramo mais próximo.

A análise dos alinhamentos, juntamente com os outros resultados, permitiu afirmar que a ORF30 do clone PI32.D09 é provavelmente um novo membro da família IV, ou seja, lipases que atuam em baixas temperaturas. Nesta ORF foi verificado o *motif* conservado HGGG (localizado *upstream* do sítio ativo) e o pentapeptídeo GDSAG, no entanto, em quase todos os micro-organismos usados no alinhamento apresentam a substituição do aspartato (D) por glutamato (E). No segundo sítio (DPLXD) também foi confirmada a substituição da prolina (P) por alanina (A) em 6 das 32 sequências utilizadas. O terceiro sítio (HGF) encontrado na ORF30 foi verificado também nas demais sequências, completando assim a tríade catalítica (Figura 26).

Figura 25. Dendrograma do agrupamento hierárquico baseado nas sequências de 32 micro-organismos extraídas do NCBI com a ORF30, encontrada no *contig* do clone PI32.D09.

250 260 270 280 290 300 310 320

Parvibaculum lavamentivorans_Y PQA**RE**IVIDGKAG--NKITLRVIA**PET**---SPKGVYLIHGGG**W**VLGGADQ**Q**-----D**PML**ERIAKNTGLACV
 Starkeya novella_YP_003692034 SYAQV**IE**IDGPHG--FVPLRIITP**PT**---PPKGVYLIHGGG**W**CLGSARE**N**-----D**GLN**H**RF**VE**RT**GF**AV**V
 Methylobacterium populi_YP_001 PRA**ET**L**TI**EGPGG--PLALRVIRP**NG**---KPGAYLIHGGG**W**WGAADE**Q**-----D**PWL**ER**IA**DACGFV**CL**
 Methylobacterium extorquens_YP PRA**ET**V**TI**EGPAG--PLALRVIRP**Q**G---KARGAYLIHGGG**W**WGAADE**Q**-----D**PWL**ER**IA**D**TC**GFV**CL**
 Methylobacterium chloromethani PRA**ET**V**TI**AGPAG--PLALRVIRP**Q**G---KARGAYLIHGGG**W**WGAADE**Q**-----D**PWL**ER**IA**D**TC**GFV**CL**
 Xanthobacter autotrophicus_YP ARARQ**ET**IGP**RG**--DIPVRIYMPD**G**---APTGVYLIHGGG**W**ITWGADE**Q**-----E**GLM**DR**IA**GC**CG**LAVV
 Methylobacterium radiotolerans ARAE**W**L**TI**DGPGG--PLKLRVVR**PK**GVKLRGAYLIHGGG**W**WGSADE**Q**-----D**PWL**ER**IA**D**TC**GFV**AL**
 Caulobacter segnis_YP_00359283 PRAYDVPVN----D**TV**SL**LV**VP**CA**---E**PK**GVLLHGGG**W**ITLGA**AA**H**Q**-----D**AM**L**ER**IA**DQ**TALT**CV**
 Methylobacterium sp_YP_0017701 PRA**ET**L**TI**PGPGG--P**IAL**RVIRP**VR**---KPGAYLIHGGG**W**ITWGADE**Q**-----D**PWL**ER**IA**D**RC**GFV**CI**
 Fulvimarina pelagi_ZP_01438446 PDAT**MS**IT**GR**DGR--EISVRI**IK**PG**R**--EARGT**FL**LIHGGG**W**VFGD**AT**EN-----D**TR**L**RR**L**AE**ATGLAT**V**
 Azorhizobium caulinodans_YP_00 PRAR**NE**T**IE**GP**RG**--PVPVRIYLPD**G**---APTGVYLIHGGG**W**ITGSARE**N**-----D**GLN**D**AV**A--AAGLV**V**
 Ahrensia_sp_ZP_07375984 TRARTLLAPRANGD**GI**VLRL**IE**P**IG**R--PMRG**LY**LIHGGG**W**MLGAAD**AQ**-----D**ER**L**Q**EA**AT**GL**GC**V
 Aurantimonas manganoxydans_ZP KNA**EE**IIHGGAGR--EIPVRV**RP**D**G**--Q**TR**GT**FL**LIHGGG**W**VFGA**DE**A-----D**PR**L**RR**MAD**NT**GLT**V**A
 Methylobacterium nodulans_YP_0 PRA**ET**L**AI**PGPGG--PVALRVIRP**VR**---K**TR**GAYLIHGGG**W**ITWGADE**Q**-----D**PWL**ER**IA**D**RC**GFV**CV**
 Pseudovibrio_sp_ZP_05085344 ENA**EE**LE**IE**GP**H**--PISLRL**FR**AD**EN**--SSKGVV**HF**HGGG**W**VFGS**AA**F**Q**-----D**PY**LE**EM**AK**NT**GLD**V**
 Streptomyces scabiei_YP_003487 PRAFD**Q**IA----D**RL**SL**LV**VP**SA**---RPRAVLLHGGG**W**ITLGA**AA**H**Q**-----D**ALL**K**HY**T**DA**ANLV**SI**
 Streptomyces viridochromogenes SHG**Q**DR**VI**AGG----V**K**VRA**FP**D---HVD**GI**YLIHGGG**W**AFGS**AD**G**Q**-----D**E**KL**W**Q**LAV**KAR**LA**V**A**
 Coccidioides immitis_XP_001242 DCAQ**S**FS**IP**S**RE**PG**R**NI**PC**RI**IK**PGSSAPSGV**FL**LIHGGG**W**VLMD**ES**S**Q**-----D**ALL**K**HY**T**DA**ANLV**SI**
 Coccidioides posadasii_XP_0030 DCAQ**S**FS**IP**S**RE**PG**R**NI**PC**RI**IK**PGSSAPSGV**FL**LIHGGG**W**VLMD**ES**S**Q**-----D**ALL**K**HY**T**DA**ANLV**SI**
 Streptomyces viridochromo_ZP_0 -----MKVRA**FP**D---HVD**GI**YLIHGGG**W**AFGS**AD**G**Q**-----D**E**KL**W**Q**LAV**KAR**LA**V**A**
 Frankia alni_YP_710823 AEA**H**DRQ**VE**D**GV**---RVGLRV**FP**---VVEGAYLIHGGG**W**AFGS**AD**G**Q**-----D**E**Q**L**W**D**LARTAR**LA**V**L**
 Uncinocarpus reesii_XP_0025448 DCG**Q**S**FM**IP**S**REAG**R**NI**PC**RV**IK**PSGD**TP**SGV**FL**LIHGGG**W**VLSD**LS**Q**PH**MRAT**DN**AS**LS**Q**D**-----D**ALL**K**HY**T**DA**ANLV**SI**
 Nectria haematococca_XP_003039 ASATA**FQ**IP**S**RD**PR**IP**CR**LL**RP**Q**NN**K**PL**CA**VF**M**HI**HGGG**W**VLQ**DE**T**Y**Q-----D**PR**L**Q**AI**AD**D**H**GL**LC**I
 Phaeosphaeria nodorum_XP_00179 PSQV**NG**T**LP**SRD**PS**RI**PM**K**PS**SGA--SK**GI**H**MI**HGGG**W**VLQ**SE**H**Y**Q-----D**PY**L**K**Y**MAD**NYQ**IT**V**I**
 Podospira anserina_XP_00191150 PAAK**D**V**V**SSRD**AG**D**IAL**RV**TP**D**NG**Q**PS**K**GL**FL**LI**HGGG**W**FM**GT**H**Q**H**Q**-----D**G**KL**RE**Y**ANT**FQ**L**T**AL**
 Pyrenophora tritici_repentis_X PEG**T**NGAL**PS**RD**AG**REIP**RV**FK**PAG**GK--SRG**I**H**MI**HGGG**W**VL**MS**E**AY**Q-----D**PY**L**K**Y**MAD**H**FE**V**T**V**V**
 Micromonas_sp_XP_002499522 AKW**SP**S**EE**K**WP**D**G**-----P**RG**IYLLHGGG**W**VFGD**AC**G**Q**-----D**ND**V**LE**EM**SE**EL**GL**V**LV**L
 Arthroderma otae_XP_002844915 PDAIDAT**LP**SRD**AG**RD**IP**LV**FK**EN**GL**PS**KQ**IF**LF**HGGG**W**VL**TH**E**GG**-----D**TP**L**K**Y**M**AN**KC**Q**L**T**AI**
 Rhodococcus opacus_YP_00278012 PPA**VR**EV**VA**AD**GR**--R**IP**L**RI**H**AP**VS**R**--AATG**VY**LEHGGG**W**VL**GS**A**AG**S-----D**V**R**NR**RL**AD**VL**GV**AV**V**
 Penicillium chrysogenum_XP_002 PEA**LD**GT**VP**SRD**AG**REIP**RI**Y**K**PN**GQ**PS**KI**FL**LI**HGGG**W**VL**TH**E**K**S-----D**ST**L**RE**Y**AN**K**C**Q**L**T**AI**
 Mycobacterium smegmatis_YP_889 PPP**S**VE**TV**S**AG**OR--SV**FP**RI**H**APIS**R**--PAA**GV**LE**IE**HSGG**W**VL**GS**A**AD**S-----D**V**R**NR**RL**LV**NAL**GV**AV**V**
 ORF30 ERART**IE**EGK**GG**--HKV**PL**RI**IA**PE---SPKGVYLIHGGG**W**VLG**AC**D**Q**Q-----D**PML**ER**IA**Q**NA**GLACV

