

UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

SECRETARIA
DE
PÓS-GRADUAÇÃO
I. B.

ISABEL NATALIA SIERRA GARCIA

**“CARACTERIZAÇÃO ESTRUTURAL E FUNCIONAL DE GENES DE
DEGRADAÇÃO DE HIDROCARBONETOS ORIGINADOS DE
METAGENOMA MICROBIANO DE RESERVATÓRIOS DE PETRÓLEO”**

Este exemplar corresponde à redação final
da tese defendida pelo(a) candidato (a)
ISABEL NATALIA SIERRA GARCIA.
e aprovada pela Comissão Julgadora.

Dissertação apresentada ao
Instituto de Biologia para
obtenção do Título de Mestre em
Genética e Biologia Molecular, na
área de Genética de Micro-
organismos.

Orientadora: Profa. Dra. Valeria Maia Merzel.

Co-Orientadora: Profa. Dra. Anete Pereira de Souza

Campinas, 2011

FICHA CATALOGRÁFICA ELABORADA POR
ROBERTA CRISTINA DAL' EVEDOVE TARTAROTTI – CRB8/7430
BIBLIOTECA DO INSTITUTO DE BIOLOGIA - UNICAMP

Si17c	<p>Sierra Garcia, Isabel Natalia, 1984- Caracterização estrutural e funcional de genes de degradação de hidrocarbonetos originados de metagenoma microbiano de reservatórios de petróleo / Isabel Natalia Sierra Garcia. – Campinas, SP: [s.n.], 2011.</p> <p>Orientadores: Valeria Maia Merzel, Anete Pereira de Souza. Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Biologia.</p> <p>1. Biodegradação. 2. Hidrocarbonetos. 3. Metagenômica. 4. Reservatórios. 5. Petróleo. I. Merzel, Valeria Maia. II. Souza, Anete Pereira de, 1962-. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.</p>
-------	--

Informações para Biblioteca Digital

Título em Inglês: Structural and functional characterization of hydrocarbon degradation genes from microbial metagenome derived from petroleum reservoirs

Palavras-chave em Inglês:

Biodegradation

Hydrocarbons

Metagenomics

Reservoirs

Petroleum

Área de concentração: Genética de Microorganismos

Titulação: Mestre em Genética e Biologia Molecular

Banca examinadora:

Valeria Maia Merzel [Orientador]

Fernando Dini Andreote

Laura Maria Mariscal Ottoboni

Data da defesa: 11-07-2011

Programa de Pós Graduação: Genética e Biologia Molecular

Campinas, 11 de julho de 2011

BANCA EXAMINADORA

Profa. Dra. Valeria Maia Merzel (Orientadora)


Assinatura

Prof. Dr. Fernando Dini Andreote


Assinatura

Profa. Dra. Laura Maria Mariscal Ottoboni


Assinatura

Prof. Dr. Welington Luiz de Araújo

Assinatura

Profa. Dra. Suzete Aparecida Lanza Destefano

Assinatura

DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de minha dissertação de Mestrado intitulada:
"Caracterização estrutural e funcional de genes de degradação de hidrocarbonetos originados de metagenoma microbiano de reservatórios de petróleo"

() não se enquadra no § 3º do Artigo 1º da Informação CCPG 01/08, referente a bioética e biossegurança.

Tem autorização da(s) seguinte(s) Comissão(ões):

(X) CIBio – Comissão Interna de Biossegurança , projeto No. **23/2009**. Instituição: **Centro Pluridisciplinar de Pesquisas Químicas, Biológicas e Agrícolas – CPQBA/UNICAMP.**

() CEUA – Comissão de Ética no Uso de Animais , projeto No. _____, Instituição: _____

() CEP - Comissão de Ética em Pesquisa, protocolo No. _____, Instituição: _____

** Caso a Comissão seja externa ao IB/UNICAMP, anexar o comprovante de autorização dada ao trabalho. Se a autorização não tiver sido dada diretamente ao trabalho de tese ou dissertação, deverá ser anexado também um comprovante do vínculo do trabalho do aluno com o que constar no documento de autorização apresentado.*

Isabel Natalia Sierra Garcia

Aluno: Isabel Natalia Sierra Garcia

Valeria Maia de Oliveira

Orientadora: Valeria Maia de Oliveira

Para uso da Comissão ou Comitê pertinente:

(X) Deferido () Indeferido

Carimbo e assinatura

Prof. Dr. MARCELO LANCELLOTTI
Presidente da Comissão Interna de Biossegurança
Instituto de Biologia - UNICAMP

Para uso da Comissão ou Comitê pertinente:

() Deferido () Indeferido

Carimbo e assinatura

Aos meus pais,

Por me dar a liberdade para escolher minha profissão,

Por que sem sua compreensão e apoio, este sonho não teria sido possível

A mis papis,

Por darme la libertad para elegir mi carrera,

Por que sin su comprensión y apoyo, este sueño no hubiera sido posible

AGRADECIMENTOS

Ao término desta etapa da minha vida, estou ciente de que a conquista deste objetivo não é só minha, por isso aqui vai minha melhor tentativa para expressar meus mais sinceros agradecimentos àquelas pessoas que fizeram parte desta aventura que significou meu mestrado, e que incondicionalmente, me apoiaram e entenderam minhas emoções, minhas alegrias, mas também minhas ausências e maus momentos.

Sinto-me afortunada de ter encontrado desde o começo desta experiência o apoio de uma pessoa que sem me conhecer, abriu suas portas e torceu por mim. Minha orientadora, a Dra. Valeria Maia cujo entusiasmo, simpatia e vitalidade são também um catalisador do meu desempenho e meu mais admirável exemplo de pesquisadora e ser humano integra. Muito obrigada Val, por ter confiado em mim, por ter me introduzido ao estudo da metagenômica, área que eu gosto tanto e que desde sempre, sem defini-la como metagenômica, estive procurando. Muito obrigada pelo acolhimento, por ter possibilitado várias oportunidades para o meu crescimento intelectual e pessoal, por me fazer sentir bem-vinda, protegida e pelo seu abraço sincero.

A próxima pessoa que merece meu mais sincero agradecimento, é a pessoa com quem compartilhei toooooodas minhas dúvidas, desassossegos e satisfações durante esses dois anos, meu *moquito* Javi, por estar comigo sempre... Desde suas sacrificadas idas ao lab de sábados e domingos, seu ouvido atento ao final de cada dia e sua especial atenção aos detalhes metodológicos das análises bioinformáticas. *Moqui*: você foi o suporte acadêmico e emocional mais importante que eu tive durante o desenvolvimento desta tese. Sem seus aportes e sem sua compreensão durante as etapas mais difíceis, deste processo, assim como durante minhas sobrecargas mentais, este objetivo teria sido muito mais difícil de alcançar. Te amo e te admiro muito!

Tenho muito por que agradecer também à Dra. Suzan P. Vasconcellos, primeiro pela confiança que depositou em mim ao deixar em minhas mãos o fruto do seu árduo trabalho de dois anos na construção da biblioteca metagenômica. Muito obrigada também por que eu pude ver em você uma amiga, que apesar do pouco o tempo que compartilhamos juntas dentro do lab, foi muito proveitoso. Seus ensinamentos, supervisão, amizade sincera, e parceria foram muito, muito valiosos para mim. Conclui que não pude ter uma melhor introdutora na DRM!. Muito Obrigada, Su! Além disso, suas participações em varias etapas do trabalho assim como na banca de qualificação e de banca previa foram fundamentais no desenvolvimento e conclusão desta dissertação.

Sinto-me muito afortunada também pela oportunidade durante estes dois anos de ser acolhida e apoiada em vários grupos de pesquisa especializados onde desempenhei parte dos experimentos deste trabalho. Por este motivo quero agradecer à professora Dra. Anita Marsaioli do Instituto de química da UNICAMP por permitir que a Suzan e eu pudéssemos no início deste trabalho realizar os ensaios da capacidade de degradação dos clones. À Profa. Dra. Anete Pereira de Souza e muito especialmente aos membros do seu Laboratório de Análise Genética e Molecular do CBMEG/UNICAMP: Lilian, Marcelo, Clelton, Alexandre, Antonio e no final, Mariana e Aline, pelas facilidades técnicas, ensinamentos, espírito de amizade e por apanharmos juntos nessa primeira subclonagem. Pelas muitas risadas e bons momentos!.

Aos professores membros da banca de qualificação e previa de defesa, Prof. Dr. Fernando Dini Andreote, Dra. Suzan P. Vasconcellos, Dra. Lara D. Sette, Dr. Renato Vicentini, pelas valiosas discussões e sugestões que levaram à conclusão deste trabalho.

À PETROBRAS pelo auxílio financeiro na pesquisa e na concessão da bolsa.

Aos todos os colegas atuais e aqueles que já passaram durante minha estadia na Divisão de Recursos Microbianos. Farei um esforço para citar todos: Mariana, Gisele, Pati, Maria Rafaella, Fernando, Rafaella, Alysson, Claudia, Michel, Leandro, Bruna, Daniela, Júlia, Barbara, Cynthia, João, Geize, Cistina, Tiago, Paula, à pesquisadora Dra. Fabiana e a super colaboração das técnicas de laboratório, Samantha e Milena, pela paciência e disposição para minhas inquietudes, materiais e reagentes. À todos, muito obrigada pelo companherismo na convivência diária, pela colaboração e por fazerem a rotina do laboratório com seus altos e baixos, mais agradável.

À galera colombiana em Campinas, pelo apoio incondicional, emocional, pelos inumeráveis momentos compartilhados, pelas discussões sobre o nosso país, pelas festas latinas, por trazerem um pouquinho da Colômbia em cada um. À galera estrangeira de começos de 2009, Franceses, Argentinos e Espanhóis, tornando meu primeiro semestre aqui muito mais feliz. Nesse mesmo ano às meninas da casa 11: Dani-gaucha, Renata, Pati e da casa 7: Joyce, Ceci, Tais e especialmente a Carol por adotar esta colômbiana e pelos esforços por compreender Natalês!

Às minhas queridas *parcerinhas* e alegrias campineiras, meu coração inteiro esta cheio de um profundo agradecimento, a mi Claudi, pelos dois anos inteiros de uma maravilhosa amizade que começava sem futuro num ponto de ônibus, mas que virou a mais valiosa e sincera companhia de VIDA no Brasil, a mi Marichi por sua amizade leal, por ser o mais recente suporte onde deslizam meus problemas, obrigada pela companhia ao final desta caminhada, pelas miutas risadas, e por tornar menos sofrível os nossos dias de isolamento na escrita das nossas dissertações. Às duas: por cada momento compartilhado, grande, pequeno, triste, alegre, pelas viagens, os desafogos, por

que não poderia ter encontrado neste caminho uma companhia melhor para abrir meu coração...
Las quiero un mundo y no sé que hubiera sido de mi sin ustedes!!

Aos meus pais, Rodolfo e Elsa, à minha tia Martha, aos meus irmãos, Rodolfo e Santiago, às minhas cunhadas, Alexandra e Karina e a minha princesa bela, minha sobrinha Mariana. À minha família inteira pelo infinito apoio, apesar dos mais de 4000 km de distância, pelas velas acesas, incentivos e boas vibes nos momentos mais desafiadores que enfretei nestas terras, pelo carinho, e principalmente por respeitar e acreditar em mim e nos meus sonhos.

Finalmente, agradeço muito a vida, por ter me permitido conhecer cada um das pessoas que nomeei e a aquele que busca seu nome nesta longa lista, por ter-nos encontrado no mesmo caminho um dia, por que gosto de pensar que cada pessoa que passa na nossa vida representa um mundo único que quando passa não nos deixa só, por que deixa um pouco do seu mundo e leva um pouquinho do nosso. Esses mundos que teve oportunidade de conhecer e de quem aprendi muito nesta caminhada fizeram quem eu sou hoje e são a prova de que as pessoas não se encontram na vida por acaso...

*Meu mais sincero agradecimento a todos,
Por me permitir ir além das fronteiras traçadas no nosso mundo,
Por ter ganhado um coração Brasileiro,*

Natalia

*“The best thing in life is to go ahead with all your plans and your dreams,
to embrace life and to live everyday with passion,
to lose and still keep the faith and to win while been grateful.
All of this because the world belongs to those who dare to go after what they
want. And because life is really too short to be insignificant.”*

Charlie Chaplin

RESUMO

A ocorrência de óleos biodegradados nos reservatórios de petróleo, juntamente com o isolamento das primeiras bactérias a partir destes ambientes, forneceu evidências de micro-organismos ativos no subsolo profundo. Os micro-organismos, habitantes comuns de reservatórios de petróleo, desenvolveram estratégias eficazes de biodegradação, baseados em sistemas enzimáticos e vias metabólicas especializadas, de maneira a acessar os hidrocarbonetos como fonte de carbono e energia. No entanto, o conhecimento atual da diversidade microbiana e dos processos metabólicos envolvidos na biodegradação em reservatórios de petróleo ainda é limitado, principalmente devido à dificuldade em recuperar a complexa comunidade microbiana presente em tais ambientes extremos. Essa limitação imposta pelo uso de técnicas de cultivo convencionais pode ser superada através do uso de abordagens independentes de cultivo, como a metagenômica, a qual permite o acesso ao potencial metabólico dos micro-organismos não cultivados. Em trabalho prévio, uma biblioteca metagenômica fosmidial derivada de enriquecimentos aeróbios e anaeróbios de petróleo foi construída e avaliada para a capacidade de crescimento microbiano em hexadecano como fonte de única de carbono. No presente estudo, foi analisada a capacidade de biodegradação de hexadecano e fenantreno pelos clones selecionados dessa biblioteca através de cromatografia gasosa acoplada a espectrometria de massas (CG-EM). As análises que se sucederam tiveram como objetivos a identificação e caracterização das sequências metagenômicas responsáveis pela biodegradação de hidrocarbonetos. A estratégia de clonagem aleatória “*shotgun*” seguida de sequenciamento dos clones foi utilizada para determinar a sequência inteira dos insertos fosmidias que mostraram relevantes capacidades de biodegradação. Cerca de 30 kb de sequência foram montados para o fosmideo 1A (91% de degradação do hexadecano e 5% do fenantreno) e 37 kb para o fosmideo 2B (98% de degradação do hexadecano e 44% do fenantreno), e em cada um foram reconhecidos 23 ORFs e 40 ORFs, respectivamente. As proteínas putativas foram identificadas por comparação das sequências de aminoácidos deduzidas. Várias proteínas envolvidas na degradação aeróbia e anaeróbia de diferentes compostos de hidrocarbonetos foram encontradas. As sequências que codificam estas enzimas não mostraram-se agrupadas em clusters completos de degradação, semelhantes aos encontrados nas bactérias degradadoras de hidrocarbonetos conhecidas. Os fragmentos metagenômicos continham apenas subconjuntos de genes pertencentes a várias vias, mostrando novos arranjos gênicos. Esses resultados reforçam o potencial da abordagem metagenômica para a prospecção e elucidação de novos genes e vias metabólicas, contribuindo para uma visão mais abrangente dos processos de biodegradação que ocorrem nos reservatórios de petróleo.

