THAÍS CARVALHO MAESTER

PROSPECÇÃO DE SEQUÊNCIAS GENÔMICAS CODIFICADORAS DE ENZIMAS LIPOLÍTICAS DEGRADADORAS DE HIDROCARBONETOS DE PETRÓLEO

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

THAÍS CARVALHO MAESTER

PROSPECÇÃO DE SEQUÊNCIAS GENÔMICAS CODIFICADORAS DE ENZIMAS LIPOLÍTICAS DEGRADADORAS DE HIDROCARBONETOS DE PETRÓLEO

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: Profa. 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

Maester, Thaís Carvalho.

Prospecção de sequências genômicas codificadoras de enzimas lipolíticas degradadoras de hidrocarbonetos de petróleo / Thaís Carvalho Maester. -- 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.

Versão do título para o inglês: Screening for genomic sequences which codify lipolytic enzymes specialized in petroleum hydrocarbons degradation.

Descritores: 1. Lipase 2. Esterase 3. Tributirina 4. Sítios catalíticos 5. Patentes I. Lemos, Eliana Gertrudes de Macedo II. Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação Interunidades em Biotecnologia USP/IPT/Instituto Butantan III. Título.

ICB/SBIB047/2011

UNIVERSIDADE DE SÃO PAULO **Programa de Pós-Graduação Interunidades em Biotecnologia** Universidade de São Paulo, Instituto Butantan, Instituto de Pesquisas Tecnológicas

Candidato(a):	Thaís Carvalho Maester.
Título da Dissertação:	Prospecção de sequências genômicas codificadoras de enzimas lipolíticas degradadoras de hidrocarbonetos de petróleo.
Orientador(a):	Eliana Gertrudes de Macedo Lemos.

() Aprovado(a) () Reprovado(a)

Examinador(a):	Assinatura: Nome: Instituição:
Examinador(a):	Assinatura: Nome: Instituição:
Presidente:	Assinatura: Nome: Instituição:

Aos meus pais, Maria Helena Carvalho Maester e Felipe Stainle Maester, por todo o amor, ajuda e inspiração.

Aos meus irmãos, Luciana e André, pela proteção e carinho dedicados.

Ao meu namorado, Diogo Cavenague Casanova, pela atenção, cuidados e amor.

Aos meus cunhados, Carlos Batista Alves e Micheli Viscardi Maester, pelo apoio e atenção.

Aos meus sobrinhos, Higor e Letícia, por serem minha inspiração.

AGRADECIMENTOS

À Profa. Dra. Eliana Gertrudes de Macedo Lemos, pela orientação, pela confiança depositada e por toda a atenção a que me foi dedicada;

Ao Dr. João Carlos Campanharo pela amizade, auxílio e pelos constantes ensinamentos;

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

Ao Prof. Dr Luiz Juliano Neto, Professor da Universidade Federal de São Paulo, pelo auxílio na condução do experimento;

A todos os integrantes do Laboratório de Bioquímica de Microrganismos e Plantas, pela amizade, pelo auxílio e troca de experiências, que contribuíram para o desenvolvimento do projeto;

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

À minha amiga mestranda, Erica Mendes Lopes, por ter me ajudado constantemente, sempre que precisei;

À futura doutoranda Elisângela Soares Gomes, pelos ensinamentos e auxílio;

À minha querida amiga Mariana Rangel Pereira, que foi fundamental para a realização deste trabalho, a quem muito me ensinou e me ajudou. Dividimos a mesma casa, mesmo laboratório por muito tempo e espero ainda poder compartilhar de muitas alegrias;

À Fundação de Amparo à Pesquisa do Estado de São Paulo pela bolsa concedida;

Aos meus pais, pelo incentivo e por toda a dedicação. Ao meu pai, por toda a preocupação e pela inspiração. À minha mãe, pelos conselhos e pela fé;

À minha irmã Luciana, pelo seu enorme coração e ao meu cunhado Carlos, pelo carinho. Obrigada pela hospedagem;

Ao meu irmão, André, pelo carinho e pela inspiração;

Ao Thor, por ter me feito parte da minha vida, pelo companheirismo, pela alegria que me trouxe sempre que precisei, por fazer sentir-me realmente especial.

A todos os meus familiares que participaram comigo de tantos momentos, trazendo amor, carinho e confiança.

Ao meu namorado, Diogo, pela atenção, incentivo, paciência e por todo seu amor;

Às minhas amigas da República Toca da Onça, Onçona, Oncinha, Sufrida, Gosminha, Páçoká, Pintada e Miss-kenta por compartilharem comigo muitos momentos bons, poucos ruins, pelas risadas, conversas e pela amizade.

A todos que contribuíram direta ou indiretamente para a realização deste projeto.

"A percepção do desconhecido é a mais fascinante das experiências.

O homem que não tem os olhos abertos para o misterioso

passará pela vida sem ver nada".

Albert Einstein

Este projeto foi realizado no Laboratório de Bioquímica de Microrganismos e Plantas, no Departamento de Tecnologia da Faculdade de Ciências Agrárias e Veterinárias da UNESP, Campus de Jaboticabal, sendo financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo- FAPESP, número do processo 2009/03772-9

RESUMO

MAESTER, Thaís Carvalho. **Prospecção de sequências genômicas codificadoras de enzimas lipolíticas degradadoras de hidrocarbonetos de petróleo**. 2011. 81 f. Dissertação (Mestrado em Biotecnologia) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2011.

As enzimas lipolíticas possuem enorme potencial biotecnológico, seja para formulação de detergentes, na indústria de couro, na produção de cosméticos, fármacos, aromas, biodiesel, dentre outros. A maior parte das lipases é de origem microbiana, apresenta baixa toxicidade e é facilmente biodegradável. 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 hidrocarbonetos de petróleo, com 4224 clones. A seleção foi feita pela atividade lipolítica através do cultivo dos clones em meio de cultura Luria-Bertani (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). As células permaneceram a 37 °C por 72 horas e depois foram transferidas a 4 °C. A avaliação foi realizada pela observação da formação de halo ao redor das colônias, sendo positiva para 30 clones, dentre os quais dois foram selectionados e tiveram o DNA sub-clonado em vetor pUC19. Os DNAs das subbibliotecas foram sequenciados, gerando um *contig* completo para o clone PL28.F10, que foi comparado com as sequências do banco National Center for Biotechnology Information (NCBI), através do programa ORF Finder. Uma ORF codificadora de esterase/lipase de 303 aminoácidos e 61% de identidade com micro-organismo não cultivável foi encontrada. Foram feitos alinhamentos com o programa Clustal W e construção de árvores filogenéticas pelo programa MEGA 4, para comparação entre a ORF encontrada, ORF15, e as sequências depositadas. As árvores indicam que o clone apresenta a ORF15 mais próxima da família IV das esterases/lipases, pois de localizou em um ramo único, diferenciando-se dos outros representantes desta família. Através dos alinhamentos foi possível identificar os sítios ativos representativos da família, confirmando o resultado das árvores filogenéticas, no entanto, foram verificadas algumas modificações nos resíduos de aminoácidos. Quando a mesma análise foi realizada com sequências já patenteadas, a ORF15 se localizou em um ramo único, sendo um grupo irmão das sequências de esterases/lipases da BASF (Patente WO0100842), relacionada à biossíntese de FK228 e seus análogos, utilizados no tratamento do câncer e de uma proteína não identificada da CAMBIA (Patente 6562958, número de acesso AAQ29806.1)., utilizada para diagnóstico e terapia, que possui o domínio conservado das esterases/lipases

Palavras-chave: Lipase. Esterase. Tributirina. Sítios catalíticos. Patentes.

ABSTRACT

MAESTER, Thaís Carvalho. Screening for genomic sequences which codify lipolytic enzymes specialized in petroleum hydrocarbons degradation. 2011. 81 p. Masters thesis (Biotechnology) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2011.

Lipolytic enzymes have show enormous biotechnological potential, such as in enzyme mixtures for the production of detergents, the processing of leather, 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. The present work was done as an attempt to find genes which codify lipolytic enzymes in a fosmid metagenomic library composed of petroleum hydrocarbons degradation microbe consortia with 4224 clones. Clones were selected according to lipolytic activity and were assessed by cultivation in Luria-Bertani (LB) medium supplemented with1% tributyrin (v/v), 1% gum arabic (w/v), 0.00125% chloramphenicol (v/v) and 0.001% arabinose (w/v). All cultures were maintained at 37°C for 72 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 were selected and had their DNA sub cloned in pUC19 vectors. DNA from the sub libraries was sequenced, generating a complete contig for clone PL28.F10 that was compared to sequences from the National Center for Biotechnology Information (NCBI) bank by the ORF Finder program. An ORF coding for esterase/lipase of 303 amino acids with 61% of identity with uncunturable microorganism was found. Alignments were done using the Clustal W program and cladograms were constructed using MEGA 4, to compare the ORF found, ORF15, and the deposited sequences. Assessment of the cladograms showed that the clone presented the ORF15 similar to that of esterase/lipase family IV, because it was located in a unique branch, differing from the other representatives of this family. The alignments made possible the identification of active sites which represent the family, confirming the results obtained with the construction of the cladograms, although it was observed some changes in amino acid residues. When the same evaluation was done using patented sequences, the ORF15 showed similarities to patented BASF esterase/lipase (Patent WO0100842) related to biosynthesis of FK228 and its analogues used in cancer treatment and an unnamed protein of CAMBIA (Patent 6562958, accession number AAO29806.1) used for diagnosis and therapy, which has conserved domain esterases/lipases.

Key words: Lipase. Esterase. Tributyrin. Catalytic sites. Patents

LISTA DE FIGURAS

Figura 1- Hidrólise da tributirina pela lipase e formação dos produtos glicerol e ácido
butírico
Figura 2- Estrutura secundária do dobramento das α/β hidrolases. α -hélices e folhas- β
estão representadas por meio de cilindros brancos e setas cinzas, respectivamente. A
localização da tríade catalítica está indicada pelos círculos pretos
Figura 3- Fluxograma metodológico demonstrando as etapas realizadas no trabalho
Figura 4- Esquema ilustrativo apresentando características do vetor pUC19DNA/SmaI
(Fermentas)
Figura 5- Placa de Petri com 30 clones previamente selecionados da biblioteca metagenômica de consórcio microbiano degradador de óleo hidrocarbonetos do natróleo. As cólulos forem incorpledes em meio LD, com artibiótico, suplemente de com
tributirina, goma arábica, arabinose, e Rhodamine-B por três dias a 37 °C. CN-
Controle negativo
Figura 6- Cultivo dos sete clones selecionados com atividade lipolítica em meio LB suplementado com tributirina goma, arábica e arabinose. As células foram incubadas
por três dias a 37 °C. A- Clone PL14.H10; B- Clone PL28.F1047

Figura 11- Perfil eletroforético em gel de agarose 0,8% de DNA plasmidial extraído de48 sub-clones originados da sub-biblioteca do clone PL28.F1052

 Figura 17- Dendrograma do agrupamento hierárquico mostrando a relação filogenética da ORF15 de esterase/lipase, encontrada no *contig* do clone PL28.F10 com outras ORFs codificadoras de esterase/lipase encontradas na mesma biblioteca metagenômica...... 69

Figura 18- Dendrograma do agrupamento hierárquico baseado em 16 sequências patenteadas extraídas no NCBI, com a ORF15 encontrada no clone PL28.F10......70

LISTA DE TABELAS

LISTA DE ABREVIATURAS

- ATP adenosina trifosfato
- BSA albumina de soro bovino
- DMF N, N' dimethyl-formamida
- DNA ácido desoxirribonucléico
- 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 ribonucléico
- RNAse ribonuclease
- SDS dodecil sulfato de sódio
- 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-galactopiranosídeo

LISTA DE UNIDADES

- g- aceleração da gravidade
- g grama
- kb kilobase
- L litro
- m– molar
- mg miligrama
- mL mililitro
- µg micrograma
- μL microlitro
- $\mu M micromolar$
- ng nanograma
- pb pares de bases
- U unidades
- V volt

SUMÁRIO

1 INTRODUÇÃO	17
2 REVISÃO DE LITERATURA	19
2.1 Enzimas	19
2.2 Enzimas microbianas	20
2.3 Diversidade microbiana do solo	20
2.4 Biotecnologia e Metagenoma	21
2.5 Enzimas lipolíticas	23
2.5.1 Dobramento das enzimas lipolíticas	24
2.5.2 Classificação das enzimas lipolíticas	26
3 OBJETIVOS	
4 MATERIAL E MÉTODOS	
4.1 Fluxograma metodológico	
4.2 Obtenção do consórcio microbiano	
4.3 Triagem dos clones com atividade para enzimas lipolíticas	
4.4 Extração do DNA fosmidial	
4.5 Quantificação e análise do DNA	
4.6 Digestão do DNA fosmidial	
4.7 Construção das sub-bibliotecas	
4.7.1 Reação de nebulização	
4.7.2 Precipitação e quantificação do DNA fragmentado	
4.7.3 Reação de reparo das extremidades dos fragmentos	
4.7.4 Reação de fosforilação dos fragmentos reparados	
4.7.5 Seleção do tamanho dos fragmentos inseridos	
4.7.6 Reação de ligação do fragmento inserido ao vetor	
4.7.7 Transformação bacteriana da célula competente	40
4.7.8 Coleta e estoque dos clones transformados	41
4.8 Cultivo dos clones e extração de DNA plasmidial em placa	41
4.9 Reação de sequenciamento	43
4.10 Análise das sequências	43
5 RESULTADOS E DISCUSSÃO	45

5.1 Obtenção do consórcio microbiano
5.2 Triagem dos clones com produção de enzimas lipolíticas
5.3 Extração do DNA fosmidial dos clones selecionados
5.4 Digestão do DNA fosmidial dos clones selecionados
5.5 Construção das sub-bibiotecas
5.6 Análise do clone PL28.F10
5.6.1 Análise da sequência do contig
5.6.2 Análise das relações filogenéticas
5.6.3 Identificação dos sítios catalíticos
5.7 Análise da ORF com sequências patenteadas
6 CONCLUSÕES
REFERÊNCIAS72
ANEXO A- Comparação do <i>contig</i> completo do clone PL28.F10, de 36.942 pb, com
as sequências do NCBI, usando a ferramenta Blastn (nucleotide collection nr/nt)80
ANEXO B- Comparação da ORF15, de 303 aminoácidos, com as sequências do
NCBI, usando a ferramenta Blastp (non-redundant protein sequence – nr)

1 INTRODUÇÃO

Estima-se que há de 10^6 a 10^8 espécies de micro-organismos na Terra (CURTIS et al., 2002) mas, a grande maioria, muito provavelmente, não pode ser cultivada por meio de técnicas convencionais (AMANN et al., 1995). Micro-organismos desempenham papéis importantes no equilíbrio ecológico e devido às alterações ambientais que ocorrem ao longo do tempo, mostram grande diversidade genética (LIAW et al., 2010).