310 320 330 340 350 360 370 380 390 400

Parvibaculum lavamentivorans_Y -----D**PML**ER**IA**KNTGLACV**SV**YRLA**PE**HP**YP**AG**PD**CEAAAL**WL**NA**KTE**FG**TD**---L**LT**IGESAG**HL**SA**VT**LL**IR**MR**K**
 Starkeya novella_YP_003692034 -----D**GLN**H**RF**VE**RT**GF**AV**SV**YRLA**PE**HP**PG**PD**CEAAAL**WL**MB**GP**AR**FG**TD---R**FA**IGESAG**AN**LS**AV**LL**IR**LR**K**
 Methylobacterium populi_YP_001 -----D**PWL**ER**IA**DACGFV**CL**SV**YRLA**PE**HP**Y**PA**AL**ED**CEAAAL**WL**AG**PG**RA**EL**GV**H**---A**LT**IGESV**GA**HL**AV**MT**LL**LR**DR**
 Methylobacterium extorquens_YP -----D**PWL**ER**IA**D**TC**GFV**CL**SV**YRLA**PE**HP**Y**PA**AL**ED**CEAAAL**WL**AG**PG**KA**EL**GI**H**---A**LT**IGESV**GA**HL**AV**MT**LL**LR**DR**
 Methylobacterium chloromethani -----D**PWL**ER**IA**D**TC**GFV**CL**SV**YRLA**PE**HP**Y**PA**AL**ED**CEAAAL**WL**AG**PG**KA**EL**GI**H**---A**LT**IGESV**GA**HL**AV**MT**LL**LR**DR**
 Xanthobacter autotrophicus_YP -----E**GLM**DR**IA**GC**CG**LAVSV**YAL**A**PE**HP**YP**AG**PD**CEAAAL**WL**V**KE**GA**AR**FG**TE**---R**FA**IGES**GA**HL**AV**MT**LL**VR**LR**K
 Methylobacterium radiotolerans -----D**PWL**ER**IA**D**TC**GFV**AL**SV**YRLA**PE**HP**Y**PA**AL**ED**CEAAAL**WL**AG**PG**K**TE**FG**IH**---A**LT**IGESV**GA**HL**AV**MT**LL**LR**DR**
 Caulobacter segnis_YP_00359283 -----D**AM**L**ER**IA**DQ**TALT**CV**SV**YRLA**PE**HP**Y**PA**AM**DD**CEAAAL**WL**VE**HA**KS**SR**FG**VE**---L**LM**IGESAG**AL**L**AV**AT**LL**LR**RP**
 Methylobacterium sp_YP_0017701 -----D**PWL**ER**IA**D**RC**GFV**CI**SV**YRLA**PE**HP**Y**PA**LD**CE**AAAL**WL**H**GB**RR**LF**GV**E**---A**LT**IGESV**GA**HL**SV**MT**LL**LR**DR**
 Fulvimarina pelagi_ZP_01438446 -----D**TR**L**RR**L**AE**ATGLAT**SV**YRLA**PE**HP**FP**AA**GD**CEAAAL**AL**AQ**Q**IA**D**LP**TD**---W**LA**IGESAG**AL**SA**VT**LL**IR**DD
 Azorhizobium caulinodans_YP_00 -----D**GLN**D**AV**A--AAGLV**V**SV**YAL**A**PE**TP**YP**AA**PD**CEAAAL**WL**RL**EG**LS**FG**TD---R**FL**IGESAG**AL**L**AV**MT**LL**LR**DR**
 Ahrensia_sp_ZP_07375984 -----D**ER**L**Q**EA**AT**GL**GC**SV**YRLA**PE**YP**Y**PA**GD**DC**DAAR**FL**--E**AE**H**GL**FN**E**---I**MT**IGESAG**AL**SV**LT**LL**LR**D
 Aurantimonas manganoxydans_ZP -----D**PR**L**RR**MAD**NT**GLT**AS**SV**YRLA**PE**FP**PA**AL**ED**CT**DA**VA**L**ANG**HL--D**LP**RG---F**LA**IGESAG**AL**L**AV**MT**LL**LR**DR**
 Methylobacterium nodulans_YP_0 -----D**PWL**ER**IA**D**RC**GFV**CS**SV**YRLA**PE**HP**Y**PA**ID**CE**AAAL**WL**H**TE**GR**KL**FG**VE**---A**LT**IGESV**GA**HL**SV**MT**LL**LR**DR**
 Pseudovibrio_sp_ZP_05085344 -----D**PY**LE**EM**AK**NT**GLD**V**SV**YRLA**PE**CP**FP**Q**PD**CE**AA**CH**W**LN**EG**AS**RF**GW**S---S**FL**IGESAG**AL**L**AL**IT**AL**LR**DR**
 Streptomyces scabiei_YP_003487 -----D**ALL**ER**MA**D**AV**GAT**CV**SV**YRLA**PE**HP**Y**PA**AM**DD**CEAA**Q**W**LL**K**H**AG**ER**FG**AE**---T**LL**IGESAG**AL**L**AV**AT**LQ**LR**NR**