ABSTRACT

The occurrence of biodegraded oil in petroleum reservoirs, along with the isolation of the first bacteria from such environments, provided evidence for active microorganisms in the deep subsurface. Microorganisms are common inhabitants of petroleum reservoirs and they may have effective biodegrading strategies, based on specialized enzyme systems and metabolic pathways, as a form to access hydrocarbons as carbon and energy sources. However, the current knowledge of the microbial diversity and metabolic pathways involved in hydrocarbon biodegradation in petroleum reservoirs is still limited, mostly due to the difficulty in recovering the complex community present in such extreme environments. This limitation imposed by conventional culturing techniques can be circumvented by the metagenomic approach, which is a culture-independent molecular method that allows the access to the metabolic potential of previously uncultured microorganisms. In a previous work, a metagenomic fosmid library derived from aerobic and anaerobic petroleum enrichments were constructed and screened for the detection of microbial growth on hexadecane as sole carbon source. In this study, the biodegradation abilities on hexadecane and phenanthrene of selected clones were analyzed using Gas Chromatography – Mass Spectroscopy (GC-MS) and subsequent analyses aimed at the identification and characterization of metagenomic sequences responsible for the hydrocarbon biodegradation. The shotgun sequencing approach was used to determine the whole insert sequence of two fosmid clones showing different and relevant biodegradation capacities, FOS1A (91% hexadecane and 5% phenanthrene degradation) and FOS2B (98% hexadecane and 44% phenanthrene degradation). About 30 kb in sequence were assembled for the fosmid 1A and 37 kb for fosmid 2B, and 23 ORFs and 40 ORFs were recognized in each one, respectively. Putative proteins were identified by comparison of deduced aminoacid sequences. Several proteins involved in the degradation of different hydrocarbon compounds by aerobic and anaerobic pathways were found. The sequences coding these enzymes were not grouped together into complete clusters of degradation pathways similar to those already described in known hydrocarbon degrading bacteria. The metagenomic fragments contained only subsets of gene belonging to different pathways, showing novel gene arrangements. The results obtained in this study reinforce the potential of the metagenomic approach for the prospection and elucidation of new genes and pathways, contributing to a broader perspective of the biodegradation processes that take place in petroleum reservoirs.

ÍNDICE

1. INTRODUÇÃO	15
2. REVISÃO BIBLIOGRÁFICA.....	17
2.1. O petróleo	17
2.2. O reservatório de petróleo.....	17
2.3. Diversidade microbiana em reservatório de petróleo	19
2.4. Degradação microbiológica do petróleo nos reservatórios	21
2.5. Degradação microbiológica de hidrocarbonetos	22
2.6. Mecanismos e genética da degradação microbiológica de hidrocarbonetos.....	24
2.6.1. Degradação aeróbia	25
2.6.2. Degradação anaeróbia	30
2.7. Métodos independentes de cultivo	35
2.8. Metagenômica	36
2.9. Ferramentas para análise de sequências metagenômicas	38
2.9.1. Análises de bibliotecas metagenômicas baseada em similaridade de sequência	40
2.9.2. Análise de bibliotecas metagenômicas baseada em função.....	40
2.11. Processamento de sequências metagenômicas	43
2.11.1. Pré-processamento de sequências (leituras)	43
2.11.2. Montagem.....	44
2.11.3. Predição de genes e anotação	46
3. JUSTIFICATIVA	48
4. OBJETIVOS	49
4.1. Objetivo geral.....	49
4.2. Objetivos específicos.....	49

5.	MATERIAL E MÉTODOS	49
5.1.	Origem dos clones fosmidiais.....	49
5.2.	Análises cromatográficas da biodegradação de Hexadecano e Fenantreno	50
5.3.	Construção de Sub-bibliotecas.....	51
5.3.1.	Extração de DNA fosmidial.....	52
5.3.2.	Produção de fragmentos aleatórios de DNA fosmidial.....	52
5.3.3.	Reação de reparo das extremidades dos fragmentos de DNA	52
5.3.4.	Reação de fosforilação	52
5.3.5.	Reação de ligação.....	53
5.3.6.	Produção e transformação de células competentes.....	53
5.4.	Verificação da presença e tamanho de inserto nos clones das sub-bibliotecas.....	55
5.5.	Sequenciamento completo da sub-biblioteca	55
5.6.	Análise das sequências geradas através de bioinformática	56
5.6.1.	Limpeza e montagem de sequências	56
5.6.2.	Identificação de quadros de leitura abertos	56
5.6.3.	Anotação de genes.....	57
5.6.4.	Determinação das relações evolutivas.....	58
6.	RESULTADOS E DISCUSSÃO	58
6.1.	Análises cromatográficas de hexadecano e fenantreno	58
6.2.	Construção de sub-bibliotecas.....	61
6.2.1.	Extração de DNA fosmidial.....	61
6.2.2.	Produção de fragmentos aleatórios de DNA fosmidial (Shotgun library).....	61
6.2.3.	Reações de reparo/fosforilação/ligação e transformação de células competentes.	62
6.3.	Avaliação de sub-bibliotecas e sequenciamento	62
6.4.	Montagem das sequências.....	63

6.5.	Características gerais dos insertos fosmidiais.....	64
6.6.	Anotação do fosmideo 2B.....	66
6.6.1.	Organização genética do fosmideo 2B.....	70
6.6.2.	ORFs relacionadas às funções metabólicas de degradação aeróbia de aromáticos.	71
6.6.3.	ORFs relacionadas às funções metabólicas de degradação anaeróbia de aromáticos.	77
6.6.4.	ORFs relacionadas a outras características metabólicas.....	80
6.6.5.	ORFs relacionadas à respiração e metabolismo energético	82
6.6.6.	Genes relacionados a outras características e funções celulares	83
6.7.	Anotação do fosmideo 1A.....	85
6.7.1.	Organização genética do fosmideo 1A.....	88
6.7.2.	ORFs relacionadas às funções metabólicas de degradação aeróbia.....	88
6.7.3.	ORFs relacionadas a outros processos celulares e metabólicos	90
6.8.	Considerações sobre a afiliação filogenética dos fosmídeos	92
6.9.	Considerações sobre as relações evolutivas dos genes de degradação de hidrocarbonetos	94
7.	CONCLUSÕES.....	96
8.	REFERÊNCIAS.....	98

1. INTRODUÇÃO

A biodegradação do petróleo e produtos derivados tem sido de enorme interesse para os cientistas ao longo do tempo. A maioria das pesquisas dedicadas à compreensão dos danos associados a este fenômeno foi feita entre os anos 50 e 70. No entanto, atualmente a degradação de hidrocarbonetos e suas implicações são uma questão urgente mundial.

Estudos realizados em reservatórios de petróleo têm evidenciado que parte da comunidade microbiana associada a este tipo de ambiente é representada por bactérias e arqueias de distribuição geográfica bastante ampla, e que diversos destes organismos têm potencial para transformar compostos orgânicos e inorgânicos, atuando na interface óleo-água dos reservatórios (Coates et al., 1997; Korda et al., 1997).

Comunidades microbianas em petróleo vêm sendo extensamente estudadas nos últimos anos, principalmente, no que diz respeito à capacidade de degradar ou biotransformar componentes específicos do óleo (Leahy e Colwell, 1990; Bressler e Fedorak, 2001). A diversidade dos microrganismos detectados em amostras de petróleo inclui grupos de bactérias e arqueias aeróbias, anaeróbias facultativas e microaerófilas, sendo sugerido que esses microrganismos são introduzidos nos reservatórios através da infiltração de águas superficiais, decorrente de características geológicas locais ou dos sistemas de injeção de água para exploração (Magot et al., 2000).

Embora as técnicas de cultivo tenham sido aprimoradas, sabe-se que apenas uma pequena fração da diversidade microbiana presente na natureza, pode ser hoje cultivada em laboratório (Amann et al., 1995). Metodologias moleculares desenvolvidas nas últimas décadas (extração de ácidos nucleicos, amplificação por PCR, clonagem e sequenciamento de DNA) têm sido otimizadas e adaptadas para superar as limitações impostas pela abordagem clássica de estudo de populações microbianas, evitando o isolamento e cultivo dos microrganismos. A utilização destas metodologias vem permitindo uma avaliação mais precisa da diversidade microbiana no ambiente e a descoberta de novos grupos de organismos, nunca antes cultivados (Hugenholtz et al., 1998).

Neste contexto, nosso grupo de pesquisa vem desenvolvendo estudos independentes de cultivo que visam a contribuir a um melhor entendimento dos micro-organismos e processos biológicos que acompanham a biodegradação em reservatórios (Projeto PETROBRAS “Estudos metagenômicos da biodegradação microbiana em reservatórios de petróleo”). Uma das abordagens teve como objetivo a construção de bibliotecas metagenômicas a partir de enriquecimentos microbianos aeróbios e anaeróbios de petróleo, e a subsequente triagem de atividades enzimáticas dos clones, com especial interesse na degradação de hidrocarbonetos, produção de biosurfactantes e de enzimas hidrolíticas. Uma coleção de cerca de 31.000 clones já

foi obtida, sendo que os resultados preliminares revelaram a presença de genes de degradação de hidrocarbonetos através de ensaios funcionais. Estes resultados precisam ser estudados com maior detalhe a fim de avaliar a capacidade degradadora dos clones, e de caracterizar a estrutura e a diversidade dos genes de degradação de hidrocarbonetos em reservatórios de petróleo brasileiros.

Os resultados do projeto citado acima representam uma contribuição inovadora com relação às estratégias de pesquisas de mecanismos de degradação em reservatórios de petróleo e serviram como base para delinear o propósito do presente trabalho. Este envolveu a caracterização estrutural e funcional de genes envolvidos na degradação de hidrocarbonetos derivados do metagenoma microbiano originado de reservatórios de petróleo brasileiros. Cabe ressaltar que este representa um estudo inédito nestes ambientes e vêm abrir fronteiras para uma compreensão mais abrangente dos processos biogeoquímicos responsáveis pela degradação do petróleo.

2. REVISÃO BIBLIOGRÁFICA

2.1.O petróleo

O petróleo cru é uma mistura complexa e heterogênea contendo dezenas de milhares de compostos orgânicos quimicamente distintos (Fernandez-Lima et al., 2009). Geralmente são agrupados em quatro frações de acordo com sua solubilidade em solventes orgânicos: hidrocarbonetos saturados, hidrocarbonetos aromáticos, resinas e asfaltenos. Até recentemente, o interesse na caracterização rápida e precisa do perfil químico dos componentes do petróleo cru, motivado pela importância deste recurso na indústria em função do custo crescente do barril de óleo e pela preferência do mercado por óleos leves cada vez mais esgotados, entre outros aspectos, vem permitindo estudos mais completos para a determinação da complexa composição do petróleo. Análises através de espectrofotometria de massas de alta resolução têm revelado a identificação de mais de 17.000 compostos químicos distintos (Marshall e Rodgers, 2004).

A composição do petróleo também é descrita em termos da proporção de (1) hidrocarbonetos saturados, como os alcanos (compostos de carbono e hidrogênio unidos por ligações simples), os isoalcanos e cicloalcanos, (2) hidrocarbonetos aromáticos e (3) compostos não hidrocarbonetos contendo oxigênio, nitrogênio, e enxofre (Magot, 2005). Os alcanos, dependendo da fonte de petróleo bruto, constituem cerca de 20 a 50% do petróleo. Os n-alcanos com menos de 5 átomos de carbono são gasosos em condições de temperatura e pressão normais, enquanto aqueles que possuem de 5 a 15 átomos são líquidos e os com mais de 15 átomos variam de líquidos a sólidos. Os hidrocarbonetos aromáticos são compostos que apresentam o anel aromático (benzeno), aqueles com mais de um anel são chamados de naftaleno (2 anéis) e fenantreno (3 anéis). O tolueno, com apenas um núcleo benzênico, é o composto aromático mais comum no petróleo, seguido pelo xileno e benzeno. Finalmente, os não hidrocarbonetos são compostos que possuem outros elementos além do carbono e hidrogênio, denominados heteroátomos. As resinas e asfaltenos são exemplos deste grupo, apresentam alto peso molecular e são pouco solúveis em solventes orgânicos (Petroleum Geoscience Technology - www.pgt.com.br).

2.2.O reservatório de petróleo

Na atualidade, as reservas mundiais de petróleo (Figura 1) são dominadas por depósitos que têm sido degradados microbiologicamente ao longo do tempo geológico e representam a manifestação mais dramática da existência de biosfera profunda (Head et al., 2010). Apesar da ocorrência generalizada de reservatórios de petróleo biodegradados pelos processos microbianos, a compreensão da degradação de hidrocarbonetos de petróleo em reservatórios de petróleo

continua sendo um grande desafio. Até hoje nenhum micro-organismo anaeróbico degradador de hidrocarbonetos foi isolado diretamente de reservatório de petróleo degradado, prevalecendo os paradigmas relacionados com os processos de biodegradação em reservatórios (Head et al., 2010).