A biotecnologia, na busca por novos biocatalisadores possibilita o desenvolvimento de novas abordagens experimentais para localizar e identificar novos genes codificadores de produtos de interesse biotecnológico. A fim de descobrir novas enzimas e estudar vias metabólicas secundárias, utilizando o genoma de comunidades microbianas complexas de habitats naturais, uma abordagem independente de cultura pode ser aplicada, tal como a construção de uma biblioteca metagenômica (LIAW et al., 2010).

O termo metagenoma foi introduzido para descrever os genomas de comunidades microbianas complexas encontradas em habitats naturais. A investigação de bibliotecas metagenômicas se tornou possível após o desenvolvimento de estratégias para o isolamento e clonagem de DNA de ecossistemas naturais. Uma vez construídas, as bibliotecas metagenômicas permitem a seleção de fenótipos interessantes com potencial ecológico e biotecnológico (HANDELSMAN et al., 1998).

Na busca de novos biocatalisadores, existem várias estratégias metagenômicas que são utilizadas para a seleção de características específicas, como um catalisador que abranja um intervalo significativo da concentração de substrato ou temperatura e pH ótimos. Uma alternativa é gerar a biblioteca metagenômica de solos ou sedimentos que frequentemente contêm alto grau de diversidade microbiana a qual pode expressar uma grande diversidade de biocatalisadores. Esta abordagem tem sido usada com sucesso para encontrar uma grande variedade de novos catalisadores e metabólitos secundários, como lipases (HENNE et al., 2000), esterases (JEON et al., 2009), celulases (KIM et al., 2008), xilanases (HU et al., 2008) e antibióticos (GILLESPIE et al., 2002).

Um aperfeiçoamento desta abordagem pode ser realizado obtendo-se uma biblioteca metagenômica de ambiente que tenha sido submetido a condições extremas, com a possibilidade de que as enzimas desse ambiente sejam capazes de atuar sob tais condições. Além disso, ambientes submetidos a condições extremas podem direcionar a busca por produtos de interesse. Há exemplos de trabalhos em ambientes adversos, como a construção de biblioteca metagenômica de consórcio microbiano degradador de óleo diesel (PAIXÃO et al., 2009), utilizada neste trabalho. Outro exemplo é o isolamento de enzimas lipolíticas de solos da Antártica (CIELINSK et al., 2009) e de lodo ativado (LIAW et al., 2010).

Enzimas lipolíticas são importantes para o uso em muitas indústrias, como biocatalisadores, na indústria alimentícia, para alimentação humana e animal, detergente, síntese química, de cosméticos, indústria farmacêutica entre outros (JAEGER e EGGERT, 2002). Lipases (EC 3.1.1.3) e esterases (EC 3.1.1.1) são enzimas lipolíticas que hidrolisam e sintetizam glicerídeos de cadeia longa e de cadeia curta (ARPIGNY e JAEGER, 1999), respectivamente. Além de sua ampla utilização, enzimas lipolíticas possuem características que facilitam sua aplicação, tais como ampla especificidade ao substrato e estabilidade em solventes orgânicos (BORNSCHEUER, 2002).

Dessa forma, a utilização da abordagem metagenômica se mostra promissora na busca de gene e/ou produtos gênicos que possam ser utilizados ou incorporados em processos industriais, visando melhorias e redução do custo destes processos. As enzimas lipolíticas, por sua ampla utilização, são excelentes candidatas nessa busca.

2 REVISÃO DE LITERATURA

2.1 Enzimas

Avanços técnicos e científicos que ocorreram, principalmente, a partir do século XX favoreceram a utilização das enzimas, sendo que estas passaram a ocupar papéis importantes em produtos e processos industriais, permitindo que fossem empregadas em larga escala (SAID e PIETRO, 2004).

A versatilidade das enzimas possibilita seu uso em diversas aplicações, incluindo processos de catálise de polímeros naturais, tais como amido, celulose e proteínas, bem como para síntese seletiva de produtos químicos assimétricos (LORENZ E ECK, 2005). No futuro, novas aplicações biotecnológicas devem impulsionar o mercado de enzimas industriais. A substituição de catalisadores químicos, em processos industriais, pelas enzimas possibilita redução de tempo e energia, de rejeitos industriais que podem comprometer o meio ambiente revertendo assim, em benefícios econômicos (SAID e PIETRO, 2004).

As enzimas evoluíram durante anos como catalisadores quirais em células vivas tornando-as aptas ao ambiente. Existem diversas reações orgânicas realizadas por enzimas e a correta identificação destas é de importância primordial. Encontrar um biocatalisador adequado em um tempo razoável depende tanto de estratégias de triagem sensíveis e eficientes como da busca pelo maior número possível de genes codificadores nos diferentes organismos (LORENZ et al., 2002).

O comércio global da produção de enzimas foi estimado em 2,3 bilhões de dólares em 2003, cujo principal montante foi divido em detergentes (US\$ 789 milhões), aplicações alimentícias (US\$ 634 milhões), agricultura (US\$ 237 milhões) e outras, incluindo enzimas para produção de tecidos e produtos químicos (US\$ 222 milhões) (LORENZ et al., 2002).

As enzimas são produzidas por todos os seres vivos, em diferentes concentrações, podendo ser de origem vegetal, animal ou microbiana. Diferentemente das enzimas vegetais e animais, a produção de enzimas microbianas não depende diretamente de fatores climáticos, que muitas vezes são incontroláveis. Além disso, soma-se a estes fatores a enorme biodiversidade dos micro-organismos em variados habitats, até mesmo aqueles que vivem em condições extremas (SAID e PIETRO, 2004).

2.2 Enzimas microbianas

O uso de micro-organismos e enzimas microbianas pelo homem no processamento de materiais naturais tem uma longa tradição. Esta capacidade de exploração de biossintéticos microbianos ocorreu naturalmente e foi facilitada pela ubiquidade dos micro-organismos. Desse modo, atualmente, as enzimas são usadas como chaves de ativação de componentes em detergentes, em vários alimentos industrializados, na manufatura de papel, entre outros (LORENZ et al., 2002).

De acordo com Lorenz e Eck (2005), as indústrias buscam genomas inteiros identificados de micro-organismos não cultivados, pois estes podem ser fonte de novos biocatalisadores, possuírem diversidade suficiente e acessível avançando na competição com a química sintética tradicional, além de ser fonte de metabólitos únicos.

A preferência pela química orgânica sintética baseada em enzimas de micro-organismos e não pela clássica catálise química é devida à especificidade (químio, régio e enantiosseletividade) que as enzimas possuem. Adicionalmente, os produtos do metabolismo secundário microbiano são compatíveis tanto em relação à eficiência quanto com a sua interferência ambiental (LORENZ et al., 2002).

Atualmente, as enzimas microbianas são as mais utilizadas industrialmente (UCHYIAMA e MIYAZAKI, 2009), pois, entre outros fatores, os micro-organismos que vivem em condições extremas, seja de pH ou temperatura, sintetizam enzimas que permitem sua sobrevivências nestes habitats. Existem também as técnicas de engenharia genética que simplificam o cultivo de micro-organismos e modificam as propriedades de enzimas já existentes para que estas sejam aprimoradas (SAID e PIETRO, 2004).

2.3 Diversidade microbiana do solo

O solo tem sido um dos ambientes mais estudados, devido à elevada biodiversidade que alguns tipos podem conter (HANDELSMAN et al., 1998; LORENZ et al., 2002). A mais complexa das comunidades reside no solo, que é composto por materiais orgânicos e minerais. Dentre os papéis dos micro-organismos, estão a fixação do nitrogênio atmosférico, reciclagem de nutrientes provenientes da decomposição de plantas e animais, conversão de elementos, como ferro e manganês, em formas que possam ser utilizadas para a nutrição de plantas, além de interferirem na infecção das plantas por certos patógenos (WHITMAN et al., 1998).

De acordo com Torsvik e Ovreas (2002), 1 grama de solo pode conter até 4000 espécies diferentes, dentre as quais, a maior parte não é cultivável. O número de microorganismos varia dentre os diferentes tipos e condições do solo, sendo que os mais numerosos são as bactérias (WHITMAN et al., 1988). Metodologias moleculares contribuem para o desenvolvimento do estudo da diversidade microbiana dos solos, evidenciando o cenário de distribuição dos micro-organismos em diferentes habitats (PEDRINHO et al., 2009). Ritz e Griffths (1994) demonstraram diferentes padrões de diversidade genética entre diferentes solos através da hibridização do DNA total extraído dos mesmos.

Por meio do isolamento, clonagem e análise do DNA microbiano do solo podem ser avaliadas mais detalhadamente a fisiologia e a função dos micro-organismos na natureza. Os genomas da microbiota total encontrada em uma comunidade são denominados de metagenoma. Estudos envolvendo sequências do gene *16S rRNA*, com abordagem metagenômica, se constituíram nos passos iniciais para a ligação entre a fisiologia e função da microbiota total (RONDON et al., 2000).

2.4 Biotecnologia e Metagenoma

Diante da crescente consciência com relação aos problemas ambientais, como a alta dos custos de energia, escassez de recursos fósseis, a poluição ambiental e uma economia globalizada, as indústrias têm buscado na biotecnologia ferramentas para a sustentabilidade (LORENZ et al., 2002). Para isso, são necessárias novas enzimas, produtos e processos que podem estar presentes em micro-organismos ainda não cultivados (LORENZ e ECK, 2005).

Por meio das técnicas atuais da biotecnologia, diversos biocatalisadores têm sido desenvolvidos, sendo cada vez maior a demanda pela descoberta de novos genes codificadores de produtos de interesse biotecnológico (VOGET et al., 2003).

As técnicas tradicionais de cultivo selecionam os micro-organismos capazes de se desenvolver em meios de cultura pré-definidos. Estimativas feitas a partir de estudos em ecologia por meio da biologia molecular apontam que a diversidade microbiana obtida através de técnicas tradicionais de cultivo abrange somente uma fração, geralmente menos de 1% da diversidade presente em amostras ambientais complexas, tais como o solo (AMANN et al, 1995; LORENZ et al., 2002). Apesar disso, por meio do desenvolvimento de culturas puras em meios com substratos específicos, muitos biocatalisadores foram obtidos e comercializados, como a nitrilase (ENGELS et al., 2000), que converte nitrilas em ácidos carboxilícos e aminas e é utilizada na degradação de herbicidas, e a lactamase (TAYLOR et al., 1999), responsável pela resistência de bactérias a antibióticos beta-lactâmicos.

O metagenoma, isolamento direto do DNA total provindo de uma amostra ambiental, foi proposto devido à discrepância existente ao comparar o número de micro-organismos encontrados *in situ*, através do isolamento de culturas puras e o número de colônias observadas *in vitro* (LORENZ et al., 2002). Como prova da existência de um número de micro-organismos ainda não estudados está a descoberta de sequências do gene *16S rRNA* obtidas de amostras ambientais não relatadas anteriormente (HANDELSMAN et al., 1998; RONDON et al., 2000). O pioneiro a utilizar a clonagem direta do genoma de todos os micro-organismos de um determinado habitat foi Schmidt em conjunto com colaboradores (1991), o que mais tarde viria a ser denominado metagenoma (LORENZ et al., 2002).

Esta potente tecnologia pode acelerar o processo de descoberta de genes que codificam novas enzimas e metabólitos secundários de comunidades microbianas sem cultivo prévio dos micro-organismos (VOGET et al., 2003). Desde então, numerosos estudos têm lidado com os métodos de extração de DNA metagenômico a partir de uma variedade de ambientes. De fato, várias enzimas novas foram identificadas que têm atividade e/ou sequências exclusivas dos mais diversos ambientes, como mar, solo, rios, entre outros. Voget et al. (2003) isolaram mais de 15 genes diferentes codificadores de enzimas como agarase, amilase, celulase e lipase. Adicionalmente, o metagenoma também tem por objetivo compreender melhor a ecologia microbiana global. As estapas para a exploração de DNAs metagenômicos são: isolamento do DNA total de determinado habitat, clonagem do DNA em micro-organismos cultiváveis, triagem de micro-organismos com atividade biológica (HANDELSMAN et al., 1998). A seleção destes micro-organismos pode se dar por identificação da atividade enzimática em substratos específicos ou por identificação de sequências, seja por PCR (Reação em Cadeia da Polimerase) ou hibridização por sondas (UCHYIAMA e MIYAZAKI, 2009).