 Streptomyces viridochromogenes -----D**E**KL**W**Q**LAV**KAR**LAV**SV**YRLA**PE**HP**FP**AG**PD**CE**AAAR**WL**V**EN**AA**EF**GT**G**---R**LL**IGESAG**AL**SV**VT**LL**LR**DR**H**
 Coccidioides immitis_XP_001242 -----D**ALL**K**HY**T**DA**ANLV**SI**SV**YRLA**PE**HP**Y**PA**GD**ED**CD**AE**W**LV**NA**EE**KG**AP**---L**K**FM**GE**SAG**HL**SALT**AL**H**LL**SH**EN**P
 Coccidioides posadasii_XP_0030 -----D**ALL**K**HY**T**DA**ANLV**SI**SV**YRLA**PE**HP**Y**PA**GD**ED**CD**AE**W**LV**NA**EE**KG**AP**---L**K**FM**GE**SAG**HL**SALT**AL**H**LL**SH**EN**P
 Streptomyces viridochromo_ZP_0 -----D**E**KL**W**Q**LAV**KAR**LAV**SV**YRLA**PE**HP**FP**AG**PD**CE**AAAR**WL**V**EN**AA**EF**GT**G**---R**LL**IGESAG**AL**SV**VT**LL**LR**DR**H**
 Frankia alni_YP_710823 -----D**E**Q**L**W**D**LARTAR**LA**VL**SV**YRLA**PE**HP**Y**PA**LD**CEAAAL**WL**V**RNA**AV**EF**GT**D**---R**LL**IGESAG**AL**SL**TS**LL**LR**DR**H**
 Uncinocarpus reesii_XP_0025448 SLS**Q**D**PL**L**K**Y**AD**ANLV**SI**IA**Y**RLA**PE**HP**FP**AA**PD**CD**AE**W**LV**NA**EE**KA**FG**AP---L**Q**FI**GE**SAG**HL**SA**VL**HL**LS**H**Q**D**P**
 Nectria haematococca_XP_003039 -----D**PR**L**Q**AI**AD**D**H**GL**LC**SV**YRLA**PE**HP**FP**AG**PD**CD**DA**EW**LV**NA**EE**IG**AP---L**Q**FI**GE**SAG**AL**SL**MT**TI**HL**L**Q**H**D**KV
 Phaeosphaeria nodorum_XP_00179 -----D**PY**L**K**Y**MAD**NYQ**IT**V**IS**IG**Y**RLA**PE**D**WP**AC**AE**DC**DA**EW**LV**NA**EE**KG**AP**---L**MF**TI**GE**SAG**HL**LA**CV**AL**HL**LA
 Podospira anserina_XP_00191150 -----D**G**KL**RE**Y**ANT**FQ**L**T**AL**SV**YRLA**PE**D**WP**AC**AE**DC**DA**EW**LV**NA**EE**KG**AP---L**LF**TI**GE**SAG**HL**LA**CV**AL**HL**LA
 Pyrenophora tritici_repentis_X PEP**Y**L**K**Y**MAD**H**FE**V**T**SV**IS**IG**Y**RLA**PE**D**WP**AC**AE**DC**DA**EW**LV**NA**EE**KG**AP**---L**MF**TI**GE**SAG**HL**LA**CV**AL**HL**LA
 Micromonas_sp_XP_002499522 -----D**ND**V**LE**EM**SE**EL**GL**V**LV**PD**YS**KA**PE**HP**Y**PA**V**NE**CE**LA**AI**W**VE**NA**TE**RE**FG**E**KA**RM**MI**IGESAG**HL**CA**AV**LL**RR**DR**AM**RP**PE**W**Y**K**Q**W**TP**
 Arthroderma otae_XP_002844915 -----D**TP**L**K**Y**M**AN**KC**Q**L**T**AI**SV**YRLA**PE**DP**Y**AG**IND**CD**AE**Y**LV**D**HA**ER**GY**AK**---L**TI**MG**ES**SG**CL**A**AL**T**AF**S**IM**RS
 Rhodococcus opacus_YP_00278012 -----D**V**R**NR**RL**AD**VL**GV**AV**SV**YRLA**PE**HP**WP**AA**PD**CEAAAL**WL**AE**HA**ER**FG**TT---R**LS**IG**FS**AG**AT**LA**MT**LL**LR**DR
 Penicillium chrysogenum_XP_002 -----D**ST**L**RE**Y**AN**K**C**Q**L**T**AI**SV**YRLA**PE**HP**Y**PA**V**HD**DA**AE**Y**MD**HA**VE**Y**GA**P---L**RF**IG**ES**AG**CL**AT**LS**AL**Q**L**IR**S
 Mycobacterium smegmatis_YP_889 -----D**V**R**NR**RL**V**NAL**GV**AV**SV**YRLA**PE**HP**WP**AA**PD**CEAAAL**WL**V**EN**AE**QR**FG**TT**---L**RV**IG**LS**AG**ST**LV**TT**LL**VR**DR
 ORF30 -----D**PML**ER**IA**Q**NA**GLACV**SV**YRLA**PE**HP**YP**AG**PD**CEAAAL**WL**V**NA**K**TE**FG**TE**---V**LT**IGESAG**HL**SA**VT**LL**IR**MR**K**