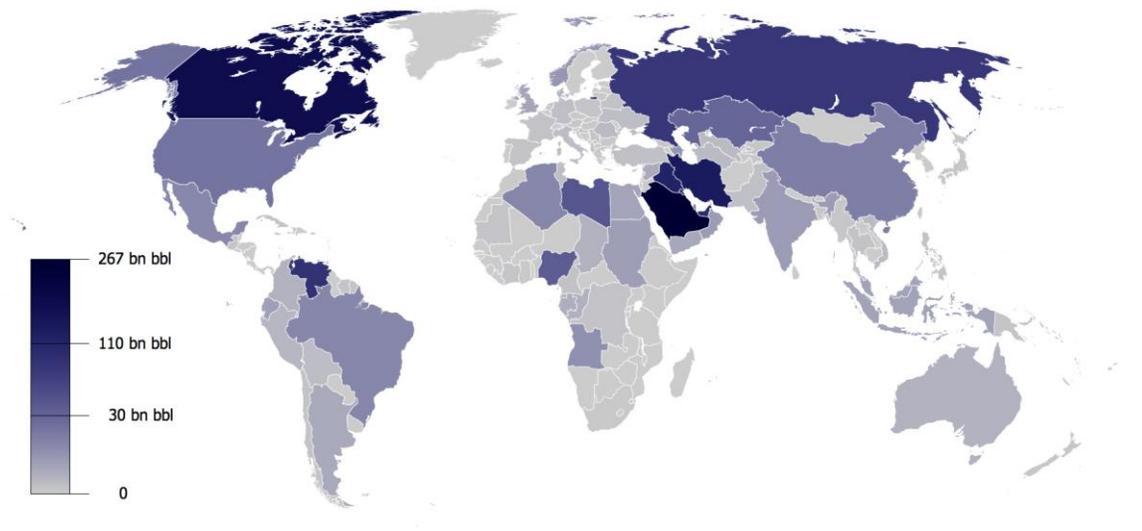


Figura 1. Reservas comprovadas de petróleo no mundo. 1 bn bbl -> 1000 milhões de barris de petróleo. Fonte: CIA – *Factbook* do mundo em 2009 (<https://www.cia.gov/library/publications/the-world-factbook/rankorder/2178rank.html>)

Os efeitos da biodegradação na composição molecular e propriedades físicas do petróleo são relativamente bem conhecidos (Head et al., 2003). Os dados a partir de óleos biodegradados indicam que os hidrocarbonetos do petróleo são biodegradados sequencialmente, na ordem: *n*-alcanos, alcanos monocíclicos, alquil-benzenos, alcanos isoprenóides, alquil-naftalenos, alcanos bicíclicos, esteranos e hopanos (Sun et al., 2005; Cunha et al., 2006). Entretanto, os mecanismos envolvidos no processo de biodegradação do petróleo nos reservatórios profundos, assim como o envolvimento de micro-organismos, as vias pelas quais eles utilizam os compostos do óleo e as condições sob as quais eles crescem, ainda permanecem pouco compreendidas (Roling et al., 2003; Cunha et al., 2006; Head et al., 2010).

As alterações físico-químicas do óleo durante o processo de degradação microbiológica nos reservatórios incluem a diminuição no conteúdo de hidrocarbonetos saturados, o consequente incremento de compostos polares, como resinas e asfaltenos, e o incremento da densidade, viscosidade e teor de enxofre e acidez, resultando na transformação de óleos leves em óleos mais pesados (Figura 2). Os óleos pesados tornam mais difíceis as operações de produção que os óleos

leves, pois leva a problemas de corrosão e complicações no refino, transporte e processamento do petróleo. Juntas, essas mudanças na qualidade do óleo têm implicações econômicas negativas para a produção e operações de refino do petróleo (Röling et al., 2003; Head et al., 2003; Cunha et al., 2006).

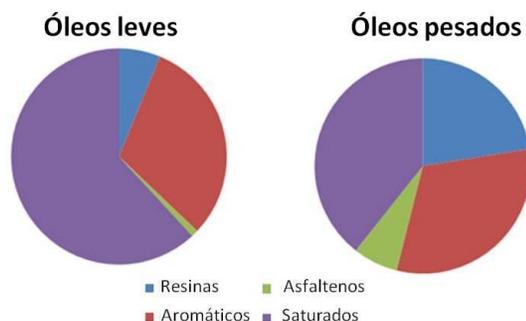


Figura 2. Proporção da composição de hidrocarbonetos em óleos leves e pesados. Fonte: <http://www.geochembio.com/ecology/petroleum/about-petroleum.html>

A avaliação do risco de encontrar reservatórios de petróleo degradados, e, portanto, menos rentáveis durante a exploração requer a previsão do grau de biodegradação que pode ser esperado em um reservatório e dependerá do conhecimento da microbiologia da degradação de hidrocarbonetos em reservatórios, dos micro-organismos envolvidos, das vias pelas quais eles utilizam os componentes do petróleo e as condições em que vivem (Röling et al., 2003).

Na atualidade, os estudos da degradação de hidrocarbonetos na indústria do petróleo são motivados não só pelas complicações econômicas da biodegradação, mas também pelo conhecimento da geoquímica dos reservatórios de petróleo (biodegradação e acidificação do óleo), pela biorremediação de locais contaminados e pela elucidação de novas reações bioquímicas empregadas nas vias de degradação. Além disso, existe também o grande interesse em biocatálise, pois várias reações enzimáticas específicas para compostos em vias de degradação individuais poderiam representar biocatalistas inovadores para aplicações industriais (Heider e Rabus, 2008) e sua possível aplicação na recuperação melhorada do petróleo - MEOR (Magot, 2005).

2.3. Diversidade microbiana em reservatório de petróleo

Desde a primeira observação de micro-organismos viáveis em águas de campo de petróleo (Bastin et al., 1926), tem havido um debate sobre a existência de uma microbiota autóctone em reservatórios de petróleo (Magot et al., 2000). A controvérsia tem sido sustentada em parte por causa da difícil amostragem asséptica em reservatórios de petróleo profundos (Magot, 2005).

Deste modo, os micro-organismos observados em fluidos de campos petrolíferos poderiam ser contaminantes introduzidos durante as operações de perfuração e/ou durante a recuperação da amostra, ou de material despreendido do biofilme que cresce nas tubulações instaladas. Outra razão para o ceticismo é a prática comum de injeção de águas superficiais ou reinjeção das águas de formação naturais (para manter a pressão do reservatório para a produção de petróleo), o que levaria à introdução de micro-organismos durante a injeção e, portanto, não constituiriam espécies autóctones (Foght, 2010).

O reservatório de petróleo é considerado um ambiente extremo, onde as condições de alta temperatura, pressão, salinidade e anaerobiose também são consideradas como inóspitas para a atividade microbiana. Só até a recente descoberta de oxidação anaeróbia de hidrocarbonetos e com o aumento da consciência da capacidade dos micro-organismos para colonizar ambientes extremos, ficou claro que os micro-organismos poderiam sobreviver e alterar o petróleo nas condições anaeróbias dos reservatórios. Na atualidade, o uso de técnicas de amostragem e de cultivo mais sofisticadas e adequadas junto com a aplicação de técnicas de biologia molecular a campos petrolíferos tem levado a dissipar a teoria do subsolo profundo estéril (Foght, 2010). Pelo contrário, tem ficado claro que muitos reservatórios de petróleo abrigam micro-organismos autóctones, como os gêneros *Petrotoga* e *Geotoga* que são isolados exclusivamente de reservatórios de petróleo (Birkeland, 2004).

Em um estudo inicial, Bernard e colaboradores (1992) examinaram 66 amostras de água conata (ou seja, a água retida nos poros e fissuras da rocha desde a sua formação sedimentar ou vulcano-sedimentar) de 24 campos de reservatórios de petróleo geograficamente distantes. Os autores enumeraram os micro-organismos totais usando microscopia de epifluorescência e o número de células viáveis em cultivo usando seis meios aeróbicos e anaeróbicos ajustados às condições químicas *in situ* e temperatura. Eles descobriram que (1) o número de células diminuiu com o aumento da salinidade e temperatura, (2) os micro-organismos do reservatório são, em sua maioria, anaeróbios (redutores de sulfato, fermentadores e metanogênicos), geralmente de espécies e gêneros não descritos, e (3) campos petrolíferos geoquimicamente distintos abrigam comunidades específicas de micro-organismos, as quais variam de acordo com as condições do reservatório, possivelmente através de seleção natural ou da história de deposição (Head et al., 2010).

Mesmo sendo controversa a evidência de quais micro-organismos seriam realmente autóctones dos reservatórios de petróleo e quais produtos de contaminações, uma grande variedade de grupos taxonômicos microbianos tem sido identificada em reservatórios de petróleo geograficamente distantes usando técnicas de estudo convencionais adaptadas às condições *in situ* do reservatório, como descrito por L'Haridon et al. (1995), Grassia et al. (1996) e revisados por

Magot et al. (2000), ou estratégias moleculares combinadas com métodos dependentes de cultivo, como reportado por Orphan et al. (2000). Os grupos fisiológicos e espécies encontrados nesses estudos incluem bactérias redutoras de sulfato (*Deltaproteobacteria*, *Firmicutes* e *Thermodesulfobacterium*) e arqueias (*Archaeoglobus*), bactérias fermentadoras (*Firmicutes*, *Thermotogae*, *Synsagittes*, *Spirochaetes*) e as arqueias (*Thermococcus*, *Pyrococcus*), redutores de ferro (*Gamaproteobacteria*, *Bacteroidetes*, *Deferribacteres*, *Firmicutes*, *Thermotogae*, *Thermococcus*), redutores de nitrato (*Betaproteobacteria*, *Gamaproteobacteria*, *Epsilonproteobacteria*, *Deferribacteres*, *Firmicutes*), homoacetogênicas (*Firmicutes*) e arqueias metanogênicas (*Methanococcales*, *Methanomicrobiales*, *Methanobacteriales*, *Methanosarcinales*) (Head et al., 2010). O inventário de micro-organismos de campos petrolíferos continua crescendo (Ollivier e Magot 2005).

2.4. Degradação microbiológica do petróleo nos reservatórios

Durante muito tempo, o mecanismo considerado para a biodegradação do óleo nos reservatórios de petróleo foi o metabolismo microbiano aeróbio utilizando o oxigênio dissolvido carregado pelas águas meteóricas. Esse mecanismo foi amplamente aceito apesar do fato de que o oxigênio provavelmente seria consumido pela oxidação da matéria orgânica e a redução de minerais nos sedimentos superficiais, e, portanto, seria muito improvável que esse oxigênio alcançasse os reservatórios de petróleo profundos (Head et al., 2010).

Nos anos recentes, a capacidade evidenciada dos micro-organismos de degradar componentes do petróleo anaerobicamente e a detecção de metabólitos resultantes da degradação em ausência de oxigênio mostraram ser informações valiosas na determinação dos processos envolvidos na degradação do óleo nos reservatórios de petróleo. Por exemplo, os micro-organismos nas condições anaeróbicas poderiam utilizar uma grande variedade de aceptores finais de elétrons, incluindo a redução de ferro, de sulfato, de nitrato, fermentação e metanogênese (Röling et al., 2003), as quais têm sido demonstradas em sedimentos superficiais e em culturas puras ou enriquecimentos em laboratórios (Widdel e Rabus, 2001).

Entre os principais parâmetros físico-químicos que limitam a degradação nos reservatórios de petróleo, encontra-se a temperatura, pois tem sido repetidamente observado que a biodegradação não ocorre em reservatórios de petróleo com temperaturas *in situ* acima de 80 e 90 °C, e as maiores degradações ocorrem em reservatórios mais frescos e superficiais (Magot, 2005). A alta salinidade é outro fator potencialmente limitante, especialmente em combinação com a temperatura (Grassia et al., 1996). A pressão parece ser o fator menos limitante, exceto

pelo fato de que pode selecionar certos tipos fisiológicos e influenciar o pH das águas porosas, aumentando a dissolução de CO₂ (Magot et al., 2000; Foght, 2010).

Embora o conhecimento dos reservatórios de petróleo biodegradados tenha avançado consideravelmente nos últimos anos, a compreensão dos processos envolvidos, dos organismos *in situ* responsáveis e da análise quantitativa dos fatores que controlam a biodegradação em reservatórios de petróleo ainda está longe de ser completa (Head et al., 2010).

2.5. Degradação microbiológica de hidrocarbonetos

Micro-organismos com capacidade de degradar alcanos estão amplamente distribuídos no ambiente e têm sido isolados de locais contaminados ou não com hidrocarbonetos (Marchant et al., 2006). Os organismos superiores podem oxidar parcialmente alguns hidrocarbonetos, ao passo que micro-organismos podem degradar completamente hidrocarbonetos da maioria das classes, incluindo alcanos, alquenos, alquinos e aromáticos. Tal degradação pode ocorrer aerobicamente com oxigênio, ou anaerobicamente com nitrato, íon férrico, sulfato ou outros aceptores de elétrons (Widdel e Musat, 2010).

No Dominio Bacteria, quase todos os isolados degradadores pertencem às Classes alfa, beta, e gama Proteobacteria e à Ordem Actinomycetales (van Beilen et al., 2003; Coleman et al., 2006). No filo Firmicutes, os isolados de *Bacillus* e *Geobacillus*, e dentro do filo Deinococcus-Thermus, o gênero *Thermus* têm sido encontrados como degradadores de alcanos (Marchant et al., 2006). Os grupos, Flavobacteria e Sphingobacteria do filo Bacteroidetes-Chlorobi foram encontrados em ambientes contaminados com petróleo (van Beilen e Funholff et al., 2007). Dentre os eucariotos, leveduras, fungos e algumas algas são conhecidos por crescer na presença de alcanos (van Beilen et al. 2003). Estudos ambientais têm demonstrado a abundância de organismos degradadores de alcanos, inclusive, alguns isolados marinhos parecem ser alcanotróficos obrigatórios, como *Alcanivorax borkumensis* (van Beilen e Funholff et al., 2007). Estima-se em mais de 100 espécies pertencentes a 40 gêneros de bactérias, leveduras e fungos filamentosos capazes de utilizar um ou mais tipos de hidrocarbonetos (Magot, 2005).