É comum o estudo de ambientes contaminados devido à seletividade de microorganismos de interesse nesses locais. Bibliotecas metagenômicas de ambientes que tenham sido submetidos a condições extremas têm maior probabilidade de que as enzimas encontradas também sejam capazes de agir em tais condições (ELEND et al., 2006).

Silveira et al. (2006) estimaram e compararam a diversidade bacteriana de comunidades em solos de floresta nativa e outra adjacente com arboreto de eucalipto. Utilizando o DNA metagenômico correspondente ao gene *16SrRNA* determinaram a diversidade bacteriana em tais locais. Análises filogenéticas revelaram diferenças entre os tipos de solos e alta diversidade em ambas as comunidades. No solo de arboreto foi encontrada maior diversidade bacteriana em comparação com o solo da área de floresta nativa.

Segundo Paixão et al. (2010), atingir maior conhecimento e compreensão da diversidade de micro-organismos pertencentes à comunidades bacterianas em ambientes adversos é importante para alcançar benefícios econômicos estratégicos, como aqueles que podem ser obtidos com a descoberta de micro-organismos que apresentam potencial para serem utilizados em processos biotecnológicos para biorremediação de áreas contaminadas.

2.5 Enzimas lipolíticas

Enzimas lipolíticas como lipases (EC 3.1.1.3) e esterases (3.1.1.1) são hidrolases que catalisam tanto a hidrólise como a síntese de acilgliceróis e outros ácidos graxos (LEE et al., 2004), conforme demonstrado na figura 1.



Figura 1- Hidrólise da tributirina pela lipase e formação dos produtos glicerol e ácido butírico. Fonte: Wu e Tsai (2004).

Esterases e lipases possuem determinadas características que as tornam o grupo de biocatalisadores com maior aplicação industrial, pois normalmente não requerem cofatores, são estáveis em solventes orgânicos e possuem ampla especificidade de substratos. Tais características permitem que sejam utilizadas na síntese de biopolímeros e biodiesel, na produção de fármacos, agroquímicos e na indústria alimentícia (JAEGER e EGGERT, 2002). Destaca-se seu potencial na biorremediação de rejeitos industriais de origem lipídica e sua utilização em etapas de produção de biodiesel a partir de óleos vegetais (JEON et al., 2009).

As lipases já foram descritas em vegetais, animais e micro-organismos, sendo que as microbianas são as mais utilizadas industrialmente, devido às diferentes atividades que apresentam e suas propriedades cinéticas. Lee et al. (2004) construíram uma biblioteca metagenômica de micro-organismos de solo de floresta e encontraram oito clones ativos em tributirina como substrato, um triacilglicerol com três moléculas de ácido butírico esterificadas com o glicerol, de um total de 33.700 clones.

São representantes da família α/β hidrolases, assim como as proteases, desalogenases, peroxidases e epóxido hidrolases, o que faz desta família uma das mais versáteis e com dobramentos protéicos mais conhecidos (NARDINI e DIJKSTRA, 1999).

2.5.1 Dobramento das enzimas lipolíticas

As enzimas que constituem a família das α/β hidrolases 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 α/β hidrolase é caracterizado por um núcleo de oito folhas β conectadas por α -hélices, para dar a disposição $\alpha/\beta/\alpha$ (Figura 2). Na maioria dos membros da família, as folhas β são paralelas, mas alguns mostram uma inversão nessa ordem, tendo por resultado a orientação antiparalela. Para manter a maquinaria catalítica conservada, grandes inserções foram toleradas, como alguns resíduos de aminoácidos ou até mesmo um domínio extra completo, dando aos membros desta família uma habilidade impressionante para adaptação e evolução (NARDINI e DIJKSTRA, 1999).

O dobramento das α/β 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 membro nucleofílico (serina, cisteína ou ácido aspártico) posicionado após a folha β 5; um resíduo ácido, posicionado quase sempre após a folha β 7; e uma histidina, que é um resíduo absolutamente conservado e está situado após a última folha β (NARDINI e DIJKSTRA, 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 oxiânion, 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 β 7. Já a histidina é o único resíduo da tríade que é absolutamente conservado, entretanto, a forma e o comprimento do local que a contém pode diferir consideravelmente entre os vários membros da família.



Figura 2- Estrutura secundária do dobramento das α/β hidrolases. α-hélices e folhas-β estão representadas por meio de cilindros brancos e setas cinzas, respectivamente. A localização da tríade catalítica está indicada pelos círculos pretos. Fonte: Nardini e Dijkstra (1999).

2.5.2 Classificação das enzimas lipolíticas

As enzimas lipolíticas são classificadas em diferentes classes por Arpigny e Jaeger (1999). São classificadas em fosfolipases, que possuem especificidade de ligação com fosfoglicerídeos, hidrolisando um grupo acila dos mesmos, carboxilesterases, que hidrolisam pequenas cadeias de ésteres com moléculas ao menos parte solúvel em água e em verdadeiras

lipases, que possuem máxima atividade em longas cadeias de de triglicerídeos insolúveis. Considerando a dificuldade de classificação baseada no mecanismo de reação, estes mesmos autores as classificaram de acordo com suas sequências, seus motivos conservados, suas estruturas tridimensionais e propriedades biológicas.

Sendo assim, de acordo com este sistema de classificação, as lipases e esterases de origem bacteriana foram classificadas em oito famílias distintas. A família I é representada pelas lipases verdadeiras e é subdividida em seis sub-famílias.

A sub-família I.1 é representada por lipases que apresentam elevada similaridade de sequências com a lipase de *Pseudomonas aeruginosa*, que inicialmente deu origem à classificação da família I. Esta sub-família inclui enzimas de *Vibrio cholerae*, *Acinetobacter calcoaceticus*, *P. wisconsinensis* e *Proteus vulgaris* e possuem massa molecular em torno de 30 a 32 kDa.

A subfamília I.2 é representada por proteínas com massa molecular um pouco maior, de 33 kDa, e inclui enzimas de micro-organismos dos gêneros *Burkholderia* e *Pseudomonas* e possuem, além da tríade catalítica, ligação a íons Ca⁺² intermediada por dois resíduos de aspartato conservados nestas enzimas e localizados próximos à tríade catalítica. Outra característica comum é a presença, adjacentes ao sítio ativo, de dois resíduos de cisteína que formam uma ponte dissulfeto e provavelmente, em conjunto com o sítio de ligação a íons cálcio, têm papel na estabilização do sítio ativo. A expressão de lipases ativas destas duas subfamílias depende da presença de uma proteína chaperona denominada foldase lipase-específica ("Lif").

A sub-família I.3 contém enzimas das espécies *Pseudomonas fluorescens* e *Serratia marcescens*. Estas lipases possuem massa molecular entre 50 a 65 kDa, e a ausência dos resíduos de cisteína e de um peptídeo sinal N-terminal. A sub-família I.4 agrupa lipases das espécies *Bacillus subtilis* e *Bacillus pumilus* e é representada por enzimas com massa molecular de 20 kDa e baixo percentual de similaridade, em torno de 15%, com as enzimas da sub-família I.5, que é representada por lipases do gênero *Staphylococcus* e de outras espécies de *Bacillus*, com massa molecular de 45 a 75 kDa e apresentam como característica comum a substituição da primeira glicina do pentapeptídeo conservado por alanina. Por último estão as lipases que possuem similaridade com as enzimas de *Propionibacterium acnes* e *Streptomyces cinnamoneus*, com 50% de similaridade entre elas, representando a sub-família I.6.

As enzimas agrupadas na família II (GDSL) não apresentam o pentapeptídeo G-X-S-X-G convencional, mas sim, um resíduo de aspartato e outro de leucina na adjacência do resíduo ativo de serina, que está muito mais próxima da extremidade N-terminal do que em outras enzimas lipolíticas. As hidrolases desta família apresentam propriedades multifuncionais e ampla especificidade de substrato, devido à flexibilidade do sítio ativo que sofre uma mudança conformacional sobre a ligação do substrato (OKAMURA et al., 2009).

A família III contém lipases extracelulares de *Streptomyces sp.* e *Moxarella sp.* e possuem o dobramento das α/β hidrolases com a tríade catalítica comum (Ser-Asp-His), enquanto a família IV engloba lipases que apresentam similaridade com lipases hormôniosensíveis de mamíferos e são adaptadas a temperaturas extremas. Membros da família V mostram semelhanças com várias enzimas bacterianas não lipolíticas e, assim como membros da família IV, são adaptadas a temperaturas extremas, além de possuírem os comuns dobramento e tríade catalítica. A família VI reúne as menores lipases, pois possuem de 23 a 26 kDa. A forma ativa dessas enzimas é um dímero com a clássica tríade catalítica e o dobramento das α/β hidrolases. São representadas por carboxilesterases e não exibem atividade em longas cadeias de triglicerídeos.

Os representantes da família VII, de 55 kDa, incluem esterases e enzimas com atividade hidrolítica sobre herbicidas com carbamato e ésteres de p-nitrobenzyl. Já na família VIII, ao contrário das outras famílias de esterases microbianas, a ordem dos resíduos catalíticos da sequência Ser-Asp-His, que é conservada na superfamília das lipases e esterases, não é a mesma nesta subfamília.

Apesar deste sistema de classificação ser muito utilizado, novas enzimas lipolíticas têm sido isoladas e identificadas de diversos ambientes pela abordagem metagenômica e novas famílias têm sido propostas (HENNE et al., 2000; KIM et al., 2009; LEE et al., 2006). Couto et al. (2010) isolaram um gene codificador de lipase, com 283 aminoácidos e massa molecular de 283 kDa, de biblioteca metagenômica de manguezal da costa brasileira. Análises filogenéticas mostraram que este gene forma um ramo único, constituindo uma nova família de lipase.

Sendo assim, com novos genes codificadores de enzimas lipolíticas sendo descobertos, existe a necessidade de acessar o recurso genético das espécies microbianas, e a metagenômica tornou-se uma abordagem que oferece uma combinação quase ilimitada para encontrar novos genes codificadores de produtos relevantes como lipases e esterases.

3 OBJETIVOS

Objetivo geral

Detectar clones produtores de enzimas lipolíticas em biblioteca metagenômica de DNA de um consórcio microbiano degradador de hidrocarbonetos de petróleo e caracterizar as sequências gênicas codificadoras de enzimas lipolíticas.

Objetivos específicos

- Realizar a triagem em meio de cultivo sólido com substrato para lipases nos clones da biblioteca metagenômica para detecção de enzimas lipolíticas;
- Realizar a sub-clonagem do(s) fragmento(s) inserido(s) no(s) clone(s) positivo(s) selecionado(s);
- Sequenciar a(s) sub-biblioteca(s) e analisar as sequências da(s) sub-biblioteca(s), obtendo a sequência completa do fragmento metagenômico do(s) clone(s) positivo;
- Identificar e selecionar as sequências codificadoras de enzimas lipolíticas;
- Construir árvores filogenéticas, analisar as sequências de aminoácidos e identificar a família lipolítica à qual o(s) gene(s) encontrado(s) pertence.

4 MATERIAL E MÉTODOS

4.1 Fluxograma metodológico

O presente trabalho foi desenvolvido no Laboratório de Bioquímica de Microorganismos e Plantas (LBMP), localizado na Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias (FCAV) de Jaboticabal. As etapas foram realizadas de acordo com o fluxograma metodológico demonstrado a seguir (Figura 3).



Figura 3- Fluxograma metodológico demonstrando as etapas realizadas no trabalho.

4.2 Obtenção do consórcio microbiano

O consórcio microbiano utilizado neste trabalho foi desenvolvido e cedido gentilmente pela pesquisadora Dr.^a Maria Benincasa Vidotti.

O solo utilizado para a obtenção do consórcio foi extraído de sítios de contaminação permanente em uma antiga fábrica de lubrificantes localizada em Ribeirão Preto – SP a 21 ° 06' S e 47 ° 49' O. O local onde as amostras de solo foram coletadas recebeu contaminação deste produto por aproximadamente 15 anos. A construção da biblioteca metagenômica de consórcio microbiano degradador de hidrocarbonetos de petróleo foi realizada como parte do projeto de Mestrado do aluno Douglas Antônio Alvoredo Paixão (PAIXÃO et al., 2010).

Utilizou-se para a construção da biblioteca metagenômica o CopyControlTM Fosmid Library Production Kit (EPICENTRE[®] Biotechnologies, Madison, Wisconsin, Estados Unidos da América), com fragmentos inseridos de 20 kb a 40 kb. O vetor utilizado na clonagem foi o pCC2FOSTM (EPICENTRE[®] Biotechnologies). Tal vetor possui 8.181 pares de bases e é linearizado no sítio de restrição único da enzima *SmaI* e desfosforilado. Foram utilizadas células competentes de *Escherichia coli* EPI300TM-T1^R (EPICENTRE[®] Biotechnologies). Dessa forma, foram obtidos 4224 clones, distribuídos em 44 placas de poliestireno (microplacas de 96 poços), que foram utilizadas para o desenvolvimento do presente trabalho.

4.3 Triagem dos clones com atividade para enzimas lipolíticas

Com o intuito de verificar a produção de enzimas lipolíticas pelos clones da biblioteca metagenômica do consórcio microbiano degradador de óleo hidrocarbonetos do petróleo, estes foram submetidos à análise em meio de cultura contendo tributirina como substrato. Foi utilizado como controle negativo células de EPI300TM-T1^R *Escherichia coli* não transformadas do Copy ControlTM Fosmid Library Production Kit (Epicentre Biotechnologies).