Figura 26. Alinhamento da sequência da ORF30 com as 32 sequências extraídas do NCBI. As caixas representam as regiões conservadas das enzimas lipolíticas; • Possíveis resíduos de aminoácidos envolvidos na tríade catalítica.

5.6 Análise das ORFs encontradas nos clones PI17.E10 e PI32.D09 com sequências patenteadas

As sequências das cinco ORFs encontradas nos clones PL17.E10 e PI32.D09 foram comparadas com sequências patenteadas depositadas no NCBI, usando a ferramenta BLASTP. A ORF30 do clone PI32.D09 teve 40% de identidade com uma proteína de *Pseudomonas syringae* relacionada com patogenicidade de plantas (número de acesso ABT06657.1, patente da CAMBIA 7220583). Quanto as ORFs do clone PI17.E110, a ORF8 apresentou 42% com uma esterase (número de acesso ABY17591.1; patente CAMBIA 7288400); e as ORF16, 17 e 21 tiveram identidade de 45%, 44% e 48%, respectivamente, com uma proteína não nomeada da Bayer Chemicals (número de acesso CAG29902.1; Patente EP1418237). No entanto, apesar desta proteína não estar classificada, o título da sequência depositada faz menção a Lipases Anti-Kazlauskas. A regra de Kazlauskas é uma teoria recente que explica a enantiosseletiva das enzimas, e prevê uma diferença na estrutura do sítio ativo das lipases.

Após as comparações, 11 sequências patenteadas de Esterase/Lipase foram selecionadas e extraídas do NCBI para análise das relações filogenéticas com as cinco ORF encontradas nos clones PL17.E10 e PI32.D09, através do Clustal W e Mega 4. 1, já descrito.

Através da construção filogenética foi possível observar que as ORFs identificadas neste trabalho não são iguais às sequências já patenteadas, fato este corroborado com o valor das identidades. Assim, a ORF8, do clone PI17.E10, ficou em um ramo com a sequência de Esterase/Lipase da Monsanto (Patente 6833447; número de acesso AAW92935.1), que está relacionada com uma região codificadora de Esterase/Lipase identificada no genoma parcial de *Myxococcus xanthus*. Já a ORF 30, do clone PI32.D09, ficou no ramo da sequência de alfa/beta hidrolase do micro-organismo *Enterobacter* sp, da Companhia DAISO (número de acesso CAG26813.1; Patente EP1408107). Quanto às ORFs 16, 17 e 21 do clone PI17.E10, todas ficaram juntas no ramo com a sequência para a proteína não nomeada da Bayer Chemicals (Patente EP1418237; número de acesso CAG29902.1).