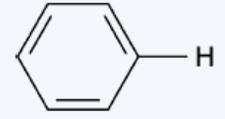
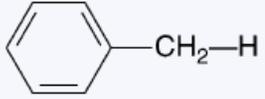
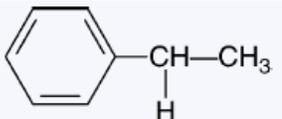
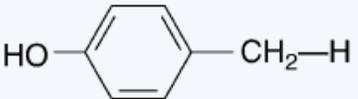
Por várias décadas esforços significativos têm sido dedicados ao estudo da biodegradação de poluentes orgânicos (Pieper et al., 2004). Sabe-se que tal degradação pode ocorrer tanto aeróbica como anaerobicamente. Portanto, qualquer tentativa de visão geral do metabolismo de hidrocarbonetos em micro-organismos não é apenas confrontada com a diversidade química de hidrocarbonetos e suas reatividades, mas também com vários estilos de vida microbiana. Geralmente, o estudo é tratado em áreas de estudo separadas, como por exemplo, degradação de hidrocarbonetos alifáticos vs. aromáticos, vias de degradação aeróbias vs. anaeróbias, fisiologia

geral e vias metabólicas vs. mecanismos enzimáticos e estruturas (Widdel e Musat, 2010). Mas apesar desse tratamento separado, cada uma dessas áreas lida com a mesma questão central que é o “desafio metabólico” que representa a natureza dos hidrocarbonetos, que são compostos apolares, desprovidos de grupos funcionais e, portanto, exibem baixa reatividade química à temperatura ambiente. Os hidrocarbonetos devem então ser ativados e, atualmente, é reconhecida a existência de uma surpreendente diversidade de reações de ativação que tem evoluído em micro-organismos (Widdel e Musat, 2010).

Os mecanismos de ativação de hidrocarbonetos são fundamentalmente diferentes em micro-organismos aeróbios e anaeróbios. Sob condições aeróbias, o metabolismo de hidrocarbonetos é sempre iniciado com o uso do oxigênio molecular como um co-substrato em reações de mono ou dioxigenases que permitem a hidroxilação terminal e sub-terminal das cadeias de alcanos alifáticos ou a mono ou di-hidroxilação dos anéis aromáticos (Boll e Heider, 2010). Na ativação anaeróbia de hidrocarbonetos, algumas reações conhecidas incluem: (1) hidroxilação por enzimas que contêm o cofator molibdênio, (2) adição de fumarato por enzimas glicil-radical, (3) metanogênese reversa envolvendo variantes da metil-coenzima M redutase, (4) metilação, (5) carboxilação catalisada por enzimas ainda não caracterizadas. O conhecimento destas enzimas varia grandemente, por exemplo, a enzima de hidroxilação do etilbenzeno já foi caracterizada estrutural e funcionalmente (Boll e Heider, 2010), ao passo que o mecanismo de ativação inicial no benzeno é ainda questionável e para outros permanece desconhecido (Tabela 1).

Vários aspectos da degradação de poluentes precisam ser estudados para obter um panorama detalhado dos processos de biodegradação no meio ambiente, a fim de otimizar e prever o desempenho de micro-organismos que realizam a degradação *in situ* (Pieper et al., 2004).

Tabela 1. Mecanismos de clivagem das ligações de C-H e tipos de enzimas e co-fatores envolvidos na ativação de hidrocarbonetos em bactérias anaeróbias. Reações e co-fatores com marcas de pontos de interrogação não foram identificados até o presente momento. Fonte: Boll e Heider, 2010.

Composto	Mecanismo de clivagem de C – H	Enzima	Co fator
 Benzeno	Carboxilação?	Desconhecido	Desconhecido
	Metilação?	Desconhecido	Desconhecido
CH₃-H Metano	Metanogênese reversa	Metil CoM redutase	F ₄₃₀ /Ni
	Adição de fumarato?	Enzima glicil-radical?	Glicil-radical / FeS?
CH₃-CH₂-CH₂-H Propano	Adição de fumarato	Enzima glicil-radical?	Glicil-radical / FeS?
 Tolueno	Adição de fumarato	Enzima glicil-radical	Glicil-radical / FeS?
 Etil-benzeno	Adição de fumarato	Enzima glicil-radical	Glicil-radical / FeS?
	Hidroxilação	Etil-benzeno desidrogenase	Co-fator Mo FeS heme b
 p – Cresol	Adição de fumarato	Enzima glicil-radical	Glicil-radical / FeS?
	Hidroxilação	p- Cresol metilhidroxilase	FAD heme c

2.6. Mecanismos e genética da degradação microbiológica de hidrocarbonetos

Os mecanismos enzimáticos envolvidos na degradação aeróbia de hidrocarbonetos por bactérias têm sido extensivamente estudados por várias décadas (Boll e Heider, 2010). Os genes que codificam as enzimas de degradação estão relativamente bem compreendidos para organismos aeróbios e cultiváveis, principalmente para um isolado de *Pseudomonas*, conhecido como *P. putida* GPo1, assim como para *Acinetobacter* sp. ADP1 e *Mycobacterium tuberculosis* H37Rv (van Beilen e Funhooff, 2007; Wentzel et al., 2007). No entanto, a degradação microbiológica anaeróbia de hidrocarbonetos vem ganhando cada vez mais atenção como o mecanismo que ocorre predominantemente nos reservatórios de petróleo, dadas as condições de anaerobiose que geralmente dominam estes ambientes, mas o seu estudo é uma área incipiente

em função das peculiaridades desse ambiente e dificuldades que surgem das tentativas de caracterizar essas comunidades (van Hamme et al., 2003) (Figura 3). Não obstante, várias bactérias derivadas de outros ambientes com capacidade de utilizar alcanos como fonte de carbono na ausência de oxigênio têm sido descritas nos últimos anos (Rojo, 2009).

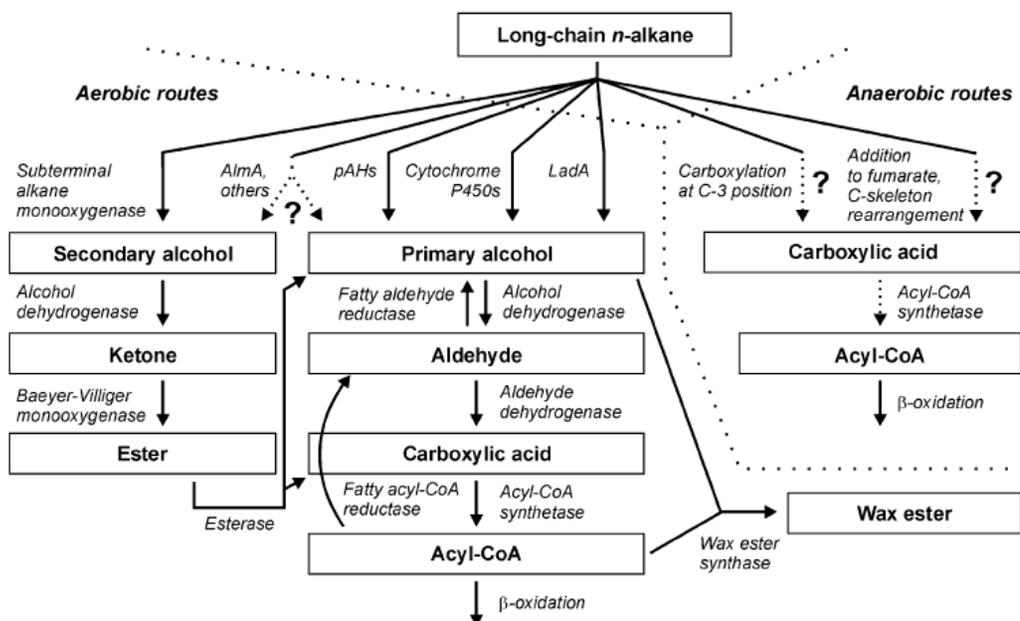


Figura 3. Vias de degradação de alcanos de cadeia longa em bactérias aeróbicas e anaeróbicas. Setas pontilhadas representam vias metabólicas sugeridas. Fonte: Wentzel et al. (2007).

2.6.1. Degradação aeróbia

As bactérias que degradam alcanos aerobicamente usam o oxigênio para a ativação da molécula de alcano e como aceptor final de elétrons (Rojo, 2009). As enzimas monooxigenases e alcano hidroxilases são as enzimas consideradas como as mais importantes no papel da biodegradação aeróbia.

2.6.1.1. Hidrocarbonetos alifáticos

Na maioria das vias de degradação descritas, o substrato *n*-alcano é oxidado ao correspondente álcool por uma enzima monooxigenase/hidroxilase terminal substrato-específica. O álcool é oxidado ao aldeído correspondente, e finalmente transformado em um ácido graxo (Figura 4). Os ácidos graxos são conjugados a CoA e posteriormente processados pela β – oxidação para gerar Acetil-CoA (van Hamme et al., 2003; Wentzel et al., 2007). A oxidação sub-terminal também tem sido descrita para alcanos de cadeia longa e curta (Wentzel et al., 2007) e, ainda, as

oxidações tanto terminais como sub-terminais podem coexistir em alguns organismos (Rojo, 2009) (Figura 4).

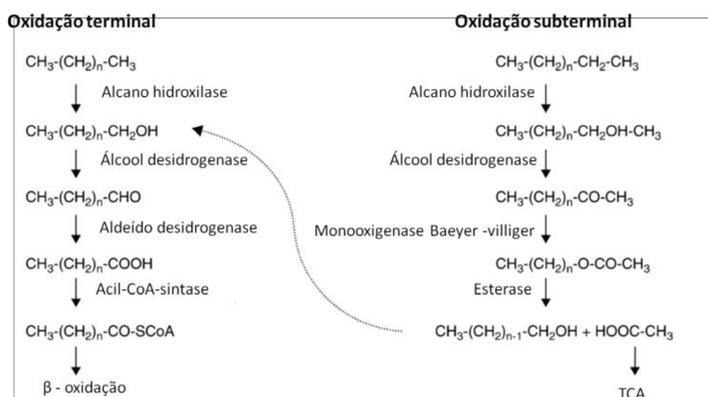


Figura 4. Vias mais freqüentes de oxidação de n-alcanos por oxidação terminal e subterminal. Fonte: Rojo (2010).

A oxidação terminal de n-alcanos pode ser realizada por enzimas de diferentes famílias (van Beilen e Funhoff, 2007; Rojo, 2010) (Tabela 2).

Tabela 2. Classes de enzimas que oxidam n-alcanos. A gama de substratos é aproximada; limites superiores e inferiores podem variar em diferentes cepas. Fonte: Rojo (2010).

Classe de enzimas	Características	Substrato	Hospedeiro
PRM, propano monooxigenase	Monooxigenase de ferro não heme	C ₃	Bactérias
sBMO, butano monooxigenase	Monooxigenase de ferro não heme	C ₂ -C ₉	Bactérias
pBMO butano monooxigenase	Monooxigenase contendo cobre	C ₂ -C ₉	Bactérias
CYP153	Citocromo P450 solúvel	C ₅ -C ₁₂	Bactérias
ALKB-relacionadas	Monooxigenase de ferro não heme	C ₃ -C ₁₃ ou C ₁₀ -C ₂₀	Bactérias
AlmA	Monooxigenase ligada a flavina	C ₂₀ -C ₃₆	Bactérias
LadA	Monooxigenase solúvel	C ₁₀ -C ₃₀	Bactérias
Dioxigenase	Dioxigenase de cobre dependente de flavina	C ₁₀ -C ₃₀	Bactérias

A degradação aeróbia de alcanos de cadeia mediana (C₅-C₁₁) e longa (>C₁₂) tem sido associada à presença da monooxigenase de ferro integral de membrana ALkB (AlkB *integral membrane non-haem iron monooxygenase*), que pertence à via de degradação melhor caracterizada codificada pelo plasmídeo OCT de *Pseudomonas putida* GPo1 (inicialmente classificada como *Pseudomonas oleovorans*). Neste sistema, o plasmídeo OCT contém dois operons: *alkBFGHJKL* e *alkST* (van Beilen et al., 1994). O primeiro operon codifica para dois

Geobacillus sp. isolado de reservatório de petróleo (Feng et al., 2007; Rojo, 2009), e a monooxigenase ligada a flavina (*flavin-binding monooxygenase* - AlmA), em *Acinetobacter venetianus* isolado de uma cultura mista com capacidade de crescer em alcanos de cadeia longa (Throne-Holst et al., 2006; Amouric et al., 2009). Atualmente, acredita-se que existem sistemas enzimáticos para degradação de alcanos que ainda não foram caracterizados e que poderiam incluir novas proteínas não relacionadas com as conhecidas atualmente (Rojo, 2009). Além disso, a propriedade observada em muitas bactérias degradadoras de alcanos de possuir múltiplos sistemas de oxidação dos mesmos, ou até mesmo de exibir sobreposição de substratos (van Beilen e Funhoff 2007; Wentzel et al., 2007), indica que para explorar e caracterizar a diversidade metabólica envolvida na degradação de alcanos deve ser levada em consideração a alta diversidade de enzimas capazes de iniciar tal metabolismo, assim como o uso de técnicas de acesso a sequências ainda não conhecidas em ecossistemas pouco estudados.

2.6.1.2. *Hidrocarbonetos aromáticos*

Em geral, o catabolismo bacteriano de compostos aromáticos envolve uma ampla variedade de vias periféricas que ativam substratos estruturalmente diversos em um número limitado de intermediários comuns que serão futuramente clivados e processados por poucas vias centrais ao catabolismo da célula (Carmona et al., 2009). O mecanismo que permite às bactérias degradar membros específicos da grande diversidade de compostos aromáticos tem sido estudado em várias cepas bacterianas isoladas, predominantemente pertencentes aos filos *Proteobacteria* e *Actinobacteria* (Brennerova et al., 2009). A degradação é tipicamente iniciada por um dos membros de três superfamílias: a oxigenase dependente de ferro em configuração não heme (*Rieske non haem iron oxygenases* - RNHO), a monooxigenase de flavoproteína (*flavoprotein monooxygenase* - FPM) e a monooxigenase di ferro solúvel (*Soluble diiron monooxygenase* - SDM) (Figura 6). O metabolismo posterior se dá com intermediários aromáticos di ou trihidroxilados. Alternativamente, a ativação é mediada por CoA ligases, onde os derivados formados são sujeitos a hidroxilações seletivas (clivagem não oxigenolítica do anel) (Perez-Pantoja et al., 2010; Vilchez-Vargas et al., 2010). No caso de poluentes como benzeno, tolueno, naftaleno, bifenila ou aromáticos policíclicos, a degradação aeróbia geralmente é iniciada pela ativação do anel aromático catalisada pelas enzimas RNHO ou, como intensivamente descrito para a degradação de tolueno, pelas enzimas SDM (Brennerova et al., 2009).