Primeiramente, para que ocorresse a recuperação das células viáveis, os clones foram retirados do estoque a -80 °C e inoculados em placas de Petri (90 mm X 15 mm) com o meio de cultura Luria-Bertani (LB) [1% de triptona (p/v), 1% de NaCl (p/v), 0,5% de extrato de

levedura (p/v), 1,5% de ágar (p/v), (pH 7,0)], que foi esterilizado a 121 °C por 20 minutos, 1 atm, e então, adicionou-se 0,00125% de cloranfenicol (v/v), o antibiótico adequado para o vetor. Os clones foram incubados a 37 °C em estufa tipo BOD (Demanda Biológica de Oxigênio) durante a noite. Posteriormente, as células recuperadas foram usadas para inoculação em meio de cultura LB suplementado 1% de tributirina (v/v), 1% de goma arábica (p/v), 0,001% de arabinose (v/v) e, em alguns casos, para facilitar a visualização, com 0,001% de Rhodamine-B (v/v).

Para a adição da tributirina ao meio, esta foi emulsificada, de acordo com Lee et al. (2004), em aparelho de ultrassom modelo Branson Sonifier 250 (Branson, Connecticut, EUA), com 30% da razão cíclica, por 3 ciclos de 15 segundos. O meio de cultura foi esterilizado em autoclave conforme descrito anteriormente e adicionado nas placas de Petri de 150 mm X 15 mm. Após serem transferidos para este meio, os clones foram cultivados por três dias em BOD a 37 °C para verificar a formação de halos claros ao redor da colônia, caracterizando a hidrólise do triacilglicerídeo em glicerol e ácido butírico.

Todos os clones que tiveram formação de halos ao redor das colônias foram selecionados como potenciais para atividade lipolítica e inoculados em uma única placa de Petri com o mesmo meio de cultura, experimento desenvolvido pela aluna Mariana Rangel Pereira como parte de seu projeto de Mestrado. Após, foram selecionados os clones a serem estudados.

4.4 Extração do DNA fosmidial

Os clones que apresentaram halo ao redor das colônias, ou seja, eram capazes de hidrolisar a tributirina, tiveram seu DNA fosmidial extraído através do kit Wizard® Plus SV Minipreps DNA Purification System (Promega, Madison, Wisconsin, Estados Unidos da América), conforme indicações do fabricante.

4.5 Quantificação e análise do DNA

O DNA fosmidial extraído foi quantificado por espectrofotometria e a relação 260/280 nm foi utilizada para avaliar a qualidade do material nucléico e a presença de proteínas, respectivamente.

O resultado das extrações também foi visualizado em gel de agarose 0,8% (p/v), fundido em tampão TBE 1X [Tris 89 mM, Ácido Bórico 89 mM e EDTA (ácido etilenodiaminotetracético) 2,5 mM, pH 8,3], acrescido de brometo de etídio (0,5 mg/mL). As amostras foram preparadas para serem aplicadas no gel, assim, uma alíquota de 3 μ L foi misturada com 3 μ L de tampão de carregamento [0,025% de azul de bromofenol (p/v) e 50% de glicerol (p/v)]. A eletroforese foi realizada em uma cuba horizontal e conduzida em tampão TBE 1X a 75 V por 2h. Também foi aplicado no gel o padrão de tamanho molecular 1 kb DNA Ladder (Fermentas, Burlington, Ontário, Canadá) e a imagem foi documentada sob luz UV em aparelho fotodocumentador L-PIX TOUCH (Loccus Biotecnologia, Cotia, São Paulo, Brasil).

4.6 Digestão do DNA fosmidial

Uma alíquota do DNA fosmidial dos clones selecionados foi digerida com duas enzimas de restrição, a fim de verificar se os clones são diferentes entre si e estimar o tamanho do fragmento inserido clonado no vetor fosmídeo.

As reações de digestão foram feitas separadamente para cada clone, e em cada reação, acrescentou-se a 1 µg do DNA, 10 U da enzima, 2 de µL tampão FastDigest[®] 10X (cat. n°; FD0504, Fermentas) e água deionizada ultra-pura para completar um volume de 20 µL. As reações foram feitas separadamente para cada enzima, FastDigest *Hind*III (Fermentas) e FastDigest *Sal*I (Fermentas). As reações foram incubadas a 37 °C por 10 minutos e em seguida, para inativação térmica das enzimas, a 80 °C por 10 minutos.

O volume total de cada digestão foi aplicado em gel de agarose 1% (p/v), sendo o preparo realizado conforme metodologia descrita no item 4.5, porém aplicou-se voltagem de
80 V por 4 horas e o marcador molecular Lambda *EcoRI* + *Hind*III (Promega) além do marcador de tamanho molecular 1 Kb DNA Ladder.

4.7 Construção das sub-bibliotecas

O DNA fosmidial extraído dos clones foi utilizado para construção de sub-bibliotecas *shotgun* em vetor pUC19, para possibilitar o acesso à informação genética por meio do sequenciamento do DNA.

4.7.1 Reação de nebulização

Com o intuito de gerar fragmentos com tamanho de 1 a 3 kb para serem sub-clonados, aproximadamente 18 μ g do DNA fosmidial dos clones foram inseridos em nebulizadores K7025-05 (Invitrogen, Carlsbad, Califórnia, EUA). Foram feitas reações independentes para cada clone positivo escolhido. Em cada nebulizador adicionou-se 50mM de Tris-HCl, pH 8,0, 15 mM de MgCl₂, ambas soluções filtradas em membrana de 0,22 μ m, 25% de glicerol a 50% (v/v) e água deionizada estéril para completar o volume de 2 mL.

O nebulizador foi conectado a um tubo de gás nitrogênio e mantido sob pressão de 3Kgf/cm² durante 16 segundos. Todo o processo de nebulização foi realizado em banho de gelo.

O nebulizador foi adaptado a um tubo de centrífuga para ser centrifugado a 162 xg, a 4 °C, por 2 minutos, a fim de dar início à precipitação do material. O volume recuperado foi transferido em alíquotas de 700 μ L para tubos de 1,5 mL.

Para cada alíquota do DNA nebulizado foram adicionados 0,1 volume de acetato de sódio 3 M, pH 5,2 e 1 volume de isopropanol absoluto gelado. As amostras foram homogeneizadas e incubadas por 22 h a -20 °C. Após a incubação, os tubos contendo DNA precipitado foram centrifugados a 15000 xg, 4 °C, por 25 minutos. Acrescentou-se 1 mL de etanol 80% (v/v) gelado para limpeza do precipitado e os tubos foram centrifugados novamente a 15000 xg, 4 °C, por 15 minutos. Os sedimentos obtidos foram secados em temperatura ambiente para que não restassem líquidos e, em seguida, dissolvidos em água deionizada estéril, incubados a 37 °C por uma hora e meia e, para cada clone, as alíquotas correspondentes foram unidas em tubos de 1,5 mL.

Uma alíquota do DNA nebulizado foi visualizado e quantificado por eletroforese em gel de agarose 0.8% (p/v) a 80V por 1 hora de 30 minutos, conforme descrito no item 4.5.

4.7.3 Reação de reparo das extremidades dos fragmentos

Após a obtenção dos fragmentos de DNA, estes foram submetidos ao reparo de suas extremidades. Para isso, utilizou-se em cada reação, aproximadamente 1,5 μ g de cada DNA, 10 μ L de Tampão NEBuffer2 1X (50 mM NaCl, 10 mM Tris-HCl, 10 mM MgCl₂ e 1 mM Dithiothreitol, pH 5,9) (New England BioLabs, Ipswich, Massachusetts, EUA), 1 μ L de BSA 100X (20 mM KPO₄, 50 mM NaCl, 0,1 mM EDTA e 5% de Glicerol, pH 7,0, New England BioLabs), 10 mM de dNTP (Fermentas), 3 U de T4 DNA Polimerase (New England BioLabs) e o volume foi completado para 100 μ L com água deionizada estéril. Uma unidade da enzima T4 DNA polimerase é definida como a quantidade de enzima necessária para incorporar 10 nmol de dNTP em material precipitado.

Em seguida, a reação foi incubada a 12 °C por 20 minutos. Depois, foram adicionadas 5 U de Klenow Fragment (New England BioLabs), sendo a reação incubada a 37 °C por 30 minutos, seguido de outra incubação para inativação térmica da enzima a 75 °C por 20 minutos.

As enzimas T4 DNA polimerase e Klenow DNA polimerase foram utilizadas para incorporação de nucleotídeos complementares livres, nos terminais 3' da dupla fita de DNA. A enzima T4 polimerase também apresenta atividade exonucleásica no sentido 3' a 5', que degrada extremidades que se sobressaem nos terminais 3' (SAMBROOK et al., 1989).

4.7.4 Reação de fosforilação dos fragmentos reparados

Para realizar a clonagem no vetor escolhido, realizou-se a fosforilação dos fragmentos inseridos. Aos fragmentos de DNA reparados (101,5 μ L) foram adicionados 12 μ L de tampão T4 DNA Ligase 10X, que contém a quantidade de ATP necessária para a reação, (50 mM Tris-HCl, 10 mM MgCl₂, 10 mM Dithiothreitol e 1 mM de ATP, pH 7,5, New England BioLabs), 10 U de T4 Polinucleotídeo Quinase (New England BioLabs) e água deionizada estéril para completar o volume da reação para 120 μ L. A enzima T4 polinucleotídeo quinase catalisa a transferência e troca de Pi a partir da posição γ do ATP para a extremidade 5 '- OH do DNA (SAMBROOK et al., 1989).

A reação foi incubada a 37 °C por 30 minutos e logo após a 65 °C por 20 minutos, para que ocorresse a inativação térmica da enzima.

4.7.5 Seleção do tamanho dos fragmentos inseridos

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

Foram aplicados, na seguinte ordem no gel de agarose: 10 μ L de cada amostra com 3 μ L de tampão de carregamento, 6 μ L do padrão molecular 1 kb DNA Ladder (Fermentas), o volume total das amostras (120 μ L) adicionado de 14 μ L de azul de bromofenol 6 X concentrado (3,3,5,5-tetrabromofenolsulfonftaleína) em três canaletas unidas. A eletroforese foi conduzida em tampão TEB 1 X isento de brometo de etídio, com voltagem inicial de 100

V por 30 minutos e posteriormente a 80 V durante 1 hora e 10 minutos. Através deste gel foi possível selecionar o tamanho do DNA entre 1 a 3 kb, para posterior recuperação.

Após a eletroforese, o gel de agarose foi dividido em duas porções por meio um corte vertical, resultando em uma porção com os insertos a serem recuperados e outra porção contendo o marcador e a amostra comparativa para serem corados. A porção do gel com o marcador foi corada com brometo de etídio (5 mg/mL) e observada sob luz UV com um transiluminador, com a finalidade de guiar a obtenção do fragmento inserido na porção não corada colocada ao lado, sendo esta cortada e retirada do restante do gel. Esta parte retirada foi dividida em tubos de 1,5 mL com 400 mg de gel por tubo e o restante do gel foi corado para que fosse possível observar se a banda foi selecionada corretamente.

A eluição da banda selecionada foi feita seguindo com o kit Wizard Plus SV Gel and PCR Clean-Up System (Promega) de acordo com instruções do fabricante e a quantificação realizada em aparelho de espectrofotometria.

4.7.6 Reação de ligação do fragmento inserido ao vetor

O DNA fosmidial extraído foi clonado no vetor pUC19 digerido com *Sma*I e desfosforilado. O vetor pUC19DNA/*Sma*I (Fermentas) possui 2.686 pares de bases e é ilustrado na figura 4.



Figura 4 - Esquema ilustrativo apresentando características do vetor pUC19DNA/SmaI (Fermentas).

A reação de ligação foi feita seguindo a razão: 300 ng de fragmentos inseridos para 100 ng de vetor. O cálculo da concentração fragmento inserido:vetor foi realizado de acordo com o Technical Manual pGEM-T and pGEM-T Easy Vector Systems (Promega), conforme demonstrado abaixo:

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

Sendo assim, a reação de ligação constituiu-se de aproximadamente 335 ng de cada DNA reparado e fosforilado, o que permitiria obter fragmentos de até 3 kb, 2 μ L do Tampão T4 DNA Ligase 10X com 10 mM de ATP (BioLabs), 1,5 μ L da enzima T4 DNA Ligase (BioLabs), 1 μ L do vetor pUC19DNA/*Sma*I a 100 ng/ μ L (Fermentas) e água deionizada estéril para completar o volume de 20 μ L. A reação foi incubada a 16 °C por 16 horas.

A enzima T4 DNA ligase catalisa a formação de uma ligação fosfodiéster entre terminais justapostos 5'-fosfato e 3'-hidroxila de DNA dupla fita. A enzima repara o DNA dupla fita, junta fragmentos com terminais coesivos ou sem corte, mas não tem nenhuma atividade em ácidos nucléicos de fita única. Esta enzima requer ATP como cofator.

4.7.7 Transformação bacteriana da célula competente

As transformações foram feitas com células competentes de *Escherichia coli*, DH5α, cedidas pelo Prof. Dr. Manoel Victor Franco Lemos, responsável pelo Laboratório de Genética de Bactérias, do Departamento de Biologia Aplicada à Agropecuária da FCAV, UNESP, Campus de Jaboticabal.

As células competentes foram previamente removidas do freezer a -80 °C e descongeladas em banho de gelo por aproximadamente 5 minutos antes de serem utilizadas. Na reação foram utilizados 6 μ L de cada amostra do DNA ligado, cuidadosamente depositados na parede de tubos de 15 mL, e 200 μ L das células competentes DH5 α . A transformação foi feita através de choque térmico e, para isso, a reação permaneceu por 20 minutos em banho de gelo e foi logo em seguida submetida a 42 °C, por exatamente 90 segundos, sendo então, recolocada no banho de gelo por mais 2 minutos.