Figura 27. Dendrograma do agrupamento hierárquico baseado nas 11 sequências patenteadas e extraídas do NCBI, com as ORFs encontradas nos clones PI17.E10 e PI32.D09.

5.7 Clonagem das ORFs 16 e 17 em vetor pET28a

As ORFs 16 e 17 identificadas no clone pl17.E10 são constituídas de 300 resíduos de aminoácidos, 906 nucleotídeos, sendo que, a primeira está localizada na posição 19250 à 20155 e a segunda na posição 20161 à 21066 dentro no *contig* de 27485 pb (descrito na Tabela 3). Elas foram selecionadas para clonagem em vetor de expressão pET28a, visando futura caracterização enzimática. As demais ORFs, apesar de não terem sido clonadas em vetor de expressão, apresentam potencial para trabalhos posteriores.

5.7.1 Reações de PCR e sequenciamento

As reações de PCR que foram feitas com os oligonucleotídeos iniciadores de síntese das ORFs 16 e 17 foram submetidas a termociclador de gradiente a fim de definir a temperatura de pareamento. Através da análise em gel de agarose foi possível selecionar a temperatura 35 °C para o oligonucleotídeo da ORF16 e a de 50 °C para o oligonucleotídeo da ORF17 (Figura 28). Estas temperaturas foram escolhidas pois propiciaram a amplificação das regiões de interesse, aproximadamente 1000 pb, e não apresentaram produtos de PCR inespecíficos.

Figura 28. Perfis eletroforéticos em gel de agarose 1% mostrando padronização da temperatura de pareamento das reações de PCR em termociclador de gradiente. A- Reação de PCR da ORF 16; B- Reação de PCR da ORF 17. 1- Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2 a 13- Variação da temperatura de pareamento (35 °C a 61 °C).

Após definir as temperaturas de pareamentos de ambos pares de oligonucleotídeos foi feita reação de PCR com a enzima *Pfu* DNA Polimerase. Os produtos desta reação foram purificados com o Kit Abstract for Wizard SV Gel and PCR Clean-Up System (Promega) e a análise em gel de agarose possibilitou verificar que tanto o produto da ORF16, quanto o da ORF17 estavam aptos para o sequenciamento e clonagem em pET28a (Figura 29).

As sequências geradas no sequenciamento das ORFs 16 e 17 foram comparadas com o banco de dados NCBI, sendo que, ambas apresentaram identidade com alfa/beta hidrolase (número de acesso YP_001412950.1) de *Parvibaculum lavamentivorans*, 85% e 77%, respectivamente, confirmando a amplificação do gene de interesse.

Figura 29. Perfil eletroforético em gel de agarose 1% da reação de PCR purificada com o Kit Abstract for Wizard SV Gel and PCR Clean-Up System (Promega). 1- Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2- Reação de PCR da ORF 16 com a enzima *Pfu* DNA polimerase; 3- Reação de PCR da ORF 17 com a enzima *Pfu* DNA polimerase.

5.7.2 Digestão dos produtos de PCR, ligação em vetor de expressão pET28a e transformação em célula competente *Escherichia coli* BL21

Os produtos de PCR purificados das ORFs 16 e 17 foram digeridos com as enzimas de restrição *EcoRI* e *XhoI*. Após a digestão foi feita purificação do material, sendo que, a quantificação no NanoDropTM 1000 Spectrophotometer permitiu verificar a concentração de 17,5 ng/μL e relação 260/280 de 1,46 para o produto de PCR digerido da ORF16; e concentração de 15 ng/μL e relação de 1,40 para o produto de PCR digerido da ORF17.

Os produtos de PCR purificados das duas ORFs foram inseridos em vetor de expressão pET28a desfosforilado e digerido com *EcoRI* e *XhoI*. Posteriormente, os plasmídeos recombinantes foram transformados na célula competente *Escherichia coli* BL21. Para cada clonagem foram coletados cinco clones, que após o cultivo foram estocados a -80 °C.

5.7.3 Extração do pET28a e quantificação

Foi realizada a extração do DNA plasmidial da construção pET28a + ORF16 e da construção pET28a + ORF17 e a quantificação no NanoDropTM 1000 Spectrophotometer

apontou a concentração de 49,6 ng/μL e relação 260/280 de 1,61 para a construção pET28a + ORF16, e a concentração de 50,2 ng/μL e relação de 1,66 para construção pET28a + ORF17.

5.7.4 Confirmação da clonagem

Os DNAs extraídos foram digeridos com as enzimas de restrição *EcoRI* e *XhoI* para a confirmação das construções, sendo que, através da análise em gel de agarose do produto digerido verifica-se a banda do vetor pET28a, de 5369 bp, e as bandas das ORFs 16 e 17, de aproximadamente 1000 pb, afirmando ambas construções (Figura 30).

Figura 30. Perfil eletroforético em gel de agarose 1% da reação de digestão com o DNA dos clones das construções pET28a + ORF16 e pET28a + ORF17, com as enzimas de restrição *EcoRI* e *XhoI*. 1 e 4- Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2- Construção pET28a + ORF16 não digerida; 3- Construção pET28a + ORF16 digerida; 5- Construção pET28a + ORF17 não digerida; 6- Construção pET28a + ORF17 digerida.