Apesar da diversidade relativamente ampla de mecanismos de ativação que é possível na degradação de aromáticos, essas vias geralmente convergem para a formação de poucos intermediários centrais di ou trihidroxilados. Estes são substratos para a clivagem do anel por duas classes de enzimas não relacionadas filogeneticamente, as intradiol e estradiol dioxigenases

Uma vez tendo tomado conhecimento sobre propriedades metabólicas das bactérias isoladas, a habilidade dos micro-organismos para mineralizar hidrocarbonetos aromáticos em solos tem sido monitorada (Junca e Pieper, 2004). Tipicamente, esses e outros estudos usam *primers* desenhados com base nas regiões conservadas e concentram-se em RNHO ou SDM como alvos para atividades de iniciação da degradação, ou sobre EXDO para a clivagem do anel aromático (Junca e Pieper, 2004). Esses estudos vão desde aqueles que procuram uma estreita gama de genes similares ou idênticos a aqueles observados nas cepas tipo usando *primers* não degenerados até aqueles que procuram subfamílias de genes homólogos usando *primers* degenerados. No entanto, devido à grande diversidade existente, nenhum par de *primers* realmente será capaz de cobrir a ampla diversidade de família de genes catabólicos na natureza (Vilchez-Vargas et al., 2010).

2.6.2. Degradação anaeróbia

Os micro-organismos e as reações aeróbias são claramente relevantes para a remineralização de contaminantes orgânicos em ambientes naturais, incluindo sítios contaminados, mas sua importância pode ter sido superestimada em certa medida. Em particular, em locais com altos teores de carbono orgânico, o oxigênio molecular é mais rapidamente consumido do que reabastecido por difusão. Assim, as condições anóxicas dominam muitos habitats naturais e sítios contaminados. Sob estas condições a remineralização de contaminantes orgânicos deve estar acoplada à respiração anaeróbia com uma variedade de aceptores de elétrons (Heider e Rabus, 2008; Rojo 2009).

Em geral, as reações sob condições aeróbicas têm sido estudadas por várias décadas, enquanto a degradação anaeróbica de hidrocarbonetos é considerada como uma capacidade microbiana recentemente descoberta que ainda aguarda um entendimento mais profundo (Heider e Rabus, 2008; McLeod et al., 2008; Carmona et al., 2009).

A oxidação anaeróbia de hidrocarbonetos tem sido acoplada a diferentes modos de conservação de energia como redução de nitrato, ferro, sulfato e manganês, e, mais recentemente, a redução de clorato (Mehboob et al., 2009), assim como a metanogênese, (Wentzel et al., 2007), fermentação e metabolismo fototrófico (Heider e Rabus, 2008). O crescimento desses micro-organismos e, portanto, a taxa de biodegradação é significativamente mais baixa em comparação com degradadores aeróbios (Wentzel et al., 2007). Como particularidade, as cepas analisadas normalmente utilizam uma faixa muito curta de alcanos como substratos (Wentzel et al., 2007; Rojo, 2009).

2.6.2.1. Hidrocarbonetos aromáticos

Para o metabolismo de compostos aromáticos já foram descritos nesta revisão os principais mecanismos para a degradação na presença de oxigênio, onde este não é só o aceptor final de elétrons como também o co-substrato de dois processos chaves, a hidroxilação e a clivagem do anel aromático por oxigenases. Em contraste, o catabolismo anaeróbico de compostos aromáticos usa uma estratégia completamente diferente, baseada em reações redutivas para atacar o anel aromático (Fuchs, 2008).

A bioquímica de algumas vias de degradação anaeróbica de compostos aromáticos tem sido estudada, no entanto, os determinantes genéticos dos processos envolvidos e a sua regulação permanecem pouco conhecidos (Carmona et al., 2009). Recentes avanços no sequenciamento de genomas levaram à completa informação genética de cinco cepas bacterianas com habilidade para degradar os compostos aromáticos usando diferentes aceptores de elétrons e pertencentes a diferentes grupos de bactérias, a saber: as betaproteobactérias desnitrificantes *Thauera aromática* e *Azoarcus* sp. EbN1 (proposta a ser renomeada como "*Aromatoleum aromaticum*" EbN1), as alfavroteobactérias, incluindo a fototrófica *Rhodospseudomonas palustris* CGA009 e a desnitrificante *Magnetospirillum magneticum* AMB-1, e duas deltaproteobactérias anaeróbias estritas, a ferro-redutora *Geobacter metallireducens* GS-15 e a fermentativa *Syntrophus aciditrophicus* SB (Carmona et al., 2009). Algumas das características destas bactérias encontram-se descritas na Tabela 3. A maioria das bactérias reportadas pertence ao novo cluster filogenético de anaeróbios facultativos dentro de betaproteobactérias desnitrificantes ou de deltaproteobactérias sulfato-ferro redutoras. As mais intensamente estudadas, por seu considerável rápido crescimento e rendimento no laboratório, são as espécies desnitrificantes *T. aromática* K172 e *A. aromaticum* EbN1. O novo gênero proposto por Wöhlbrand et al. (2007), "*Aromatoleum*", compreende desnitrificantes de vida livre com notáveis capacidades de degradação, distintos dos gêneros relacionados *Thauera* e *Azoarcus*, estes últimos endófitos fixadores de nitrogênio em arroz e outras plantas (Heider e Rabus, 2008) (Figura 7). Vale lembrar, que importantes generalizações têm sido feitas nos anos recentes sobre os genes envolvidos no metabolismo de hidrocarbonetos a partir das descrições feitas com estas bactérias que puderam ser isoladas sob condições de laboratório padronizadas. Entretanto, no caso de bactérias não cultivadas, novos genes, enzimas e vias metabólicas não descritas provavelmente são responsáveis pelos processos de degradação de hidrocarbonetos.

Tabela 3. Comparação de características relevantes em genomas bacterianos completamente sequenciados de micro-organismos degradadores anaeróbios.
Fonte: Carmona et al., 2009.

Organismo	Classificação taxonómica	Habitat	Principal acceptor de eletróns	T (°C) ótica crescimento	Algumas fontes aromáticas de carbono (anaerobiose)	Tam. genoma (pb)	Tam. plasmídeo (Kb)	Conteúdo G + C (%)	ORFs preditas
<i>R. palustris</i> CGA009	Alfaproteobacteria	Aquático	O ₂ fototrófica	25-30	p-cumarato, cinamato, ácido ferúlico, cafeato, 3 - fenilpropionato, 5 - fenilvalerato, mandelato, 4 - hidroxibenzoato, 3 - hidroxibenzoato, benzoato	5.549,213	8,4	65	4.836
<i>Magnetospirillum magneticum</i> AMB-1	Alfaproteobacteria	Aquático	O ₂ (microaerófilo), NO ₃ ⁻	30	Benzoato, fenilacetato	4.967,148		65,1	4.559
<i>Aromatoleum aromaticum</i> EbN1	Betaproteobacteria	Sedimento água doce	O ₂ , NO ₃ ⁻ , NO ₂ ⁻	31	Tolueno, etilbenzeno, fenilacetato, antranilato, fenol, p-cresol, petilfenol, fenilalanina, álcool benzílico, 3 - hidroxibenzoato, 4 - hidroxibenzoato, benzoato	4.296,230	207/223	65,1	4.603
<i>G. metallireducens</i> GS-15	Deltaproteobacteria	Múltiplo	Fe (III), Mn (IV), NO ₃ ⁻ , NO ₂ ⁻ , Co (III), U (VI), Cr (VI)	30	Tolueno, fenol, p-cresol, benzila, benzaldeído, álcool, 4 - hidroxibenzoato, benzoato	3.997,420	13,7	59,5	3.532
<i>S. aciditrophicus</i> SB	Deltaproteobacteria	Múltiplo	Fermentador	35	Benzoato	3.179,300		51,4	3.169

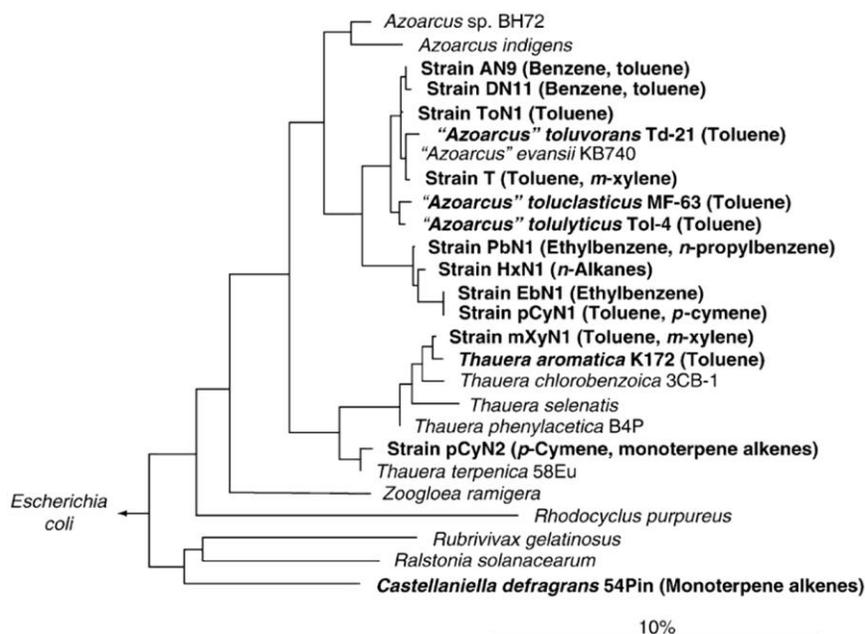


Figura 7. Betaproteobacterias degradadoras anaeróbicas de hidrocarbonetos descritas até o momento. Várias estão classificadas como espécies de *Azoarcus*, no entanto, elas diferem metabolicamente das espécies endofíticas “reais” de *Azoarcus*, e constituem um grupo separado (embora intimamente relacionado), sugerido ser renomeado como espécies do novo gênero *Aromatoleum* (Wöhlbrand et al., 2007). Vide texto para mais detalhes.

O mecanismo de ativação melhor compreendido e, aparentemente, mais comum dentre os mecanismos anaeróbicos é a adição de fumarato a hidrocarbonetos, produzindo adutos de succinato. Essa reação tem sido reconhecida para a ativação de vários alquil-benzenos, bem como para *n*-alcanos (Kube et al., 2004). No entanto, o entendimento dessa ativação de hidrocarbonetos dependente de fumarato é mais avançado apenas para o caso do tolueno (Figura 8). A enzima chave nesse processo é a enzima benzilsuccinato sintase. Todas as enzimas requeridas para a β oxidação de benzilsuccinato são codificadas pelo operon *bbs*. A subsequente degradação do benzoilCoA procede via desaromatização redutiva, clivagem hidrolítica do anel, β oxidação a unidades de acetil-CoA e oxidação terminal a CO_2 (Boll et al., 2002).

Bactérias como *A. aromaticum* EbN1 têm mostrado habilidade de utilizar o etilbenzeno como substrato hidrocarboneto alternativo. A via de degradação do etilbenzeno difere totalmente da degradação do tolueno, apesar da similaridade estrutural entre os dois compostos. A organização dos genes que codificam esta via também se encontra em um cluster diferente nesta bactéria. O etilbenzeno é hidroxilado e desidrogenado a acetofenona, a qual é posteriormente carboxilada e convertida em benzoil-CoA, intermediário comum entre as duas vias (Kube et al., 2004).

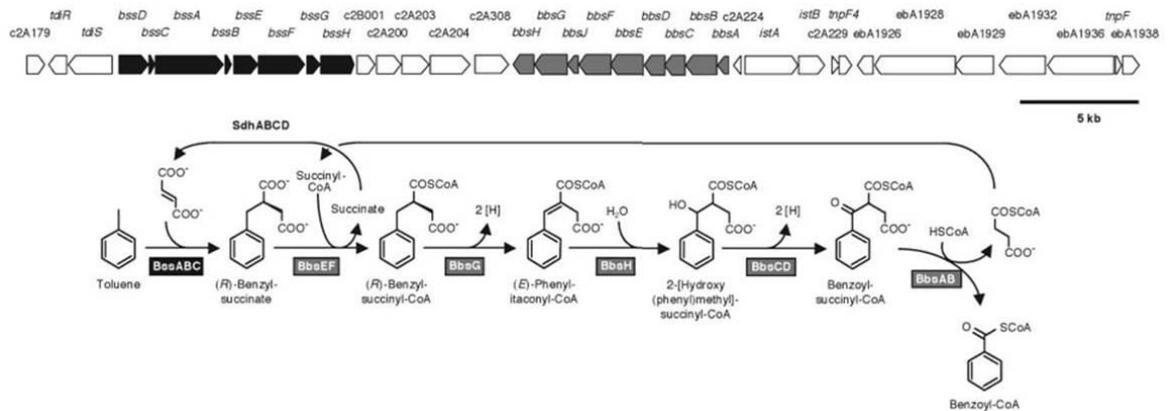


Figura 8. Via de degradação anaeróbica do tolueno e o cluster de genes que a codifica em *Aromatoleum aromaticum* EbN1. BssABC: Benzilsuccinato sintase; BbsEF: succinil-CoA-benzilsuccinatoCoA transferase; BbsG: benzilsuccinil-CoA desidrogenase; BbsH: feniltacanoil-CoA hidratase; BbsCD: 2 [hidroxi(pre)nil]metil]-succinil-CoA desidrogenase; BbsAB, benzoilsuccinil-CoA tiolase; SdhABCD: succinato desidrogenase. Fonte: Heider e Rabus (2008).

2.6.2.2. Hidrocarbonetos alifáticos

O metabolismo anaeróbico de alcanos até agora tem sido investigado em menor grau que o metabolismo de aromáticos, como alquil-benzenos, tolueno e etilbenzeno. A razão é presumivelmente porque o crescimento anaeróbico com n-alcanos é ainda mais lento do que com o alquilbenzeno. Além disso, o fato dos alcanos serem compostos muito hidrofóbicos e menos solúveis em água em comparação ao tolueno ou etilbenzeno, muitas vezes impede o cultivo de células de forma homogênea no meio (Widdel e Grundmann, 2010). Várias bactérias anaeróbicas capazes de degradar n-alcanos com 6 ou mais carbonos de comprimento, em particular o hexadecano (C₁₆), usando sulfato ou nitrato como aceptores de elétrons, têm sido isoladas (Grossi et al., 2008; Widdel e Grundmann, 2010).