Aos tubos foram adicionados 970 μ L de meio SOC [2% de triptona (p/v), 0,5% de extrato de levedura (p/v), 1 mL de NaCl 1 M, 0,25 mL de KCl 1 M, 1 mL de Mg²⁺ 2 M filtrado em membrana de poro 0,22 μ m e 1 mL de Glicose 2 M filtrada sob as mesmas condições] a fim de propiciar a recuperação das células competentes, que foram então, submetidas a agitação de 230 rpm a 37 °C, por uma hora e trinta minutos.

Após incubação das células transformadas, alíquotas de 90 µL de cada cultura foram distribuídas em placas de Petri contendo o meio LB sólido seletivo, com o auxílio de uma alça de Drigalski esterilizada. Para promover a seleção das células transformadas, ao meio LB

foram adicionados 70 µg/mL de ampicilina, de acordo com especificações do vetor pUC19, e então foram espalhados por sobre o meio com o auxílio da alça de Drigalski estéril, 100 µL de IPTG 0,1M [2,4% em água (p/v)] e 20 µL X-GAL [5-bromo-4-cloro-3-indolil- β -*D*-galactopiranosídeo, 100 mg dissolvidos em 2 mL de DMF (N,N'dimethyl-formamida)].

As placas de Petri com o meio LB e as células transformadas permaneceram em estufa BOD por 16 horas a 37 °C.

4.7.8 Coleta e estoque dos clones transformados

Após o período de incubação, as placas foram incubadas a 4 °C por 30 minutos. Este processo facilita a diferenciação da coloração das colônias azuis (não transformadas) e brancas (transformadas). As colônias brancas foram coletadas com o auxílio de palitos de madeira esterilizados por autoclavagem.

Os clones foram inoculados em placas estéreis de poliestireno (microplacas de 96 poços) contendo 120 μ L de meio LB líquido suplementado com ampicilina, em uma concentração de 70 μ g/mL. As placas foram incubadas a 37 °C durante 22 horas. Após esse período, foram adicionados 110 μ L de glicerol 40% (v/v) estéril em cada poço da placa. As placas foram então seladas com adesivos e estocadas a -80 °C.

4.8 Cultivo dos clones e extração de DNA plasmidial em placa

Os clones obtidos com a sub-clonagem foram cultivados em placas de cultivo de bactérias, de 96 poços, com 1,1 mL de meio LB suplementado com 70 µg/mL de ampicilina. Em seguida, foram submetidos a uma agitação contínua de 230 rpm a 37 °C durante 22 horas.

Para realizar a extração do DNA plasmidial, após a incubação, os clones foram centrifugados a 3220 xg, 20 °C, durante 6 minutos, sendo o sobrenadante descartado e a placa invertida em papéis absorventes durante 5 minutos. O material precipitado foi dissolvido em 140 μ L de solução GTE [Glicose 50 mM, Tris- HCl 23 mM (pH 8,0) e EDTA 10 mM] por

meio de agitação vigorosa por 2 minutos, ou até que todas as células ficassem totalmente dissolvidas de forma homogênea.

Repetiu-se a centrifugação anterior. Novamente os sobrenadantes foram descartados e a placa colocada invertida em papel para a secagem. Foram adicionados 50 μ L da solução GTE e o material foi agitado vigorosamente, sendo que a suspensão celular foi transferida para microplacas de 250 μ L, contendo 2,5 μ L de RNAse (10 mg/mL) e 60 μ L de solução de lise [NAOH 0,2 N e SDS 1% (p/v)] em cada poço. As placas foram seladas e invertidas 20 vezes. Após a inversão, foram adicionados 60 μ L de Acetato de Potássio 3 M, pH 5,2, em cada poço e as placas foram incubadas por 30 minutos a 90 °C. Então, o material foi resfriado em gelo por 7 minutos e centrifugado a 3220 xg, por 6 minutos, a 20 °C.

O sobrenadante obtido foi transferido para novas placas de 250 μ L, de polipropileno e foram adicionados 110 μ L de isopropanol absoluto, sendo a solução invertida por 20 vezes. As placas foram centrifugadas por 45 minutos a 3220 xg, a 20 °C, sendo o sobrenadante descartado. Foram acrescentados 200 μ L de etanol 70% gelado ao DNA precipitado para limpeza e as placas foram novamente centrifugadas a 3220 xg, por 5 minutos, a 20 °C. O sobrenadante foi descartado e a placa colocada invertida em papel toalha para absorver o excesso de líquidos.

Para secagem do precipitado formado, a placa permaneceu durante 1 hora em temperatura ambiente e em seguida cada amostra de DNA foi dissolvida em 30 µL de água deionizada estéril.

O DNA plasmidial extraído foi visualizado por eletroforese em gel de agarose 0,8%, conforme descrito no item 4.5.

4.9 Reação de sequenciamento

A reação de PCR para sequenciamento dos fragmentos de DNA plasmidial extraídos dos clones das sub-bibliotecas foram sequenciados em microplacas nas seguintes condições: 1 μ L de DYEnamic ET Terminator (GE Healthcare, Amersham, Buckinghamshire, Reino Unido), 10 pmoles do oligonucleotídeo iniciador M13, sendo que as reações foram feitas separadamente para *Forward* e *Reverse*, 3 μ L de tampão 2,5 X (400 mM Tris-HCl, pH 9, e 10

mM MgCl2), 100 a 120 ng de DNA e água deionizada esterilizada para completar um volume de 10 μ L.

As placas foram seladas com um adaptador de silicone e levadas ao termociclador, sendo submetidas ao seguinte ciclo: desnaturação a 94 °C por 5 minutos; 35 ciclos de 55 °C por 2 minutos, 72 °C por 3 minutos; ciclo final de extensão de 72 °C por 10 minutos e as amostras permaneceram a 4 °C até serem retiradas do termociclador.

Após a reação, as amostras foram preparadas para o sequenciamento do produto de PCR. Para a precipitação do DNA amplificado e marcado, foram adicionados 80 μ L de isopropanol 75% (v/v) em cada poço da placa, agitando vigorosamente durante alguns segundos. As placas foram incubadas em temperatura ambiente por 15 minutos e logo depois centrifugadas por 45 minutos a 1699 *xg* a uma temperatura de 20 °C. Descartou-se o sobrenadante e retirou-se o excesso de isopropanol por inversão da placa em papel absorvente. Foram adicionados 180 μ L de etanol a 70% e as placas foram centrifugadas por 5 minutos, na mesma temperatura e força centrífuga descritas anteriormente. Descartou-se o sobrenadante, sendo que o excesso de etanol foi retirado por inversão da placa novamente em papel absorvente. Para secar o líquido restante das placas, estas permaneceram por 2 minutos em aparelho com função de dessecador para soluções alcoólicas, a 45 °C.

As lavagens têm a finalidade de precipitar o DNA e retirar os nucleotídeos marcados por fluorescência que não foram incorporados.

Para aplicação no sequenciador, o DNA deve estar em fitas simples, então as amostras foram dissolvidas em 9 μ L de formamida e desnaturadas em a 95 °C por 5 minutos, e colocadas no gelo em seguida.

4.10 Análise das sequências

Após o sequenciamento, para verificar a qualidade das sequências geradas, foi utilizado o programa *Sequecing Analysis* 3.4 que gerou os eletroforogamas submetidos à análise pelo programa Phred/Phrap/Consed (GORDON et al., 1998) para a limpeza e remoção dos vetores e agrupamento das sequências *Forward* e *Reverse*. A seleção das sequências adequadas foi realizada com os programas do pacote Phred/Phrap, o qual analisa a qualidade das sequências visualizando graficamente e gerando arquivos no formato FASTA, (EWING et

al., 1998). Nos casos em que não há o agrupamento, sequências *Forward* e *Reverse* permanecem separadas gerando os *singlets*.

O programa ContGEN selecionou somente aquelas sequências que apresentaram mais de 350 bases com qualidade Phred maior ou igual a 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).

Baseadas em sua similaridade, as sequências são agrupadas em sequências consenso conhecidas como *contigs*. Então, os *contigs* gerados foram submetidos à consulta de similaridade de nucleotídeos, com sequências depositadas no banco de dados GenBank do *National Center for Biotechnology Information* (NCBI), acessado através da ferramenta de alinhamento local Blastn (ALTSCHUL et al., 1997), em um servidor SUNBLADE 1000 (Oracle, Redwood Shores, Califórnia, Estados Unidos da América).

Para a anotação gênica das sequências foi utilizado o programa Open Reading Frame Finder (ORF Finder) do NCBI (http://www.ncbi.nlm.nih.gov/projects/gorf) e as ORFs (Open Reading Frames) encontradas foram comparadas com sequências depositadas no banco de dados do próprio NCBI. A predição da função foi realizada através do Blastx, no banco de dados nr (proteínas não redundantes), no qual, sequências de aminoácidos no formato FASTA são comparadas com sequências homólogas de proteínas depositadas no banco de dados.

Para análise das relações filogenéticas, as sequências distintas de aminoácidos foram alinhadas usando o programa ClustalW, através do programa BioEdit Sequence Alignment Editor (versão 7.0). As árvores filogenéticas foram construídas com o auxílio do programa Mega 4.1 (TAMURA et al., 2007) usando o algoritmo do vizinho mais próximo *neighborjoining* (SAITO e NEI, 1987), com *bootstrap* de 500.

5 RESULTADOS E DISCUSSÃO

5.1 Obtenção do consórcio microbiano

O consórcio microbiano o qual foi fonte do DNA metagenômico utilizado para realização dos experimentos descritos neste trabalho, é resultado de uma parceria entre a Dra. Maria Benincasa Vidotti e a Profa. Dra. Eliana Gertrudes de Macedo Lemos.

5.2 Triagem dos clones com produção de enzimas lipolíticas

Os ensaios *in vitro* das 44 placas da biblioteca metagenômica de consórcio microbiano degradador de óleo hidrocarbonetos do petróleo foram realizados em placas de Petri contendo o meio LB com tributirina como substrato.

O desenvolvimento dos clones foi acompanhado diariamente com o intuito de verificar a formação de halos claros ao redor da colônia, até que os halos não pudessem mais ser dinstinguidos uns dos outros. Esta etapa fez parte do projeto de Mestrado da aluna Mariana Rangel Pereira, que observou a formação de halos em 30 clones de um total de 4224 (0,7%). Juntamente com os 30 clones, foi inoculado na mesma placa de Petri o controle negativo *Escherichia coli* EPI300TM-T1^R do Copy ControlTM HTP Fosmid Library Production Kit (EPICENTRE Biotechnologies) não transformada, ou seja, sem o fragemento inserido. Notase que não houve a formação de halo ao redor da colônia (figura 5), confirmando a viabilidade do ensaio, visto que já foram encontrados falsos positivos em trabalhos de seleção por atividade funcional de genes de interesse biotecnológico em biblioteca metagenômica (JONES; SUN; MARCHESI, 2007).



Figura 5- Placa de Petri com 30 clones previamente selecionados da biblioteca metagenômica de consórcio microbiano degradador de hidrocarbonetos de petróleo. As células foram inoculadas em meio LB, com antibiótico, suplementado com tributirina, goma arábica, arabinose, e Rhodamine-B por três dias a 37 °C. CN- Controle negativo.

É notável o número alto de clones que apresentaram atividade lipolítica, contrastando com os dados de Wu et al. (2004), que encontraram apenas oito clones de biblioteca metagenômica de solo de floresta com atividade lipolítica, de um total de 33.700 (0,02%). A expressão em *Escherichia coli*, é dificultada por alguns fatores, como: dificuldade de expressão dos genes lipolíticos em hospedeiro heterólogo em função do não reconhecimento em sistema regulador, ao requerimento de dobramento de proteínas e transportadores e a toxicidade da proteína lipolítica expressada (RANJAN et al., 2005).

Foram contabilizados até mesmo os clones que apresentaram halos com menor destaque, sendo esta uma possibilidade para a contagem de maior número de clones com atividade. Além disso, fatores como o DNA dos micro-organismos ter sido extraído de solo contaminado com lubrificantes, a correta representatividade da amostra, eficiência da ligação e transformação e o método utilizado para selecionar os clones positivos também podem ter contribuído para este resultado. Os halos formaram-se no terceiro dia de cultivo das placas a 37 °C, mas ficaram ainda mais visíveis ao serem incubados a 4 °C. O substrato, tributirina, por constituir-se de triacilglicerol, ao resfriar-se, torna-se mais opaca, facilitando o contraste com os halos claros.

O vetor utilizado, pCC2FOS, possui marcador de resistência a cloranfenicol, sítio *cos* para empacotamento no fago lambda, duas origens de replicação, *ori*2, para manutenção do

fosmídeo numa única cópia em *E. coli*, e *ori*V que, segundo Wild et al. (2002), promove alto número de cópias do fosmídeo (10-200) pelo sistema *ori*V/*trf*A, após indução com arabinose. A goma arábica é um emulsificante amplamente utilizado para cadeias longas de triacilgliceróis (TISS; CARRIÈRE; VERGER, 2001) e a Rhodamina-B facilita a visualização da formação dos halos.

Com a análise dos resultados obtidos nos ensaios, foram selecionados, devido ao tamanho dos halos, sete clones para que fossem confirmados os resultados obtidos anteriormente e selecionar os clones para construção de sub-bibliotecas. Foi possível verificar que todos os clones selecionados apresentaram a formação de halos claros ao redor das colônias (figura 6). Após a medida dos halos, foram selecionados os clones PL14.F10 e PL28.H10 para o prosseguimento deste trabalho. Dos outros 28 clones, dois foram estudados no trabalho da aluna Mariana Rangel Pereira, o que vem a contribuir com a busca de novas enzimas lipolíticas.



Figura 6- Cultivo dos sete clones selecionados com atividade lipolítica em meio LB suplementado com tributirina, goma arábica e arabinose. As células foram incubadas por três dias a 37 °C. A- Clone PL14.H10; B- Clone PL28.F10.