5.7.5 Análise *in vitro* dos clones pET28a + ORF16 e pET28a + ORF17

Os clones das construções pET28a + ORF16 e pET28a + ORF17 e a *Escherichia coli* BL21 contendo o vetor pET28a fechado (controle negativo) foram analisadas em placa de petri com meio de cultura adicionado de tributirina, a fim de avaliar presença de atividade lipolítica.

Da mesma forma que os clones da biblioteca que foram selecionados, as duas construções apresentaram formação de halos a partir do terceiro dia de cultivo, sendo mais visível após estoque a 4 °C (Figura 31). O controle negativo não apresentou formação de halo, já a construção pET28a + ORF17 formou halo nítido e maior do que a construção pET28a + ORF16. As análises *in vitro* dos clones pET28a + ORF16 e pET28a + ORF17 permitiram conformar novamente o êxito das clonagens e verificar atividade lipolítica à baixa

temperatura. O estudo de enzimas que atuam em temperatura baixa é importante pois, em comparação com as mesofílicas, as psicofílicas possuem maior flexibilidade ao redor do sítio ativo, devido a fatores como: diminuição das pontes dissulfeto e das interações iônicas eletrostáticas (SIDDIQUI e CAVICCHIOLI, 2006).

Dentro deste contexto, estas enzimas apresentam adaptações das suas propriedades para elevada atividade a baixas temperaturas, baixa temperatura ótima para catálise e relativa termoestabilidade (D'AMICO et al., 2006), que oferecem grande potencial à indústria e biotecnologia.

Figura 31. Análise *in vitro* em placa de petri dos clones das construções pET28a + ORF16 e pET28a + ORF17 e da *Escherichia coli* BL21 contendo o vetor pET28a fechado (controle negativo), no quarto dia a 4 °C, em meio LB suplementado com: 1% de tributirina (v/v); 1% de goma arábica (p/v); 0,001% de Rhodamine-B (v/v); 50 µg/mL de kanamicina; e adição de 100 µL de IPTG [2,4% (p/v) em água] após a solidificação do meio. A- Controle negativo; B- Clone pET28a + ORF16; C- Clone pET28a + ORF17.

6 CONCLUSÕES

- ❖ A seleção dos clones da biblioteca metagenômica de consórcio degradador de óleo diesel foi feita através da expressão funcional da atividade lipolítica e este método, juntamente com as análises de bioinformática, permitiu a descoberta de novos genes;
- ❖ Foram selecionados dois clones que apresentam formação de halo nítido e o DNA fosmidial fragmentado dos mesmos foi subclonado em plasmídeo. A técnica mostrou-se eficiente para obter o *contig* completo de ambos;
- ❖ O clone P117.E10 apresenta 25 ORFs, sendo que, quatro foram confirmadas como novas representantes das enzimas lipolíticas. Uma ORF pertence à família IV e as outras três à família V;
- ❖ O clone P132.D09 possui 32 ORFs, no qual uma é representante da família IV das enzimas lipolíticas;
- ❖ A caracterização dos genes encontrados no clone P117.E10 e P132.D09 possibilitou verificar a presença da tríade catalítica na sequência de aminoácidos, que é essencial para lipases e esterases. Esta análise foi de encontro com o resultado obtido nas construções das árvores filogenéticas;
- ❖ Este estudo demonstrou que a abordagem metagenômica pode ser muito utilizada para expandir o conhecimento de enzimas, especialmente de lipases bacterianas.
- ❖ Futuros estudos de caracterização enzimática destas novas enzimas lipolíticas serão interessantes a fim de prever as possíveis aplicações biotecnológicas.

REFERÊNCIAS*

ARPIGNY, J. L.; JAEGER, K. E. Bacterial lipolytic enzymes: classification and properties. **Biochemical Journal**, v. 343, p. 177-183, 1999

ASTSCHUL, S. F.; MADDEN, T. L.; SCHAFFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D.; Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. **Nucleic Acids Research**, v. 25, p. 3389-3402, 1997.

CANTÃO, M. E.; FERREIRA, J. E.; LEMOS, E. G. M. Optimal clone identifier for genomic shotgun libraries: OC Identifier tool. **Genetics and Molecular Research**, v. 6, n. 4, p. 743-755, 2007.

CHU, X.; HE, H.; GUO, C.; SUN, B. Identification of two novel esterases from a marine metagenomic library derived from south China sea. **Applied Microbiology and Biotechnology**, v. 80, p. 615-625, 2008.

CIESLINSKI, H.; BIALKOWSKA, A.; TKACZUK, K.; DLUGOLECKA, A.; KUR, J.; TURKIEWICZ, M. Identification and molecular modeling of a novel lipase from an antarctic soil metagenomic library. **Polish Journal of Microbiology**, v. 58, n. 3, p. 199-204, 2009.

COUTO, G. H.; GLOGAUER, A.; FAORO, H.; CHUBATSU, L. S.; SOUSA, E. M.; PEDROSA, F. O. Isolation of a novel lipase from a metagenomic library derived from mangrove sediment from the south Brazil coast. **Genetics and Molecular Research**, v. 9, n. 1, p. 514-523, 2010.

D'AMICO, S.; COLLINS, T.; MARX, J. C.; FELLER, G.; GERDAY, C. Psychrophilic microorganisms: challenges for life. **EMBO Reports**, v. 7, p. 385-389, 2006.

ELEND, C.; SCHMISSER, C.; LEGGEWIE, C.; BABIAK, P.; CARBALLEIRA, J. D.; STEELE, H. L.; REYMOND, J. -L.; JAEGER, K. -E.; STREIT, W. R. Isolation and biochemical characterization of two novel metagenome-derived esterases. **Applied and Environmental Microbiology**, v. 72, n. 5, p. 3637-3645, 2006.