Dois mecanismos principais de degradação de n-alcanos sob condições anaeróbicas têm sido descritos (Figura 9). O primeiro envolve a ativação por adição de fumarato como co-substrato para a formação de alquil-succinatos, em analogia à formação de benzil-succinato durante a degradação anaeróbica do tolueno. No entanto, as reações subsequentes são completamente diferentes, envolvendo reações de desidrogenação e hidratação (Widdel e Grundmann, 2010). E o segundo é a carboxilação, sugerida na degradação de hexadecano pela cepa sulfato redutora Hxd3 (So et al., 2003), tentativamente denominada como *Desulfococcus oleovorans*. Essa cepa difere das outras alcano degradadoras por converter alcanos de cadeias pares em ácidos graxos de cadeias ímpares em suas membranas lipídicas, enquanto os degradadores de alcanos conhecidos

(que provavelmente usam a adição de fumarato como reação inicial) formam ácidos graxos de cadeia ímpar a partir de alcanos de cadeia ímpar e ácidos graxos de cadeia par a partir de alcanos de cadeia par (Heider e Rabus, 2008).

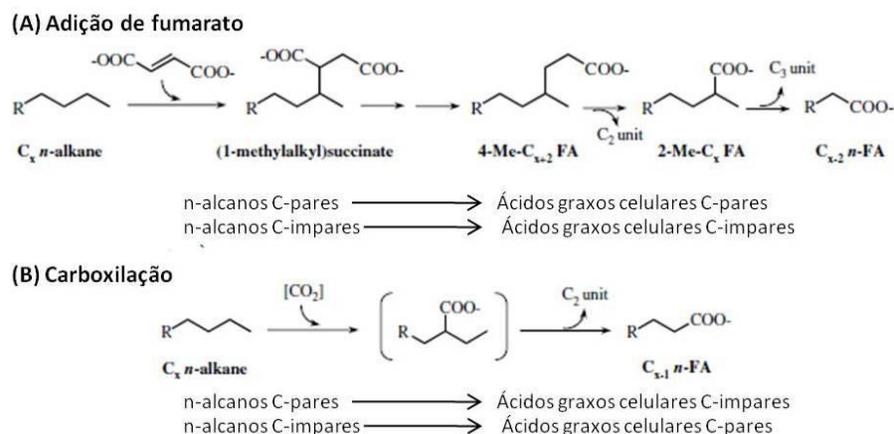


Figura 9. Mecanismos de degradação de n-alcanos em bactérias anaeróbicas e seu impacto sobre os principais ácidos graxos celulares. (A) e (B) mecanismos conhecidos; (C) mecanismo desconhecido. Adaptação de: Grossi et al., 2008.

2.7. Métodos independentes de cultivo

Uma das razões do incompleto entendimento das vias de degradação microbianas consiste na dificuldade em recuperar as comunidades microbianas em ambientes associados ao petróleo. Outra limitação para esse entendimento tem sido o contínuo isolamento das mesmas espécies quando se tenta estudar os micro-organismos relacionados com a degradação (Eyers et al., 2004). Essa limitação tem sido imposta pelas técnicas baseadas em cultivo, as quais foram por muitas décadas tradicionalmente as ferramentas primárias usadas para o estudo da microbiologia de diferentes ambientes (Chandler et al., 1997). Esta metodologia, apesar de ser extremamente importante para o entendimento do potencial fisiológico de micro-organismos isolados, não fornecem necessariamente informações abrangentes sobre a diversidade e abundância de comunidades microbianas presentes em amostras ambientais (Rölling et al., 2003; van Hamme et al., 2003).

As técnicas baseadas em cultivo ofereceram informações valiosas das interações microbianas com hidrocarbonetos no ambiente, assim como permitiram a recuperação de uma ampla coleção de bactérias capazes de utilizar hidrocarbonetos e a obtenção de operons que codificam essas vias de degradação. No entanto, na atualidade, sabe-se que apenas uma pequena fração da diversidade microbiana na natureza, em torno de 1 a 10%, pode ser cultivada em laboratório

(Torsvik et al., 1990; Amann et al., 1995; Torsvik et al., 1998). Apesar dos esforços em melhorar as técnicas de cultivo para recuperar os micro-organismos ainda não cultivados, o conhecimento das condições de crescimento na natureza (por exemplo, química do ambiente original, a vida em comunidades complexas, as interações com outros organismos) permanece insuficiente para cultivar a maioria deles (Eyers et al., 2004), e portanto, as funções ecológicas e possíveis aplicações em biotecnologia desses micro-organismos continuam obscuras (Kellenberger, 2001).

Poderosas ferramentas moleculares têm sido desenvolvidas nas últimas décadas, otimizadas e adaptadas para superar as limitações impostas pela abordagem clássica (isolamento e cultivo de micro-organismos), explorando de forma mais precisa a diversidade taxonômica e funcional dos micro-organismos ambientais (Eyers et al., 2004), e permitindo a descoberta de novos grupos de organismos nunca antes cultivados (Giovannoni et al., 1990; Amann et al., 1995; Hugenholtz et al., 1998).

Ao longo da última década, muita atenção tem sido dada aos habitats microbianos associados a estratos subterrâneos ricos em petróleo, devido à sua significância científica e às aplicações industriais, tais como recuperação de ambientes contaminados com petróleo (Van Hamme et al., 2003; Aitken et al., 2004; Liu et al., 2005). O uso de metodologias independentes de cultivo para investigar comunidades microbianas associadas aos ambientes petrolíferos vem sendo disseminado nos anos recentes (Head et al., 2006) e vem permitindo uma maior e mais completa caracterização da microbiologia de reservatórios de altas e baixas temperaturas (Li et al., 2007).

2.8. Metagenômica

O termo metagenômica, também conhecido como genômica de comunidades, ecogenômica ou genômica ambiental, foi introduzido em 1998 por Handelsman e colaboradores. Este campo de pesquisa relativamente novo permite o estudo genômico e acesso ao potencial metabólico de organismos que não são cultivados ainda no laboratório em misturas com outros micro-organismos. Chen e Patcher (2005) definiram metagenômica como a aplicação de técnicas modernas de genômica ao estudo de comunidades de micro-organismos diretamente em seus ambientes naturais, ignorando a necessidade de isolamento e cultivo das espécies individuais no laboratório. Apesar dos diferentes nomes utilizados para descrever os mesmos métodos, o termo “metagenômica” foi aprovado pela revista científica *Environmental Microbiology*. A definição exclui o uso de *primers* para amplificar cassetes de genes ou *primers* aleatórios para acessar genes de interesse (Handelsman, 2004; Riesenfeld et al., 2004; Handelsman, 2007). O propósito da metagenômica inclui (1) entendimento dos processos ecológicos (por exemplo, reações catabólicas) atuando em um ecossistema de interesse e (2) busca de novos genes a partir de

organismos não cultivados que poderiam ser biotecnologicamente importantes (Watanabe e Kassai, 2008).

Muitos ambientes têm sido foco da metagenômica, incluindo solo (Fierer et al., 2007), intestino distal em humanos (Gill et al., 2006), fezes humanas (Zhang et al. 2006) e habitats aquáticos (Venter et al., 2004). O acesso aos genes dos micro-organismos ainda não cultivados representa um grande potencial para a descoberta de novos produtos gênicos como metabólitos secundários com atividade biológica ou vias de degradação de xenobióticos. Isso foi claramente ilustrado por Venter e colaboradores (2004), que, usando abordagens moleculares independentes de cultivo, encontraram milhares de novas espécies bacterianas e mais de um milhão de novos genes codificadores de proteínas em 200 L de água do mar Sargasso (Eyers et al., 2004).

Na última década, a metagenômica tem se convertido numa das ferramentas mais poderosas para a detecção e identificação de novos genes, produtos, vias metabólicas, metabólitos secundários, além da sua utilidade na identificação da composição filogenética e funcional das comunidades microbianas nos diversos ecossistemas (Handelsman, 2004; Reisenfeld et al., 2004; Steele e Streit, 2005; Daniel, 2005).

No caso de comunidades microbianas biodegradadoras, as bibliotecas metagenômicas constituem-se como uma excelente ferramenta para a descoberta de novas vias catabólicas, principalmente de micro-organismos fastidiosos ou ainda não-cultivados, ou mesmo daqueles que vivem em consórcios e não podem ser recuperados isoladamente. Ainda, a partir de uma biblioteca de metagenoma, uma variedade maior de atividades catabólicas pode ser obtida simultaneamente, em comparação ao método tradicional baseado em isolamento, cultivo e triagem de linhagens puras de micro-organismos. Os estudos têm mostrado que os genes responsáveis pela biodegradação obtidos a partir de metagenoma de solo são diferentes daqueles acessados através do isolamento e cultivo dos micro-organismos (Margesin et al., 2003).

A estratégia de construção da biblioteca metagenômica depende fortemente do objetivo principal do estudo. A decisão do tamanho dos insertos a serem clonados depende se são os genes individuais ou seus produtos o alvo do estudo, ou se o objetivo é detectar operons inteiros e *clusters* de genes codificando vias biossintéticas ou degradadoras (Schmeisser et al., 2007). Bibliotecas metagenômicas têm sido geradas usando insertos pequenos (2-15 kb) clonados em vetores do tipo plasmídeos (Li et al., 2008), insertos maiores (15 – 40 kb) clonados em vetores do tipo cosmídeos (Voget et al., 2003) ou fosmídeos (Hardeman e Sjöling, 2007) ou insertos bem grandes (maiores de 100 kb) clonados em cromossomos bacterianos artificiais – BAC (Wu e Sun, 2009). As duas últimas estratégias permitem a clonagem de vias catabólicas ou biossintéticas inteiras, e oferecem uma alternativa para a exploração do potencial metabólico de micro-organismos que não são recuperados por métodos baseados em cultivo. Quanto à construção de

bibliotecas de insertos pequenos, eles são mais apropriados para a geração de grandes quantidades de sequências de DNA e obtenção de uma varredura da diversidade microbiana no ambiente, mas não são muito úteis nas pesquisas funcionais propriamente ditas (Venter et al., 2004; Schmeisser et al., 2007)

As bibliotecas metagenômicas se destacam como um dos importantes avanços em matéria de comunidades microbianas, contribuindo para melhorar a perspectiva filogenética e funcional sobre essas comunidades, além de permitir a exploração tecnológica de seu potencial metabólico. No caso dos ambientes petrolíferos, a metagenômica permitirá o acesso a novas sequências, genes, rotas metabólicas completas e seus produtos (Hardeman e Sjöling, 2007). Isto permitirá ampliar o conhecimento sobre a diversidade das vias de degradação utilizadas por microorganismos nestes ambientes, que permanece ainda recente e escasso. Além disso, a prospecção metagenômica de genes catabólicos da microbiota presente nestes ambientes considerados como extremos apresenta-se como excelente ferramenta na exploração de novas formações petrolíferas, onde o conhecimento da microbiologia do reservatório permitirá uma predição do grau de biodegradação do mesmo, incluindo preferências microbianas à degradação de determinados hidrocarbonetos, rotas metabólicas e condições nas quais as mesmas ocorrem num reservatório determinado (Magot et al., 2000).

2.9. Ferramentas para análise de sequências metagenômicas

A análise de metagenomas geralmente envolve os seguintes passos: (1) isolamento de DNA de alto peso a partir da amostra ambiental, (2) clonagem dos fragmentos de DNA a um vetor apropriado, (3) transformação de uma bactéria hospedeira e (4) análise ou *screening* dos clones transformantes através de duas estratégias principais: a primeira baseada na expressão de alguma atividade biológica de interesse (*function-based screening*), e a segunda baseada na detecção de sequências nucleotídicas (*sequence-based screening*) (Handelsman, 2004). Os clones obtidos também podem ser sujeitos ao sequenciamento por *shotgun* ou por plataformas de sequenciamento de segunda geração, como o pirosequenciamento – 454 GS FLX (Roche) ou *Genome Analyser* (Illumina) (Figura 10).

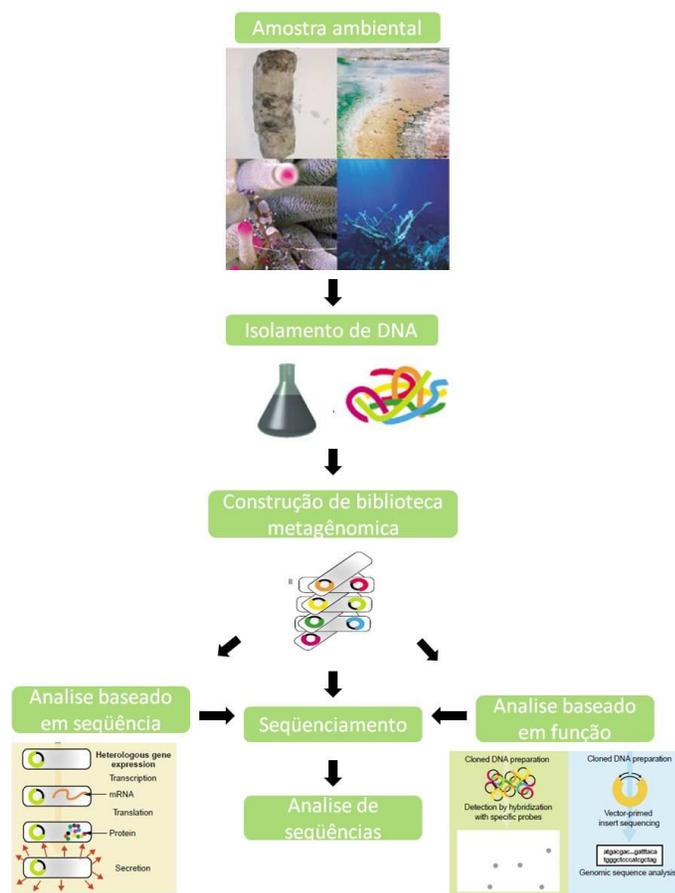


Figura 10. Diferentes abordagens para análise de metagenomas. Adaptado de Schloss e Handelsman (2003).

Mais de 130 projetos de sequenciamento de metagenomas usando a plataforma de 454 estão atualmente em andamento ou finalizados (Liolios et al., 2008). Dada a enorme quantidade de dados de seqüências geradas, esta última abordagem representa um desafio metodológico e computacional (Li, 2009), e vários métodos e recursos têm sido desenvolvidos para responder a este desafio, através de conjuntos de dados simulados como IMG/M (Markowitz et al., 2008), CAMERA (Seshadri et al., 2007), MG-RAST (Aziz et al., 2008), com ferramentas taxonômicas como MEGAN (Huson et al., 2007), com comparações estatísticas (Rodriguez-Brito et al., 2006), com análises de diversidade funcional (Schloss e Handelsman, 2008) entre outros (Li, 2009).