5.3 Extração do DNA fosmidial dos clones selecionados

O DNA fosmidial extraído dos clones PL14.H10 e PL28 pode ser visualizado na Figura 7. Pode-se verificar, assim, que o DNA não estava fragmentado e não continha impurezas. Para o clone PL14H.10, a concentração foi de 152,7 ng/µL com relação 260/280 nm de 1,81. Já o clone PL28.F10 apresentou concentração de 131 ng/µL, com relação 260/280 nm de 1,79, o que demonstra concentração suficiente para a realização das subclonagens e ausência de contaminação por proteínas, respectivamente.



Figura 7- Perfil eletroforético em gel de agarose 0,8% da extração do DNA fosmidial dos clones selecionados. 1- Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2- DNA extraído do clone PL14.H10; 3- DNA extraído do clone PL28.F10.

5.4 Digestão do DNA fosmidial dos clones selecionados

A digestão do DNA fosmidial dos clones PL14.H10 e PL28.F10 com as enzimas *Hind*III e *Sal*I permitiu afirmar que os clones são diferentes entre si, visto que apresentaram padrões diferentes de restrição. Esta análise possibilitou calcular o tamanho aproximado do fragmento inserido clonado em cada vetor, sendo aproximadamente 36 kb para o clone PL14.H10 e 32 kb para o clone PL28.F10 (Figura 8).



Figura 8- Perfil eletroforético em gel de agarose 0,8% da restrição do DNA fosmidial dos clones PL14.H10 e PL28.F10 com as enzimas *Hind*III e *Sal*I (Fermentas). 1, 5 e 9- Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2, 6 e 10- Marcador de DNA Lambda *EcoRI/Hind*III (Fermentas); 3 e 4- DNA fosmidial não digerido dos clones PL14.H10 e PL28.F10, respectivamente; 7 e 8- DNA dos clones PL14.H10 e PL28.F10, respectivamente, digeridos com *Hind*III; 11 e 12- DNA dos clones PL14.H10 e PL28.F10, respectivamente, digeridos com *Sal*I.

5.5 Construção das sub-bibiotecas

As sub-bibliotecas dos clones PL14.H10 e PL28.F10 foram construídas com fragmentos variando seu tamanho de 1 a 3 kb.

Por meio da análise em eletroforese em gel de agarose de uma alíquota de cada material nebulizado, foi possível verificar que a nebulização de ambos foi bem sucedida, já que os fragmentos gerados estavam mais concentrados na região de interesse (figura 9). A concentração do material nebulizado foi estimada pelo gel de agarose sendo que para o clone PL14.H10 foi de 25 ng/µL e para o clone PL28.F10 foi de 15 ng/µL, aproximadamente, de acordo com a relação ng/µL do marcador de tamanho molecular utilizado 1 kb DNA Ladder (Fermentas).



Figura 9- Perfil eletroforético da nebulização dos clones PL14.H10 e PL28.F10 em gel de agarose 0,8%. 1 e 3- Marcador de tamanho molecular 1 kb DNA Ladder (Fermentas); 2- DNA nebulizado do clone PL14.H10; 4- DNA nebulizado do clone PL28.F10.

Após seleção da região de interesse no gel de agarose de baixo ponto de fusão, os fragmentos de 1 a 3 kb, já reparados e fosforilados, puderam ser selecionados. Após a eluição foi realizada a leitura da concentração por espectrofotomeria, resultando em 59,8 ng/µL e 24,4 ng/µL para os clones PL14.H10 e PL28.F10, respectivamente.

Os fragmentos de DNA fosmidial dos clones recuperados da eluição foram inseridos em vetor pUC19 desfosforilado e digerido com *Sma*I. Posteriormente, os plasmídeos recombinantes inseridos por transformação nas células competentes de *Escherichia coli* DH5α. Para a sub-biblioteca relativa ao clone PL14.H10, foram coletados 464 sub-clones em cinco placas de poliestireno (microplacas de 96 poços). Para a sub-biblioteca relativa ao clone PL28.F10, foi possível coletar 672 sub-clones, em sete placas de poliestireno.

A extração do DNA plasmidial dos sub-clones pode ser visualizada nas figuras 10 e 11. Conforme observado, a extração resultou em qualidade e concentração suficientes para o sequenciamento.



Figura 10 - Perfil eletroforético em gel de agarose 0,8% de DNA plasmidial extraído de 48 sub-clones originados da sub-biblioteca do clone PL14.H10.



Figura 11 - Perfil eletroforético em gel de agarose 0,8% de DNA plasmidial extraído de 48 subclones originados da sub-biblioteca do clone PL28.F10.

5.6 Análise do clone PL28.F10

5.6.1 Análise da sequência do contig

Após o sequenciamento dos sub-clones referentes ao clone PL28.F10, os eletroferogramas foram analisados com o auxílio de programas do pacote Phred/Phrap/Consed e, utilizando o programa contGEN foi possível construir um *contig* de 36.942 pb.

O formato fasta do *contig* completo foi comparado com as sequências do NCBI, utilizando a ferramenta Blastn (*nucleotide collection* nr/nt), obtendo como resultado 80% de identidade com o micro-organismo *Pseudomonas aeruginosa* (número de acesso CP000438.1), 35% de *Query-coverage* e *E-value* de 0.0, seguido de uma identidade de 84% para o mesmo micro-organismo, com *Query-coverage* de 35% e *E-value* de 0.0, conforme demonstrado na tabela 1. Por possuir apenas pequenas regiões com similaridade (ANEXO A), não é possível afirmar se a espécie é realmente a que foi encontrada, indicando somente o gênero e adicionalmente, pode-se concluir que o *contig* obtido do clone PL28.F10 ainda não foi estudado.

		1	1	(continua)	
N° de acesso		icro-organismo Identidade Query- (%) coverage (%)			
NCBI	Micro-organismo			E-value	
	Pseudomonas			0.0	
CP000438.1	aeruginosa	80	35		
CP000744.1	Pseudomonas	0.4	25	0.0	
	aeruginosa	84	35	0.0	
FM209186.1	Pseudomonas	92	35	0.0	
	aeruginosa				

 Tabela 1 – Resultado do sequenciamento completo do contig formado do clone PL28.F10, em comparação com sequências do NCBI, através da ferramenta Blastn.

N° de acesso NCBI	Micro-organismo [%]		Query- coverage (%)	E-value	
AE004091.2	Pseudomonas aeruginosa	86	35	0.0	
CT573326.1	Pseudomonas entomophila	84	19	0.0	

 Tabela 1 – Resultado do sequenciamento completo do contig formado do clone PL28.F10, em comparação com sequências do NCBI, através da ferramenta Blastn.

 (continuação)

A alta identidade com micro-organismos do gênero *Pseudomonas* corrobora com os resultados obtidos por Paixão et al. (2010), que analisaram a frequência dos micro-organismos na biblioteca de consórcio microbiano degradador de hidrocarbonetos do petróleo, a mesma estudada neste trabalho. De acordo com estes autores, o grupo *Pseudomonas* foi o gênero com maior frequência na biblioteca metagenômica, com 32,2% de clones totais. Trabalhos como o de Higashioka et al. (2009) descrevem a elevada frequência deste gênero em solos contaminados, assim como a sua eficiente capacidade na degradação de hidrocarbonetos (OBUEKWE et al., 2008). Além disso, um grande número espécies de *Pseudomonas* é capaz de utilizar hidrocarbonetos do petróleo como fonte de carbono (WHYTE et al., 1997). Wunsche et al. (1995), em um estudo realizado com solo exposto durante anos à contaminação de hidrocarbonetos, observaram a dominância do gênero *Pseudomonas* entre os micro-organismos presentes naquele ambiente, mais especificamente *P. boxihidrogena*.

Tal resultado pode estar relacionado com a maior facilidade de adaptação destes micro-organismos em ambientes extremos, sendo menos sensíveis aos derivados de petróleo. A tolerância e predominância destes gêneros bacterianos em ambientes contaminados com compostos de hidrocarbonetos podem estar associadas a vias metabólicas específicas que envolvem diferentes genes e enzimas as quais transformam complexas moléculas contidas no petróleo em intermediários comuns das suas rotas catabólicas. Vários são os genes envolvidos na via metabólica de degradação dos hidrocarbonetos e entre estes muitos já foram caracterizados para um grupo extenso e diversificado de micro-organismos, os quais são capazes de decompor uma imensa gama de compostos que formam o petróleo, incluindo os alcanos (WHYTE et al., 2002), tolueno (PARALES et al., 2000) e hidrocarbonetos poliaromáticos (WHYTE et al., 1997).

Após análise do resultado do sequenciamento do *contig* completo, foi feita a predição das ORFs com o programa ORF Finder (Open Reading Frame Finder) do NCBI. O programa gerou 227 ORFs e todas foram analisadas. Mapeados os genes, a etapa seguinte consiste em identificar quais proteínas são codificadas, e nisso consiste o processo de anotação das sequências protéicas. Nessa etapa, procura-se montar um catálogo dos genes presentes no micro-organismo estudado, dando-lhes nomes e associando-os a prováveis funções. As 227 ORFs foram analisadas com relação à presença ou ausência de domínios conservados, se há ou não similaridade com micro-organismos depositados no banco de dados, *Query-coverage* e *E-value*. Por meio dessa análise, foi possível 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 predizer qual o papel biológico de cada uma.

Após a análise, foi possível selecionar 32 ORFs que estão relacionadas na tabela 2. A enzima lipolítica está representada pela ORF de número 15. Todas as ORFS encontradas foram representadas através de um mapa físico ilustrado na Figura 12.

						(00)	ininua)
ORFs	Início e término da ORF (nucleotídeo)	ORF (pb)	ORF (aa)	Descrição e número de acesso	Organismo	Identidade (%)	E- value
				Putative	Pseudomonas		
1	22894-25119	2226	741	oxidoreductase	aeruginosa	85	0.0
				(YP_002439103.1)	LESB58		
				Hypothetical	Parvibaculum		
2	10690-12729	2040	679	protein	lavamentivorans	57	0.0
				(YP001411377.1)	DS-1		
3	18470-20287	1818	605	Dihydroxy-acid dehydratase (YP_614480.1)	Ruegeria sp. TM1040	79	0.0
				Putative			
		1390	463	decarboxylase	Pseudomonas		0.0
4	33890-35281			family protein	aeruginosa	86	
				(YP_002438464.1)			
				Hypothetical	Pelohactor		
5	20345-21643	3 1299	432	protein Pcar_2748	carbinolicus	47	1e-20
				(YP_358153.1)	caromonicus		
				Argininosuccinate	Pseudomonas		
6	31828-33045	1218	405	synthase	aeruginosa	95	0.0
				(YP_001347006.1)	PA7		
				Hypothetical	Pseudomonas		
		1164	207	protein	aeruginosa		
7	25642-26805	1164	387	PSPA7_1615	DA7	62	3e-102
				(YP_001346999.1)	1 / /		

 Tabela 2 – Predição das 32 ORFs identificadas no contig do clone PL28.F10 através do ORF Finder (NCBI).

 (ncBI).

						(con	tinuaçao)
8	15595-16719	1125	374	Extracellular solute-binding protein family 1 (YP_003755558.1)	<i>Hyphomicrobium</i> <i>denitrificans</i> ATCC 51888	57	8e-116
9	29524-30570	1047	348	Dihydroorotase (YP_001347004.1)	Pseudomonas aeruginosa PA7	92	0.0
10	17027-18058	1032	343	integral membrane sensor signal transduction histidine kinase (YP_001414513.1)	Parvibaculum lavamentivorans DS-1	45	1e-80
11	1531-2547	1017	338	short-chain dehydrogenase/ reductase SDR (YP_001411651.1)	Parvibaculum lavamentivorans DS-1	80	9e-105
12	14602-15540	938	312	Binding-protein dependent transport systems inner membrane component (YP_774935.1)	Burkholderia ambifaria AMMD	60	2e-98
13	13775-14704	930	309	Binding-protein dependent transport system inner membrane protein (YP_426103.1)	Rhodospirillum rubrum ATCC 11170	63	6e-92

 Tabela 2 – Predição das 32 ORFs identificadas no contig do clone PL28.F10 através do ORF Finder (NCBI).
 (continuação)

				n		(COII	(iniuação)
14	21793-22710	918	305	ornithine carbamoyl- transferase (YP_001346993.1)	Pseudomonas aeruginosa PA7	86	3e-155
15	2415-3326	912	303	Lipase/esterase (AAX37296.1)	uncultured bacterium	61	1e-95
16	12915-13799	885	294	Metal-dependent hydrolase-like protein (YP_001411363.1)	Parvibaculum lavamentivorans DS-1	60	4e-90
17	6539-7318	780	259	Two component transcriptional regulator, winged helix family (YP_001411360.1)	Parvibaculum lavamentivorans DS-1	72	4e-99
18	707-1477	771	256	3-oxoacyl-[ACP] reductase (YP_002129961.1)	Phenylobacterium zucineum HLK1	75	2e-102
19	4198-4944	747	248	Short chain dehydrogenase/ reductase family oxidoreductase (YP_761048.1)	Hyphomonas neptunium ATCC 15444	77	2e-104
20	28847-29527	681	226	Ribonuclease T (YP_001347003.1)	Pseudomonas aeruginosa PA7	88	4e-115
21	32921-33592	672	223	Lactoyl- glutathione lyase (YP_001186892.1)	Pseudomonas mendocina ymp	88	3e-61

 Tabela 2 – Predição das 32 ORFs identificadas no contig do clone PL28.F10 através do ORF Finder (NCBI).

 (ncBI).