* De acordo com: ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: Informação e documentação: referências: elaboração. Rio de Janeiro, 2002.

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

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

GAO, B.; SU, E.; LIN, J.; JIANG, Z.; MA, Y.; WEI, D. Development of recombinant *Escherichia coli* whole-cell biocatalyst expressing a novel alkaline lipase-coding gene from *Proteus* sp. for biodiesel production. **Journal of Biotechnology**, v. 139, p. 169-175, 2009.

HAMME, J.; SINGH, A.; WARD, O. Recent advances in petroleum microbiology. **Microbiology and Molecular Biology Reviews**, v. 67, n. 4, p. 503-549, 2003.

HANDELSMAN, J.; RONDON, M. R.; BRADY, S. F. ; CLARDY, J.; GOODMAN, R. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. **Chemistry & Biology**, v. 5, n. 10, p. R245-R249, 1998.

HARDEMAN, F.; SJOLING, S.; Metagenomic approach for the isolation of a novel low-temperature-active lipase from uncultured bacteria of marine sediment. **FEMS Microbiology Ecology**, v. 59, p. 524-534, 2007.

HENNE, A.; SCHMITZ, R. A.; BOMEKE, M.; GOTTSCHALK, G.; DANIEL, R. Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. **Applied Environmental Microbiology**, v. 66, p. 3113-3116, 2000.

HONG, K. S.; LIM, H. K.; CHUNG, E. J.; PARK, E. J.; KIM, J. C.; CHOI, G. J.; CHO, K. Y.; LEE, S. W. Selection and characterization of forest soil metagenome genes encoding lipolytic enzymes. **Journal of Microbiology and Biotechnology**, v. 17, p. 1655-1660, 2007.

HU, Y.; FU, C.; HUANG, Y.; YIN, Y.; CHENG, G.; LEI, F.; LU, N.; LI, J.; ASHFORTH, E. J.; ZHANG, L.; ZHU, B. Novel lipolytic genes from the microbial metagenomic library of the South China Sea marine sediment. **FEMS Microbiology Ecology**, v. 72, p. 228-237, 2010.

JAEGER, K. E.; DIJKSTRA, B. W.; REETZ, M. T. Bacterial Biocatalysts: Molecular Biology, Three-Dimensional Structures, and Biotechnological Applications of Lipase. **Annual Review Microbiology**, v. 53, p. 315-351, 1999.

JONES, B. V.; SUN, F.; MARCHESI, J. R. 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.

KIM, E. Y.; OH, K. H.; LEE, M. H.; KANG, C. H.; OH, T. K.; YOON, J. H. Novel cold-adapted alkaline lipase from an intertidal flat metagenome and proposal for a new family of bacterial lipases. **Applied and Environmental Microbiology**, v. 75, n. 1, p. 257-260, 2009.

KOBLITZ, M. G. B.; PASTORE, G. M. Purificação parcial, por dois diferentes métodos cromatográficos, da lipase produzida por *Rhizopus* sp. **Ciência e Tecnologia de Alimentos**, v. 24, n. 2, p. 287-292, 2004.

KOUKER, G.; JAEGER, K. E. Specific and sensitive plate assay for bacterial lipases. **American Society for Microbiology**, v. 53, n. 1, p. 211-213, 1987.

LEE, S. W.; WON, K.; LIM, H. K.; KIM, J. C.; CHOI, G. J.; CHO, K. Y. Screening for novel lipolytic enzymes from uncultured soil microorganisms. **Applied Microbiology and Biotechnology**, v. 65, p. 720-726, 2004.

LEE, M. H.; LEE, C. H.; OH, T. K.; SONG, J. K.; YOON, J. H. Isolation and characterization of a novel lipase from a metagenomic library of tidal flat sediments: Evidence for a new family of bacterial lipases. **Applied and Environmental Microbiology**, v. 72, p. 7406-7409, 2006.

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.

LORENZ, P.; LIEBETON, K.; NIEHAUS, F.; ECK, J.; Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. **Current Opinion in Biotechnology**. Elsevier v. 13, p. 572-577, 2002.

LORENZ, P.; ECK, J. Metagenomics and industrial applications. **Nature**, v. 3, p. 510-516, 2005.

MARTÍNEZ, A.; SOBERÓN-CHÁVEZ, G. Characterization of the *lipA* gene encoding the major lipase from *Pseudomonas aeruginosa* strain B83. **Applied Microbiology and Biotechnology**, v. 26, p. 731-735, 2001.

NARDINI, M.; DIJKSTRA, B. W. α/β Hydrolase fold enzymes: the family keeps growing. **Current opinion in Structural Biology**, v. 9, p. 732-737, 1999.

OLLIS, D. L.; CHEAH, E.; CYGLER, M.; DIJKSTRA, B.; FROLOW, F.; FRANKEN, S. M.; HAREL, M.; REMINGTON, S. J.; SILMAN, I.; SCHRAG, J. The alpha/beta hydrolase fold. **Protein Engineering, Design and Selection**, v. 5, n. 3, p. 197-211, 1992.

PAIXÃO, D. A.; DIMITROV, M. R.; PEREIRA, R. M.; ACCORSINI, F. R.; VIDOTTI, M. B.; LEMOS, E. G. M. Análise molecular da diversidade bacteriana de um consórcio degradador de óleo diesel. **Revista Brasileira de Ciência do Solo**, v. 34, p. 773-781, 2009.