Progressos recentes revelaram que a captura de recursos genéticos de comunidades microbianas complexas em bibliotecas metagenômicas permite a descoberta de uma diversidade genética nova que não tinha sido previamente imaginada (Ferrer et al., 2005). No entanto, somente alguns estudos claramente tentaram identificar genes catabólicos diretamente de DNA ambiental pela abordagem metagenômica. Brennerova e colegas (2009) usaram a coloração amarela dos produtos da clivagem do anel aromático pela atividade da enzima catecol dioxigenase

como triagem funcional em uma biblioteca de ambiente contaminado de benzeno, tolueno e xilenos (BETX). Os autores encontraram uma alta abundância de genes catecol dioxigenase e, interessante, apenas um quarto das enzimas estradiol dioxigenases observadas pertenciam às famílias consideradas como predominantes pelo conhecimento obtido com micro-organismos isolados.

2.9.1. Análises de bibliotecas metagenômicas baseada em similaridade de sequência

Nesta abordagem os clones da biblioteca são selecionados por métodos de amplificação por PCR ou hibridização usando *primers* e sondas derivadas de sequências específicas de DNA de regiões conservadas de genes conhecidos baseando-se na informação contida nas bases de dados (Handelsman, 2004; Daniel, 2005). A vantagem desta estratégia é a independência da expressão na célula hospedeira da biblioteca (Lorenz et al., 2002). As desvantagens relacionam-se à detecção de genes incompletos e à pouca praticidade em bibliotecas grandes, além disso, só aqueles genes com regiões similares às sequências de sondas e *primers* conhecidos podem ser recuperados e, portanto, não é provável identificar novos genes (Simon e Daniel, 2009). No entanto, esta estratégia tem permitido a detecção de novas enzimas funcionais, como álcool oxidorreductase (Knietsch et al., 2003b) e quitinase (LeCleir et al., 2004), e como vantagem permite o uso de âncoras filogenéticas para análise da biodiversidade (Sjöling et al., 2007) (Tabela 4). Um estudo recente baseado em similaridade de sequência foi publicado por Jogler e colaboradores (2009). Neste estudo uma biblioteca metagenômica de 5.823 clones foi construída a partir de um enriquecimento de bactérias magnetotáticas (MTB) e avaliada por hibridização com sondas para genes *mam* de alfaproteobactérias magnetotáticas conhecidas. A similaridade com magnetossomos de MTB cultiváveis foi detectada e a organização desse magnetossomo na MTB não cultivada foi elucidada.

2.9.2. Análise de bibliotecas metagenômicas baseada em função

A abordagem baseada em função não é dependente da informação da sequência ou da similaridade de sequência com genes conhecidos. Assim, essa é uma abordagem que leva à descoberta de novas classes de genes que codificam funções conhecidas ou novas (Rees et al., 2003). Uma limitação significativa desta técnica é a dependência da expressão de genes e produtos de genes funcionais na célula hospedeira, que na maioria dos estudos é *Escherichia coli* (Simon e Daniel, 2009). Além disso, essa abordagem requer freqüentemente a análise de mais clones que o método baseado em sequência para a recuperação de poucos clones positivos

(Daniel, 2005). No entanto, essa parece ser a melhor opção para procurar novas proteínas que não compartilham nenhuma homologia com proteínas conhecidas. Isto é de especial interesse, pois se sugere que cerca de 3×10^{30} enzimas em nível de espécie, distribuídas entre 10.500 – 15.000 famílias de proteínas, esperam para serem descobertas (Beloqui et al., 2008). Neste aspecto, potenciais processos industriais podem ser muito mais diversificados do que se pode imaginar, contando todas as famílias de proteínas existentes e, portanto, abrindo a perspectiva da descoberta de milhares de diferentes enzimas capazes de realizar as biotransformações industriais de miríades de diferentes substratos no produto desejado (Guazzaroni et al., 2010).

Através desta abordagem, a identificação de funções enzimáticas de clones individuais é realizada por meio de corantes, cromóforos ou fluoróforos como substratos que podem ser incorporados ao meio de crescimento (Daniel, 2005). Exemplos destas análises incluem a detecção de atividade lipolítica empregando a tributirina no ágar (Hardeman e Sjöling 2007) e a detecção de proteases e celulasas através da formação de halos em meio indicador solidificado (Simon e Daniel, 2009) (Tabela 4).

Acredita-se que a abordagem funcional é a mais apropriada para encontrar novos genes catabólicos para o processo de degradação de xenobióticos (Eyers et al., 2004). Como já foi ressaltado acima, novos mecanismos metabólicos e enzimáticos envolvidos na degradação de poluentes ainda estão sendo descobertos, e mesmo abordagens baseadas em similaridade de sequência com base na diversidade metabólica e mecanismos descritos estão revelando uma maior diversidade de enzimas do que se pensava anteriormente (Vilchez-Vargas et al., 2010).

Tabela 4. Alguns exemplos recentes de biocatalistas derivados de metagenomas e a estratégia de *screening* empregada.

Alvo	Fonte	Método de <i>screening</i>	Referencia
Lipases	Sedimento do mar Báltico	Funcional	Hardeman e Sjöling 2007
	Solo de campo não cultivado	Funcional	Voget et al., 2003
Esterases	Wadi Natrun (Egito), lago Nakuru e cratera (Quênia)	Funcional	Rees et al., 2003
	Solo	Funcional	Li et al., 2008
Celulasas	Solo	Funcional	Voget et al., 2006
	Wadi Natrun (Egito), lago Nakuru e cratera (Quênia)	Funcional	Rees et al., 2003
Quitinases	Diversos habitats aquáticos	Funcional e sequência	LeCleur et al., 2004
Mono e Dioxigenases	Aqüífero leste do rio Snake	Sequência	Erwin et al., 2005
Polifenoloxidase	Rúmen de vaca	Funcional	Beloqui et al., 2006

Glicerol desidratase	Solos de campo de açúcar (Alemanha)	Sequência	Knietsch et al., 2003a
Álcool e aldeído oxidorreductase	Solos de campo de beterraba (Alemanha)	Funcional e sequência	Knietsch et al., 2003b
Carboxipeptidases	Intestino de inseto praga do milho <i>Helicoverpa armigera</i>	Sequência	Bown et al., 2004
Resistência a antibióticos	Placa dental e saliva humana	Sequência	Diaz-Torres et al., 2003

2.10. Detecção de novas enzimas com papel na biodegradação

As enzimas monooxigenases e alceno hidroxilases são as enzimas mais estudadas pelo seu papel na biodegradação do petróleo, hidrocarbonetos clorados, aditivos de combustível e muitos outros compostos. Essas enzimas vêm recebendo cada vez mais atenção como marcadores para prever o potencial de diferentes ambientes à degradação de petróleo (van Beilen e Funhoff, 2005), e, ao mesmo tempo, por serem consideradas como biocatalisadores versáteis, que realizam uma vasta gama de reações de oxidação (Smits et al., 2002; van Hellemond et al., 2007). No entanto, a aplicação industrial destas enzimas é restrita devido à complexa bioquímica, desafiando as exigências do processo, bem como o número limitado de enzimas clonadas e expressas. A AlkB, por exemplo, é uma proteína de membrana que requer rubredoxina e rubredoxina redutase para sua atividade (Feng, et al., 2007).

Abordagens de engenharia genética e evolutiva começaram a produzir sistemas mais robustos e versáteis, ampliando o portfólio de oxigenases de alcanos. Além disso, como já foi dito, a abordagem metagenômica fornece o acesso a novas sequências e sistemas de oxidação não descritos até agora. Entretanto, as enzimas derivadas de ambientes considerados como extremos oferecem maiores vantagens biotecnológicas. Os reservatórios de petróleo constituem, em geral, um habitat natural de micro-organismos extremófilos, os quais suportam elevadas temperaturas, pressão, salinidade, anóxia, escassez de nutrientes, oscilações de pH, presença de compostos tóxicos, entre outras condições adversas. Assim, micro-organismos capazes de sobreviver nestes ambientes constituem arsenais enzimáticos únicos, resistentes a condições extremas, com elevado potencial industrial e ambiental. Os micro-organismos extremófilos degradadores de hidrocarbonetos são candidatos ideais para o tratamento biológico de locais poluídos (Margesin e Schinner, 2001). Além disso, as vias de degradação também são importantes como fontes de novas atividades catalíticas úteis em química verde e biotecnologia branca (Galvão et al., 2005). Portanto, o potencial biocatalítico de linhagens microbianas derivadas de ambientes extremos vem ganhando crescente interesse, sendo que a identificação de novos compostos e rotas metabólicas vêm estimulando estudos metagenômicos destes nichos. O exemplo mais conhecido

de extremoenzima atualmente comercializada inclui a DNA polimerase obtida a partir de *Thermus aquaticus*, de essencial aplicação em biologia molecular nas reações em cadeia da polimerase (PCR) (Satyanarayama et al., 2005).

2.11. Processamento de sequências metagenômicas

O processamento das sequências genômicas e o das metagenômicas apresentam várias etapas em comum: o pré-processamento de leituras, a montagem e a predição de genes e a anotação. Alguns detalhes destes procedimentos estão descritos a seguir, com especial ênfase em dados de sequências metagenômicas provenientes do sequenciamento pelo método de Sanger.

2.11.1. Pré-processamento de sequências (leituras)

O pré-processamento compreende a chamada das bases (*base calling*) dos dados brutos que saem das máquinas de sequenciamento, remoção da sequência do vetor de clonagem, a imposição de qualidade corte para remover bases de baixa qualidade e a análise de contaminantes para remover sequências contaminantes. O pré-processamento das leituras antes da montagem, predição e anotação de genes é um aspecto crítico da análise, cujos possíveis erros poderiam causar maiores consequências na análise subsequente em metagenomas do que em genomas (Kunin et al., 2008).

O programa predominantemente usado para o *base calling* é o Phred (Ewing et al., 1998; Ewing e Green, 1998) o qual dá uma pontuação de qualidade para cada base como estimativa da sua probabilidade de erro. A pontuação de qualidade (q) atribuída a uma base está relacionada com a probabilidade (p) da base ser chamada erroneamente, seguindo a fórmula: $q = -10 \times \log_{10}(p)$. Assim, uma pontuação de qualidade phred de 20 corresponde a uma probabilidade de erro de 1% (Ewing e Green, 1998).

A trimagem do vetor é o processo de remoção de sequências de vetor de clonagem das bases nucleotídicas chamadas. A remoção completa e precisa das sequências de vetor de clonagem é especialmente importante em conjuntos de dados com regiões de cobertura baixa, sendo que os dados não trimados podem produzir contigs quiméricos onde as sequências do vetor, sendo as mais comuns em relação à maioria das leituras, atuam para reunir sequências independentes. Além disso, a predição de genes do vetor pode introduzir genes fantasma nas análises subsequentes. A baixa qualidade em sequências inteiras ou em regiões das leituras deve ser removida, caso contrário os dados de baixa qualidade serão indistinguíveis para o usuário médio e podem passar para a etapa de predição de genes. No pior cenário, esses genes fantasmas

em regiões de baixa qualidade podem passar sem controle em bases de dados públicas (Figura 11). O reconhecimento de outras sequências contaminantes como de *E. coli* é importante especialmente em dados procedentes do sequenciamento com Sanger, dado que *E. coli* é o organismo hospedeiro da clonagem e pequenas quantidades de seu genoma pode ter passado na purificação do plasmídeo.

Existem numerosas ferramentas para cada uma destas etapas, por exemplo, na triagem do vetor o mais comumente empregado é o *cross_match* (www.prahp.org). No entanto, esse programa falha em seu desempenho ao remover o vetor, dado que exige correspondência exata com o vetor e não possui nenhuma expectativa da localização do vetor na leitura (Kunin et al., 2008). Outra ferramenta desenvolvida, o programa LUCY (Chou e Holmes 2001), combina a triagem do vetor e de qualidade em um mesmo programa, e tem mostrado resultados substancialmente melhores que os encontrados com *cross_match* (Kunin et al., 2008). Recentemente, o programa *BD trimmer* foi desenvolvido e usado com sucesso nas etapas de pré-processamento, removendo sequências de baixa qualidade, contaminantes provenientes de *E. coli* e sequências de vetores contaminantes, usando UniVec como base de dados referência para eliminar os contaminantes (Mondego et al., 2011).

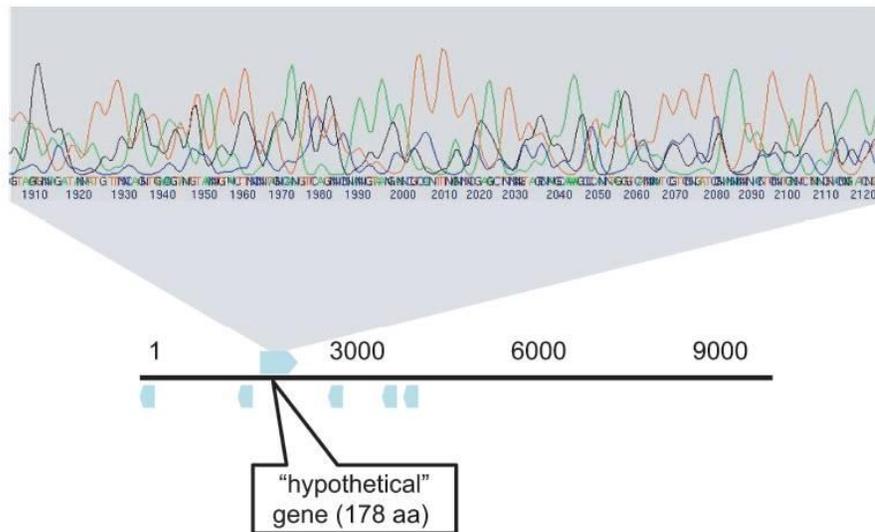


Figura 11. Parte do cromatograma de uma leitura de baixa qualidade onde genes não existentes poderiam ser preditos Fonte: Kunin et al. (2008).