						(co	ntinuação)
22	35877-36515	639	212	Endonuclease III (NP_252185.1)	Pseudomonas aeruginosa PAO1	94	4e-114
23	3584- 4204	621	206	DSBAParvibaculumoxidoreductaselavamentivorans(YP_001414905.1)DS-1		69	5e-74
24	28032-28634	603	200	Peroxidase (NP_252219.1)	Pseudomonas aeruginosa PAO1	94	4e-106
25	31195-31695	501	166	Putative outer membrane protein precursor (ZP_06877446.1)	Pseudomonas aeruginosa PAb1	78	1e-53
26	26854-27330	477	158	Bacterioferritin (NP_252221.1)	Pseudomonas aeruginosa PAO1	99	2e-83
27	8845-9318	474	157	Hypothetical protein Plav_1521 (YP_001412797.1)	Parvibaculum lavamentivorans DS-1	71	7e-54
28	5138-5602	465	154	Hypothetical protein Arad_0425 (YP_002543046.1)	Agrobacterium radiobacter K84	59	8e-50
29	25252-25578	327	108	Hypothetical protein PA3533 NP_252223.1	Pseudomonas aeruginosa PAO1	93	3e-53
30	36512-36823	312	103	Hypothetical protein PaerPA_01004093 (ZP_01366942.1)	Pseudomonas aeruginosa PACS2	77	2e-33

Tabela 2 – Predição das 32 ORFs identificadas no *contig* do clone PL28.F10 através do ORF Finder (NCBI)

						(00	onclusão)
31		216	71	PA3530			
				(synthetic	Pseudomonas	84	
	27551-27766			construct)	aeruginosa		3e-22
				(AAT50652.1)			
				Thioesterase	Parvibaculum		
32	8641-8841	201	66	superfamily protein	lavamentivorans	49	1e-15
				(YP_001412597.1)	DS-1		
	1						

 Tabela 2 – Predição das 32 ORFs identificadas no contig do clone PL28.F10 através do ORF Finder (NCBI).



Figura 12 - Mapa físico gerado pelo programa Pages'09, versão 4.04 (Apple Inc., Cupertino, Califórnia, Estados Unidos da América) do clone PL28.F10. Foram encontradas 32 ORFs no *contig*, além de três ORFs sem domínio conservado (sem numeração). A ORF representada pelo número 3 (em vermelho) é a ORF15 da tabela, codificadora de lipase/esterase.

O resultado do Blastp aponta uma ORF, representada pelo número 15, codificadora de lipase/esterase e uma ORF, de número 18, codificadora de 3-oxoacyl-[ACP] redutase, uma enzima que participa do segundo passo da biossíntese dos ácidos graxos, sugerindo que tais ORFs podem estar em um *cluste*r gênico, envolvido na biossíntese de lipídeos.

Para analisar com mais acurácia a ORF 15, a sequência em aminoácidos obtida pelo ORF Finder foi submetida ao BLASTp (*non-redundant protein sequence* – nr). Ela possui 303 aminoácidos e apresentou 61% de identidade com micro-organismo não cultivável (Tabela 3). O resultado encontrado está melhor ilustrado no ANEXO B.

ORF	Descrição e número de acesso	Organismo	Identidade (%)	E-value
15	lipase/esterase (AAX37296.1)	uncultured bacterium	61	1e-95
	Est10 (ADA70030.1)	Uncultured marine bacterium	62	2e-93
	lipase/esterase (AAS77233.1)	uncultured bacterium	61	3e-91
	lipase/esterase (AAX37295.1)	uncultured bacterium	62	5e-91

 Tabela 3– Análise da ORF 15 utilizando Blastp (non-redundant protein sequence – nr) codificadora de esterase/lipase, encontrada no *contig* da sub-bilioteca PL28.F10. São mostrados os quatro primeiros resultados fornecidos pelo NCBI.

O sequenciamento de grandes fragmentos de DNA metagenômico, obtidos aleatoriamente, ou de fragmentos específicos, como o gene *16S rRNA*, revelou numerosas ORFs, muitas delas codificadoras de enzimas. Como resultado da busca por enzimas específicas foram encontradas enzimas como quitinase, 4-hidroxi-butirato desidrogenase, lipase/esterase, proteases, oxigenases, amilase, DNAse, xilanase e policetídeos sintases (HENNE et al., 2000; RONDON et al., 2000). Porém, sabe-se que o acesso a novas enzimas e biocatalisadores biológicos têm sido limitado pelo número relativamente pequeno de micro-organismos cultiváveis (LORENZ e ECK, 2005).

Diversas moléculas de interesse biotecnológico já foram obtidas por culturas puras de micro-organismos. Porém, técnicas tradicionais recuperam apenas uma fração, geralmente menos de 1% da diversidade presente em amostras ambientais complexas, como o solo (VOGET et al., 2003). Já a diversidade molecular natural encontrada pela abordagem metagenômica de micro-organismos não cultiváveis é tão grande que a probabilidade de recuperar genes conhecidos é muito pequena (LORENZ et al., 2002). Sendo assim, no caso deste trabalho, a probabilidade da ORF15 ser codificadora de uma nova enzima lipolítica é alta, devido ao fato de o micro-organismo encontrado não ser cultivável, de acordo com os resultados obtidos pelo NCBI.

Segundo Arpigny e Jaeger (1999) a classificação deste grande conjunto de dados resultante do sequenciamento e a identificação das famílias e sub-famílias de enzimas lipolíticas têm sido feita a fim de identificar as sequências de motivos conservados em

enzimas lipolíticas provenientes de uma ampla variedade de organismos, e relacioná-las a elementos estruturais de terceira dimensão envolvidos no reconhecimento do substrato e catálise, sendo assim, essencial para a compreensão da função das enzimas.

A sequência da ORF15 também foi analisada com as sequências depositadas no banco de dados do NCBI de amostras ambientais [environmental samples (envi_nr)], através do Blastp. O resultado foi de 48% de identidade com uma proteína hipotética de metagenoma marinho (número de acesso ECL38806.1), no entanto, foi verificado o domínio conservado da Superfamília de esterases/lipases.

5.6.2 Análise das relações filogenéticas

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

As relações filogenéticas foram construídas com o auxílio do programa Mega 4, usando as sequências que representam as famílias lipolíticas. Foram feitos diferentes alinhamentos e construções filogenéticas, baseando-se nas sequências extraídas de trabalhos de Arpigny e Jaeger (1999), Wu e Sun (2009) e 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 sub-divisões. Estes autores utilizaram 53 sequências de enzimas lipolíticas, e estas foram comparadas com a sequência da ORF15 obtida neste trabalho para estabelecer as relações filogenéticas entre elas, utilizando o algoritmo do vizinho mais próximo, *neighbor-joining*.

Por meio da análise da árvore (Figura 13), é possível observar que a ORF15 está mais próxima da Família IV das esterases/lipases, em um ramo único, diferenciando-se dos outros representantes desta família, *Cupriavidus necator* (L36817) e *Pseudomonas sp* (AF034088).



Figura 13 - Dendrograma do agrupamento hierárquico baseado nas sequências de membros das famílias lipolíticas, mostrando a relação filogenética da ORF15 de esterase/lipase, encontrada no *contig* do clone PL28.F10 em relação às sequências utilizadas por Arpigny e Jaeger (1999).

Já no trabalho de Wu e Su (2009), as oito famílias também foram propostas, incluindo as enzimas lipolíticas não definidas, que podem vir a constituir nova(s) família(s). A construção da relação filogenética foi feita utilizando as mesmas sequências de enzimas lipolíticas que Wu e Sun relataram no trabalho e a sequência da ORF15 da sub-biblioteca do clone PL28.F10. A metodologia utilizada foi a mesma relatada anteriormente.

Mais uma vez, a ORF15 agrupou-se com os membros da Família IV das esterases/lipases, em um ramo único, diferenciando-se dos outros representantes dessa família, *Cupriavidus necator* (L36817) e *Pseudomonas sp* (AF034088) (Figura 14).



Figura 14 - Dendrograma do agrupamento hierárquico baseado nas sequências de membros das famílias lipolíticas, mostrando a relação filogenética da ORF15 de esterase/lipase, encontrada no *contig* do clone PL28.F10 em relação às sequências utilizadas por Wu e Sun (2009). Outra árvore filogenética foi construída, com a mesma metodologia, porém as sequências utilizadas foram baseadas no trabalho de Couto et al. (2010). Estes autores utilizaram as mesmas sequências, mas acrescentaram sequências "lip G", uma nova família de esterases/lipases sugerida por Lee et al. (2006). A sequência da ORF15 foi relacionada com as sequências para a construção de árvore filogenética (Figura 15). Mais uma vez esta ORF se relacionou com a espécie *Moraxella sp.* (X53868), mostrando-se próxima a este micro-organismo pertencente à família IV, mas se dispôs em um ramo único, podendo se constituir de uma nova sub-família dentro da família IV.

Dessa forma, existe mais uma vez a possibilidade de a ORF15 ser identificada como uma nova sequência gênica codificadora de enzimas lipolíticas.



Figura 15 - Dendrograma do agrupamento hierárquico baseado nas sequências de membros das famílias lipolíticas, mostrando a relação filogenética da ORF15 de esterase/lipase, encontrada no *contig* do clone PL28.F10 em relação às sequências utilizadas por Couto et al. (2010).

5.6.3 Identificação dos sítios catalíticos

Para obter resultados que confirmem a localização da ORF15 dentro da família IV, foi realizado o alinhamento da ORF em questão com representantes desta família (Figura 16).

Pode-se afirmar que o alinhamento corrobora com os resultados obtidos, já que é possível observar sequências do sítio catalítico correspondentes à família IV.

Os membros desta família 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 Serina como resíduo dentro do pentapeptídeo GDSAG, além de apresentarem um motivo estritamente conservado HGGG, localizado *upstream* ao sítio ativo (JAEGER; DIJKSTRA; REETZ, 1999), cuja função está relacionada à formação das pontes de hidrogênio para estabilidade durante a catálise (WEI et al., 1999).

Com os alinhamentos no ClustalW realizados com a sequência da ORF15 e de alguns micro-organismos da família IV foi possível identificar as sequências de aminoácidos <Ser-Asp-His> para esta ORF, que corresponde à tríade catalítica conservada, presente nas enzimas lipolíticas da superfamília alfa/beta hidrolase.

A sequência de aminoácidos HGGG e os três sítios catalíticos (GDSAG; DPLXD e HGG) foram identificados na ORF15. 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 ORF15, sugere-se que esta seja uma ORF representante da família IV. No entanto, foram verificadas algumas modificações nos resíduos de aminoácidos. O resíduo de aminoácido fenilalanina (F) do motivo conservado HGGGF foi substituído por tirosina (Y), no sítio *upstream* ao sítio catalítico da serina, sendo a sequência deste bloco, HGGGY. Já o motivo GDSAG permaneceu inalterado. Outra modificação foi vista no segundo sítio catalítico, do aspartato (DPLXD), no qual o resíduo de aminoácido aspartato (D) está trocado por glutamato (E), e o resíduo de aminoácido prolina foi trocado por treonina (T), formando o bloco ETLLD.

	10	20	30	40	50	60	70
Moraxella sp X53868	MPTIPVPAINALLY	RTT TRTT RTCA	ARNAHOHHVI H	HTIRGI DNI	PAPVI ERINER	RASTAROVI	
Cupriavidus necator L36817		MSR	CGRRAELLHAA	AGTGSGSASE	PMSAAPSSSPE	APATPGLPSS	PLDPOVAAL
Pseudomonas sp AF034088						1	PLDKOIAAV
ORF 15							-MSDPEIGAI
	120	130	140	150	160	170	180
Moraxella sp X53868	OONADASGSTENAL	SWODKTIANA	DGGDMTVRCYC	KSTONSERK	STDEAAMLETE	GGGFCIGDI	THHEFCHTV
Cupriavidus necator L36817	INPPPV	MAEDLLAPAR	DGHAIPLRLY	PREASW	TEPLPLLVYF	IGGGETVGSVI	SHDPLCRLL
Pseudomonas_sp_AF034088	GDPMI	EVRNLRVAAA	AG-ELDARLYF	RPLEE	-DNLPLLVFFF	IGGGFVMGNLI	THDNLCRSL
ORF 15		VTRERVSANG	VSA <mark>E</mark> WSSTP <mark>G</mark> A	AARDR	VVMYVE	IGGGYVIGSVA	SHRHLVTEL
	230	240	250	260	270	280	290
Moraxella_sp_X53868	EHSQSLGASPSRIV	LSGDSAGOCL	AALVAQQVIK	DALWODNN	QAPAA <mark>DKK</mark> VN <mark>I</mark>	TFRNSLADLI	RPLAQLPLY
Cupriavidus_necator_L36817	AEAGRLGADPARIA	VGGD SAGGTL	AAACA-VEAR	NAG		1	APVLQLLIY
Pseudomonas_sp_AF034088	EHAAELGVDGRRL/	LAGDSAGGNL	ALAVSRLAAQF	RQG		I	RISYQCLFY
ORF 15	AQGVEPGNIA	IAGDSAG9GL	TVATL-LAIR)AG		L	QPACGICIS
	340	350	360	370	380	390	400
Moraxella_sp_X53868	SGLPQSHPLISVM	GDNTQL	CPSYIVVAEL	DILRDEGLAY	AELLQREGVQV	QTYTVLGAP	IGFINLMSVE
Cupriavidus_necator_L36817	EASRDDWRFAPLDO	GGAGAEVRGT	CPAWIAVAGY	PLHDEGVAY	AEKLRAAGVAA	AT LADYPGMI	IDFFKLGRFV
Pseudomonas_sp_AF034088	TGQGDDPLASPLR-	AETLADL	PPTTLITAEFE	PLRDEGEAF.	ALRLQQAGVSV	RVQRCEGMI	IGFI SMAPFV
ORF 15	GANPRTPLAAPLY-	AEFRGI	APLLIQVGAA	TLLDD STRL	ASRAGAADVDV	TLQIWPEMI	IVWHFFHPML

Figura 16- Alinhamento da sequência da ORF15 com as sequências de aminoácidos de representantes da famíla IV, *Moraxella sp.* (número de acesso X53868), *Pseudomonas sp.* (número de acesso AF034088) e *Cupriavidus necator* (número de acesso L36817). As caixas representam as regiões conservadas de enzimas lipolíticas, e o símbolo (•) representa possíveis resíduos de aminoácidos envolvidos na tríade catalítica.