PEDRINHO, E. A. N.; LEMOS, E. G. M.; PEREIRA, R. M.; SCAQUITTO, D. C.; SILVEIRA, É. L.; VAL-MORAES, S. P.; CARARETO-ALVES, L. M.; WICKERT, E.; VALARINI, M. J. Avaliação do impacto do lodo de esgoto na microbiota do solo utilizando o gene 16S rRNA. **Arquivos do Instituto Biológico**, v. 76, n. 3, p. 443-448, 2009.

RANJAN R.; GROVER, A.; KAPARDAR, R. K.; SHARMA, R. Isolation of novel lipolytic genes from uncultured bactéria of pond water. **Biochemical and Biophysical Research Communications**, n. 335, p. 57-65, 2005.

REES, H. C.; GRANT, S.; JONES, E. B.; GRANT, W. D.; HEAPHY, E. S. Detecting cellulase and esterase enzyme activities encoded by novel genes present in environmental DNA libraries. **Extremophiles**, v. 7, p. 415-421, 2003.

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. **Applied and Environmental Microbiology**, v. 71, n. 2, p. 817-825, 2005.

ROBLES, J.; DOERS, M. PGEM^R-T Vector Systems troubleshooting guide. **Promega Notes**, Madison, v. 45, p. 19, 1994.

ROH, C.; VILLATTE, F. Isolation of a low-temperature adapted lipolytic enzyme from uncultivated micro-organism. **Journal of Applied Microbiology**, n. 105, p. 116-123, 2008.

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 Environmental Microbiology**, v. 66, n. 6, p.2541–2547, 2000.

RUIZ, C.; BLANCO, A.; PASTOR, F. I. J.; DIAZ, P. Analysis of *Bacillus megaterium* system and cloning of LipA, a novel subfamily I.4 bacterial lipase. **FEMS Microbiology Letters**, v. 217, p. 263-267, 2002.

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

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

SAXENA, R. K.; SHEORAN, A.; GIRI, B.; DAVIDSON, W. S. Purification strategies for microbial lipases. **Journal of Microbiological Methods**, v. 52, n. 1, p. 1-18, 2003.

SCHLEHECK, D.; DONG, W.; DINGER, K.; HEINZLE, E.; COOK, A. M. An α -Proteobacterium Converts Linear Alkylbenzenesulfonate Surfactants into Sulfophenylcarboxylates and Linear Alkyldiphenyletherdisulfonate Surfactants into Sulfodiphenylethercarboxylates. **Applied and Environmental Microbiology**, v. 66, p. 1911-1916, 2000.

SIDDIQUI, K. S.; CAVICCHIOLI, R. Cold-adapted enzymes. **Annual Review of Biochemistry**, v. 75, p. 403-433, 2006.

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

STEELE, H.; STREIT, W. R. Metagenomics: Advances in ecology and biotechnology. **FEMS Microbiology Letters**, v. 247, p. 105-111, 2005.

TAMURA, K.; DUDLEY, J.; NEI, M.; KUMAR, S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. **Molecular Biology and Evolution**, v. 24, p. 1596-1599, 2007.

TIRAWONGSAROJ, P.; SRIPRANG, R.; HARNPICHARNCHAI, P.; THONGARAM, T.; CHAMPREDA, V.; TANAPONGPIPAT, S.; POOTANAKIT, K.; EURWILAICHITR, L. Novel thermophilic and thermostable lipolytic enzymes from a Thailand hot spring metagenomic library. **Journal of Biotechnology**, v. 133, p. 42-49, 2008.

TREVISAN, H. C. Lipases. In: SAID, S.; PIETRO, R. C. L. R. **Enzimas como agentes biotecnológicos**. Ribeirão Preto: Legis Summa, 2004.

VAKHLU, J.; KOUR, A. Yeast lipases: enzyme purification, biochemical properties and gene cloning. **Journal of Biotechnology**, v. 9, n. 1, p. 69-85, 2006.

VOGET, S.; LEGGEWIE, A.; UESBECK, A.; RAASCH, C.; JAEGER, K. -E.; STREIT, W. R. Prospecting for novel Biocatalysts in a soil metagenome. **Applied and Environmental Microbiology**, v. 69, n. 10, p. 6235-6242, 2003.

WEI, Y.; CONTRERAS, J. A.; OSTERLUND, T.; DEREWENDA, U.; KNEUSEL, R. E.; MATERN, U.; DEREWENDA, Z. S. Crystal structure of brefeldin A esterase, a bacterial homolog of the mammalian hormone sensitive lipase. **Nature Structural Biology**, v. 6, n. 4, p. 340-345, 1999.

WILD, J.; HRADECNA, Z.; SZYBALSKI, W. Conditionally amplifiable BACs: Switching from single-copy to high-copy vectors and genomic clones. **Genome Research**, v. 12, p. 1434-1444, 2002.

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.

WU, C.; SUN, B. Identification of novel esterase from metagenomic library of Yangtze River. **Journal of Microbiology and Biotechnology**, v. 19, n. 2, p. 187-193, 2009.

WU, H. S.; TSAI, M. J. Kinetics of tributyrin hydrolysis by lipase. **Enzyme and Microbial Technology**, v. 35, p. 488-493, 2004.

ANEXO A - Comparação do *contig* completo do clone PI17.E10, de 27485 pb, com as sequências do NCBI, usando a ferramenta BLASTN (nucleotide collection nr/nt).

ANEXO B - Comparação do *contig* completo do clone PI132.D09, de 35631 pb, com as sequências do NCBI, usando a ferramenta BLASTN (nucleotide collection nr/nt).