Com o intuito de comparar as ORFs e confirmar que a ORF15 deste trabalho difere das ORfs encontradas pela aluna Mariana Rangel Pereira, foi feita uma árvore para mostrar a relação filogenética entre elas. A metodologia utilizada para a construção da árvore foi a mesma citada anteriormente, com *booststrap* de 500. O resultado pode ser observado na figura 17.



Figura 17- Dendrograma do agrupamento hierárquico mostrando a relação filogenética da ORF15 de esterase/lipase, encontrada no *contig* do clone PL28.F10 com outras ORFs codificadoras de esterase/lipase encontradas na mesma biblioteca metagenômica.

Por meio da análise da árvore filogenética é possível verificar que a ORF15 ficou localizada em um ramo único, agrupando-se com outros membros da mesma família. Sendo assim, é possível afirmar que esta ORF difere das outras ORFs encontradas em clones diferentes da mesma biblioteca metagenômica. Adicionalmente, a árvore corrobora com a localização da ORF na família IV das enzimas lipolíticas, visto que agrupou-se com representantes desta família.

5.7 Análise da ORF com sequências patenteadas

A sequência da ORF encontrada no clone PL28.F10 foi comparada com sequências patenteadas depositadas no NCBI, usando a ferramenta Blastp. A ORF15 teve 44% de identidade com uma proteína desconhecida utilizada para diagnóstico e terapia de infecções bacterianas, da companhia CAMBIA, cuja sequência está relacionada a *Acinetobacter baumannii*. (número de acesso AAQ29806.1, patente US 6562958). No entanto, apesar de estar caracterizada como desconhecida há presença do domínio conservado de esterases/lipases (BRETON e BUSH, 2003).

O próximo resultado indica 44% de identidade com uma proteína não noemada de *Burkholderia thailandensis* (número de acesso CAV33019.1, patente: WO 2008098199-A2), da UWM Research Foundation, com o dobramento das α/β hidrolases, relacionada à biossíntese de FK228 e seus análogos, utilizados no tratamento do câncer (CHENG, 2008). O

FK228 é um depsipeptídeo que inibe deacetilases, sendo que a acetilação de histonas está ligada à formação de tumores (FURUMAI et al., 2002).

Além das cinco sequências obtidas pela ferramenta Blast, foram selecionadas onze sequências de patentes conhecidas de esterases/lipases extraídas do NCBI para análise das relações filogenéticas com a ORF15, através do ClustalW e Mega 4.1, já descrito anteriormente (figura 18).

Através da construção filogenética foi possível observar que a ORF identificada neste trabalho não é igual a nenhuma sequência já patenteada, fato corroborado com o valor das identidades. A ORF15 se localizou em um ramo único, sendo um grupo irmão das sequências de esterase/lipase da BASF (Patente WO 0100842) e de uma proteína da companhia CAMBIA descrita anteriormente.



Figura 18 – Dendrograma do agrupamento hierárquico baseado em 16 sequências patenteadas extraídas no NCBI, com a ORF15 encontrada no clone PL28.F10.
6 CONCLUSÕES

- A análise da expressão funcional da atividade lipolítica, juntamente com as ferramentas de bioinformática, permitiram a caracterização de uma nova possível sequência gênica (ORF15) responsável pela expressão de enzimas lipolíticas da família IV das esterases/lipases;
- Embora com alguma modificações nos resíduos de aminoácidos, a tríade catalítica, essencial para a compreensão da atividade das lipases e esterases, esteve presente na caracterização da nova sequência gênica encontrada;;
- A técnica de sub-clonagem mostrou-se eficiente para obtenção do *contig* completo de clones produtores de enzimas lipolíticas;
- A biblioteca metagenômica do consórcio microbiano degradador de hidrocarbonetos de petróleo mostrou-se fonte de micro-organismos produtores de enzimas lipolíticas, bem como útil para expansão dos conhecimentos de tais enzimas, especialmente lipases e esterases;
- A caracterização da enzima lipolítica encontrada predizerão suas potenciais aplicações biotecnológicas futuras.

REFERÊNCIAS*

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

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.

BASF AG GERMAN PATENT. D. Engels, R. F. Mattes, B. Hauer, T. Friedrich, M. Ress-Loeschke. New nucleic acid sequence encoding *Alcaligenes faecalis* nitrilase polypeptide useful for converting racemic nitriles to chiral carboxylic acids. DE 19848129, 2000.

BORNSCHEUER, U. T. 2002. Microbial carboxyl esterases: classification, properties and application in biocatalysis. **FEMS Microbiology Review**, v. 26, p. 73–81, 2002.

CAMBIA PATENT LENS. Gary Breton e David Bush. Nucleic acid and amino acid sequences relating to Acinetobacter baumannii for diagnostics and therapeutics. US 6562959, 17 de agosto de 2003.

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.

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.

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

CURTIS, T. P.; SLOAN, W. T.; SCANNELL, J. W. Estimating prokaryotic diversity and its limits. **Proceeding of the National Academy of Sciences of the United States of America**, v. 99, p. 10494–10499, 2002.

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.

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.

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

FURUMAI, R.; MATSUYAMA, A.; KOBASHI, N.; LEE, K.-H.; NISHIYAMA, M.; NAKAJIMA, H.; TANAKA, A.; KOMATSU, Y.; NISHINO, N.; YOSHIDA, M.; HORINOUCHI, S. FK228 (Depsipeptide) as a Natural Prodrug That Inhibits Class I Histone Deacetylases. **Cancer Research**, v. 62, p. 4916–4921, 2002

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.

GILLESPIE, D. E.; BRADY, S. F.; BETTERMANN, A. D.; CIANCIOTTO, N. P.; LILES, M. R.; RONDON, M. R.; CLARDY, J.; GOODMAN, R. M.; HANDELSMAN, J. Isolation of antibiotics Turbomycin A and B from a Metagenomic Library of Soil Microbial DNA. **Applied and Environmental Microbiology** v. 68, n. 9, p. 4301–4306, 2002.

GORDON, D.; ABAJIAN, C.; GREEN, P. Consed: A graphical tool for sequence finishing. **Genome Research**, Toronto, v. 8, p. 195-202, 1998.

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.

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

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.

HU, Y.; ZHANG, G.; LI, A.; CHEN, J.; MA, L. Cloning and enzymatic characterization of a xylanase gene from soil derived metagenomic library with an efficient approach. **Applied Microbiology Biotechnology**, v. 80, p. 823-830, 2008

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.

JAEGER, K. E.; EGGERT, T. Lipases for Botechnology. Current Opinion in Biotechnology, v. 13, p. 390–397, 2002;

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

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, S.-J.; LEE, C.-M.; HAN, B.-R., KIM, M.-Y.; YEO, Y.-S.; YOON, S.-H.; KOO, B.-S.; JUN, H.-K. Characterization of a gene encoding cellulase from uncultured soil bacteria. **FEMS Microbiology Letters.** v. 282, n. 1, p. 44-51, 2008.

KIM, E. Y.; OH, K. H.; LEE, M. H.; KANG, C. H.; OH, T. K.; YOON, J. H. Novel coldadapted 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. 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**, v. 13, p. 572-577, 2002.

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

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

OBUEKWE, C. O.; AL-JADI, Z. K.; AL-SALEH, E. S. Comparative hydrocarbon utilization by hydrophobic and hydrophilic variants of *Pseudomonas aeruginosa*. Journal of Applied Microbiology. v. 105, p. 1876–1887, 2008.

OKAMURA, Y.; KIMURA, T.; YOKOUCHI, H.; MENESES-OSORIO, M.; KATOH, M.; MATSUNAGA, T.; TAKEYAMA, H. Isolation and Characterization of a GDSL Esterase from the Metagenome of a Marine Sponge-associated Bacteria. **Marine Biotechnology**, v. 12, p. 395-402, 2010.

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.

PARALES, R. E.; DITTY, J. L.; HARWOOD, C. S. Toluene-Degrading Bacteria Are Chemotactic towards the Environmental Pollutants Benzene, Toluene, and Trichloroethylene. **Applied and Environmental Microbiology**, v. 66, p. 4098-4104, 2000.

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.

RITZ, K.; GRIFFTHS, B. S. Potential application of a community hybridization technique for assessing changes in the population structure of soil microbial communities. **Soil Biology and Biochemistry**, v. 26, p. 963-971, 1994.

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.

SAID, S.; PIETRO, C. L. R. Generalidades sobre aplicação industrial de enzimas. In:_____. **Enzimas como agentes biotecnológicos.** Ribeirão Preto: Legis Summa, 2004. p. 1-7.

SAITO, N.; NEI, M. The neighbor-joining metod: 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.

SCHMIDT, T.; DELONG, E.; PACE, N. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. Journal of Bacteriology, v. 173, p. 4371-4378, 1991.

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.

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.

TAYLOR, S.J.; BROWN, R. C.; KEENE, P. A.; TAYLOR, I. N. Novel screening methods – the key to cloning commercially successful biocatalysts. **Bioorganic & Medicinal Chemistry**, v. 7, p. 2163-2168, 1999.

TISS, A.; CARRIÈRE, F.; VERGER, R. Effects of Gum Arabic on Lipase Interfacial Binding and Activity. **Analytical Biochemistry**, v. 294, p. 36–43, 2001.

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

UCHIYAMA, T.; MIYAZAKI, K. Functional metagenomics for enzyme discovery: challenges to efficient screening. **Current Opinion in Biotechnology**, v. 20, p. 616–622, 2009.

UWM RESEARCH FOUNDATION, INC. Yi-Giang Cheng. Sequences for fk228 biosynthesis and methods of synthesizing fk228 and fk228 analogs. WO 2008098199-A2, 14 de Agosto 2008.

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.

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

WHYTE, L. G.; SMITS, T. H. M.; LABBE, D.; WITHOLT, B.; GREER, C. W.; VAN BEILEN, J. B. Gene Cloning and Characterization of Multiple Alkane Hydroxylase Systems in *Rhodococcus* Strains Q15 and NRRL B-16531. Applied and Environmental Microbiology, v. 68, p. 5933-5942, 2002.

WHYTE, L. G.; BOURBONNIERE, L.; GREER, C. W. Biodegradation of Petroleum Hydrocarbons by Psychrotrophic *Pseudomonas* Strains Possessing Both Alkane (*alk*) and Naphthalene (*nah*) Catabolic Pathways. **Applied and Environmental Microbiology**, v. 63, p. 3719-3723, 1997.

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.

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.

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

ANEXO A - Comparação do *contig* completo do clone PL28.F10, de 36.942 pb, com as sequências do NCBI, usando a ferramenta BLASTN (nucleotide collection nr/nt).



Descriptions

Legend for links to other resources: 🛄 UniGene 🖪 GEO 🖸 Gene 🗧 Structure 🕅 Map Viewer 🖉 PubChem BioAssay

Accession	Description	<u>Max</u> score	Total score	Query coverage	$\Delta \frac{\underline{E}}{value}$	<u>Max</u> ident
CP000438.1	Pseudomonas aeruginosa UCBPP-PA14, complete genome	3260	1.213e+04	35%	0.0	91%
CP000744.1	Pseudomonas aeruginosa PA7, complete genome	3245	1.216e+04	35%	0.0	91%
FM209186.1	Pseudomonas aeruginosa LESB58 complete genome sequence	2473	1.221e+04	35%	0.0	92%
AE004091.2	Pseudomonas aeruginosa PAO1, complete genome	2473	1.223e+04	35%	0.0	91%
CT573326.1	Pseudomonas entomophila str. L48 chromosome, complete sequence	1862	6529	19%	0.0	89%
CP000680.1	Pseudomonas mendocina ymp, complete genome	1814	9832	27%	0.0	97%
AE015451.1	Pseudomonas putida KT2440 complete genome	<u>1773</u>	7668	26%	0.0	89%
CP000712.1	Pseudomonas putida F1, complete genome	1746	7230	25%	0.0	89%

ANEXO B – Comparação da ORF15, de 303 aminoácidos, com as sequências do NCBI, usando a ferramenta BLASTp (*non-redundant protein sequence* – nr).



Legend for links to other resources: 🛄 UniGene 🖪 GEO 🖸 Gene 🗧 Structure 🕅 Map Viewer 🌌 PubChem BioAssay

Accession	Description	<u>Max</u> <u>score</u>	<u>Total</u> <u>score</u>	<u>Query</u> coverage	$= \frac{\underline{E}}{value}$
AAX37296.1	lipase/esterase [uncultured bacterium]	353	353	90%	1e-95
ADA70030.1	Est10 [uncultured marine bacterium]	347	347	92%	2e-93
AAS77233.1	lipase/esterase [uncultured bacterium]	339	339	90%	3e-91
AAX37295.1	lipase/esterase [uncultured bacterium]	338	338	88%	5e-91
AAS77238.1	lipase/esterase [uncultured bacterium]	338	338	90%	7e-91
AAS77236.1	lipase/esterase [uncultured bacterium]	337	337	90%	1e-90
ACL67845.1	lipolytic enzyme [uncultured bacterium]	330	330	88%	2e-88
ACL67849.1	lipolytic enzyme [uncultured bacterium]	329	329	88%	3e-88