2.11.2. Montagem

A montagem é o processo de combinação de leituras de sequências em trechos contíguos de DNA chamados contigs, baseando-se na similaridade de sequências sobrepostas entre as leituras. De modo geral, a maioria dos programas de montagem segue uma abordagem

de três fases: sobreposição, desenho (esboço) e consenso (Lapidus, 2009). Um dos programas de montagem mais comumente utilizados é o Phrap (Green, 1994), o qual analisa o conjunto de leituras para encontrar pares com sequências perfeitamente correspondentes e usa um algoritmo para encontrar o melhor alinhamento entre as sequências. Uma grande variedade de programas de montagem vem se tornando cada vez mais disponíveis, tanto para sequências derivadas do sequenciamento por Sanger como para sequências derivadas de plataformas de segunda geração, incluindo CAP3, Celera Assembler, Newbler, Mira, Euler e Velvet, entre outros.

Dado que os insertos podem ser sequenciados de ambos os extremos, resultando em pares de leituras irmãs (orientadas uma para a outra), o conhecimento da distância entre essas leituras pareadas é uma informação importante para a montagem. O CAP3 é capaz de usar a informação dessas leituras pareadas *forward* e *reverse* para corrigir erros de montagem e gerar os contigs (Huang e Madan, 1999). O Celera Assembler foi desenvolvido especificamente para permitir a montagem do genoma humano e foi reportado pela primeira vez quando o sequenciamento e a montagem do genoma da mosca da fruta foi concluído em 2000 (Myers et al., 2000). Porém, as montagens de genomas bacterianos com Celera não têm produzido melhores resultados que o Phrap para Sanger e Newbler para a plataforma de sequenciamento 454. O Newbler, por sua vez, foi desenvolvido pela empresa 454 Life Sciences (Roche) para a montagem de sua própria plataforma de sequenciamento e otimizado para o tamanho de sequências geradas (Margulies et al., 2005). O programa consiste de três módulos executando as operações de sobreposição, desenho e consenso usando apenas fluorogramas (equivalentes aos eletroferogramas da química Sanger) de alta qualidade para um maior desempenho (Lapidus, 2009). No primeiro módulo, *Overlapper*, o montador utiliza as sobreposições de leituras para a formação de contigs por comparação direta entre os fluorogramas. O módulo *Unitigger* agrupa as leituras de sequências que possuem sobreposições consistentes em unitigs. Por fim o módulo *Multialigner* faz o processo de otimização dos unitigs em contigs onde qualquer deficiência identificada nas outras etapas é corrigida e os contigs são estendidos (Weis, 2010).

Ganhando conhecimento sobre os diversos programas montadores, uma coisa é clara: não existe uma fórmula mágica para montar conjuntos de dados, e todos os programas de montagem vão cometer inúmeros erros. O ideal, portanto, é a inspeção manual antes da sequência ser lançada em bases públicas. Os erros na montagem podem ser facilmente identificados com ferramentas de visualização, como CONSED (Gordon et al., 1998). No entanto, a escala em que são produzidos os dados de metagenômica impede a inspeção manual e a identificação de erros de montagem. Uma abordagem para resolver esta limitação é fazer duas ou mais montagens com os mesmos dados em montadores diferentes (Kunin et al., 2008).

2.11.3. Predição de genes e anotação

O primeiro passo para a análise detalhada de uma sequência de DNA é a identificação de regiões funcionais com a aplicação de ferramentas bioinformáticas que permitam a predição de genes funcionais que codificam proteínas e genes estruturais, que não codificam proteínas, como RNAs transportadores (RNAt), genes de RNA ribossomal (RNAr) e sítios de ligação ao ribossomo (RBS). Existem duas abordagens principais para a predição de genes. O primeiro deles é o método “baseado em evidência”, que usa buscadores de homologia para identificar genes similares àqueles reportados previamente, como BLAST ou CRITICA (Badger e Olsen, 1999). Ao contrário, a segunda abordagem “ab initio” baseia-se nas características intrínsecas de uma sequência de DNA para discriminar entre sequências codificantes e não codificantes, permitindo a identificação de genes sem homólogos disponíveis em bases de dados. Alguns programas com esta abordagem são o GLIMMER (Delcher et al., 1999), o GeneMark (Besemer e Borodovsky, 1999) e MetaGene (Noguchi et al., 2006).

Dado que a identificação de genes é fundamental em projetos metagenômicos, o emprego de um método baseado em evidência estaria influenciado fortemente pela disponibilidade de sequências publicadas nas bases de dados e muitos genes novos de interesse (elementos chave em metagenômica) são perdidos (Kunin et al., 2008). No intuito de resolver esse problema, um dos programas mais conhecidos, criado com a estratégia “ab initio” e próprio para análise de sequências de metagenoma, é o MetaGene (Noguchi et al., 2006). Este programa leva em conta as frequências di-codon estimadas pelo conteúdo de GC de uma sequência e várias outras medidas derivadas da natureza fragmentada dos dados, qualidade do sequenciamento e montagem. Recentemente, vários métodos novos para predição de genes em sequências metagenômicas têm sido descritos, como Orphelia (Hoff et al., 2009) e MetaGeneMark, sendo que este último tem mostrado maior precisão na predição de genes que Metagene (Zhu et al., 2010).

Os genes de RNA (RNAt e RNAr) são preditos usando ferramentas como tRNAscan-SE (Lowe e Eddy, 1997) para RNAt e buscadores similares para RNAr. Outros tipos de RNA, como siRNAs (do inglês: *small interfering RNA*) podem ser detectados por comparação com modelos de covariância e motivos estruturais (Kunin et al., 2008). No entanto, essas buscas são computacionalmente dispendiosas e proibitivamente longas para dados de metagenoma. Em geral, a identificação de genes de RNAs não codificantes é difícil desde que essas sequências não são conservadas e os métodos “ab initio” confiáveis são escassos, inclusive para genomas de organismos isolados. Atualmente os genes que codificam RNAs não codificantes são excluídos das análises a jusantes, no entanto, espera-se que a situação mude nos próximos anos com o advento dos dados de transcriptômica que enriquecerá os inventários destes genes (Kunin et al., 2008).

A predição de genes é usualmente seguida da anotação funcional. A anotação funcional em sequências metagenômicas é similar à anotação genômica e baseia-se nas comparações de genes preditos com os genes existentes, ou seja, as sequências previamente anotadas. O BLAST (Altschul et al., 1997) é a ferramenta mais freqüentemente usada para a previsão computacional da função conhecida com a anotação por similaridade de sequência (Friedberg, 2006). De fato, o artigo sobre BLAST de Altschul (1997) tem mais de 35.500 citações até os dias de hoje (Maio de 2011). Apesar desta ferramenta se mostrar como uma rotina nas pesquisas, nos últimos anos tem sido sugerido que o uso de BLAST para a atribuição da função de proteínas deve ser evitado, principalmente por causa da potencial propagação de erros através das bases de dados (Kuning et al., 2008).

Uma alternativa é a anotação por identificação de domínios protéicos em bancos de dados de famílias de proteínas como TIGRFAM (Selengut et al., 2007), PFAM (Finn et al., 2008) e COGs (Tatusov et al., 1997), entre outras, usando também o alinhamento das sequências. Quando esses métodos de alinhamento de sequências falham, as famílias de proteínas putativas podem ser encontradas com buscas de homólogos estruturais usando a informação da composição dos aminoácidos e ignorando a ordem em que os aminoácidos estão na sequência. O PROPSEARCH (Hobohm e Sander, 1995) é provavelmente um dos primeiros métodos a explorar essas outras características, como peso molecular, conteúdo de resíduos de aminoácidos volumosos, conteúdo de resíduos de aminoácidos pequenos, hidrofobicidade média, carga média, conteúdo de grupos de dipeptídeos selecionados, entre outras características, possibilitando a predição da função em proteínas com similaridade de sequência abaixo de 25% (Hobohm e Sander, 1995).

3. JUSTIFICATIVA

Apesar dos estudos sobre comunidades microbianas com capacidade de degradação de hidrocarbonetos terem se intensificado nos últimos anos, a informação sobre as vias catalíticas responsáveis pelos processos de degradação em reservatórios de petróleo permanece escasso. O emprego de estratégias moleculares modernas, como a metagenômica, para o estudo da diversidade genética e elucidação de genes ou vias de degradação em reservatórios de petróleo, os quais constituem ambientes conhecidos como extremos, poderá trazer resultados relevantes na identificação mais detalhada dos processos envolvidos na biodegradação de hidrocarbonetos. Além disso, o sequenciamento e análise estrutural de operons responsáveis pela degradação de alcanos e aromáticos permitirão identificar novos biocatalistas com grande potencial biotecnológico, tornando o projeto inovador na área de petróleo no Brasil. A aplicação de novos biocatalistas em processos de descontaminação de ambientes poluídos ou de recuperação secundária de petróleo (MEOR – *Microbial Enhancement Oil Recovery*), oferece enorme potencial tecnológico em áreas estratégicas para a indústria de exploração e processamento de petróleo mundial.

Cabe salientar ainda que, o laboratório da Divisão de Recursos Microbianos do CPQBA/UNICAMP começou os primeiros estudos dos processos biogeoquímicos de ambientes petrolíferos há quase 10 anos e que, no decorrer desse tempo, esses estudos têm evoluído e hoje estão consolidados pelo grupo de Metagenômica do Petróleo, coordenado pela Dra. Valéria Maia de Oliveira. Atualmente, diversas abordagens de biologia molecular estão sendo utilizadas pelo referido grupo para o estudo destes ambientes e têm possibilitado parcerias frutíferas com o grupo de pesquisa em Biocatálise coordenado pela Profa. Dra. Anita J. Marsaioli (Instituto de Química/UNICAMP) e com o Laboratório de Análise Genética e Molecular coordenado pela Profa. Dra. Anete Pereira de Souza (CBMEG/UNICAMP), os quais contribuem positivamente à multidisciplinariedade das pesquisas e permitem ampliar os conhecimentos e expertise da equipe nas áreas de biodegradação, engenharia genética, e, futuramente, de expressão de enzimas de interesse.

4. OBJETIVOS

4.1. Objetivo geral

Caracterizar estrutural e funcionalmente genes envolvidos na degradação de hidrocarbonetos, originados de metagenoma microbiano, contribuindo para esclarecer os mecanismos de degradação que ocorrem em reservatórios de petróleo.

4.2. Objetivos específicos

- Avaliar por cromatografia a capacidade de degradação de clones selecionados preliminarmente em triagens funcionais de alto desempenho;
- Extrair o DNA fosmidial de clones selecionados da biblioteca metagenômica;
- Construção de sub-bibliotecas de plasmídeos (bibliotecas shotgun) para sequenciamento completo de cada um dos fosmídeos selecionados que apresentaram melhores capacidades degradadoras;
- Analisar as sequências dos insertos metagenômicos para a montagem, predição e anotação de genes;
- Identificar e caracterizar estruturalmente arranjos genômicos responsáveis pela degradação;
- Analisar as relações evolutivas das sequências gênicas responsáveis pelas vias de degradação com relação àquelas já conhecidas através de inferência filogenética;
- Contribuir para o conhecimento das vias de degradação aeróbia e anaeróbia atuantes em processos biogeoquímicos que ocorrem nos reservatórios de petróleo.

5. MATERIAL E MÉTODOS

5.1. Origem dos clones fosmidiais

Os clones utilizados neste estudo foram selecionados a partir de uma biblioteca metagenômica fosmidial obtida em estudos prévios pelo nosso grupo de pesquisa (Vasconcellos et al., 2010) e mantida no acervo de pesquisa da Coleção Brasileira de Micro-organismos de Ambiente e Indústria (CBMAI), da Divisão de Recursos Microbianos (CPQBA/UNICAMP). Esta biblioteca foi construída com técnicas de enriquecimentos microbianos anaeróbios e aeróbios a partir de amostras de petróleo bruto originada do reservatório de petróleo GMR75 na Bacia Potiguar, classificado como moderadamente biodegradado. A construção da biblioteca foi feita utilizando o kit de clonagem *Cloning-Ready Copy Control pCC2FOS* (EPICENTRE®), o qual possui

vetores do tipo fosmídeos que suportam fragmentos de até aproximadamente 45 Kb e célula hospedeira *E.coli* EPI-300, resistente ao fago T₁. Os clones foram selecionados por apresentarem atividades positivas em triagens preliminares de alto desempenho (tolerância à presença de hidrocarbonetos no meio e de utilização de hexadecano como principal fonte de carbono) (Vasconcellos et al., 2010).

5.2. Análises cromatográficas da biodegradação de hexadecano e fenantreno

Os clones foram avaliados quanto à capacidade de biodegradação, através de cromatografia gasosa acoplada à espectrometria de massas (CG-EM). Essas análises foram realizadas em colaboração com o grupo de pesquisa em biocatálise da Profa. Dra. Anita J. Marsaioli (Instituto de Química/UNICAMP). Inicialmente, foi utilizado o hexadecano como composto modelo para estudar a capacidade de degradação de hidrocarbonetos alifáticos e, posteriormente, para aqueles clones com capacidade de degradar hexadecano foi avaliada a habilidade de degradar fenantreno, representando um composto modelo de hidrocarbonetos poliaromáticos.

No ensaio de degradação do hexadecano, os clones previamente selecionados em triagens colorimétricas de alto desempenho (Vasconcellos et al., 2010) foram cultivados em “pools” de clones aleatórios no meio de cultura LB (500 mL) acrescido com cloranfenicol (12,5 µg/mL), durante 24 h a 37 °C. Após o crescimento das células, as mesmas foram centrifugadas (5.000 rpm, 20 min, 4°C) e lavadas em água Milli-Q. Seguiu-se a suspensão das células em água Milli-Q, na concentração de 10⁸ NMP/mL. Erlenmeyers (125 mL), contendo 40 mL de Meio Mineral Modificado (14 g de NaCl, 2,8 g de K₂HPO₄, 2,8 g de KH₂PO₄, 2,8 g de (NH₄)₂SO₄, 0,56 g de MgSO₄ ·H₂O e 8,4 g de NaNO₃, em 1000 mL de água destilada), e cloranfenicol (12,5 µg/mL), foram inoculados com 4 mL da suspensão celular dos clones e 400 µL de hexadecano (Sigma) filtrado, em 5 replicatas experimentais. 40 µL de uma solução de vitaminas (Nakayama, 2005) também foi adicionada ao meio, como forma de ativar o crescimento microbiano e dar início à degradação ou biotransformação dos substratos adicionados como fonte de carbono (Zinder et al., 1984). Os ensaios foram amostrados nos tempos 0, 7, 14, 21 e 28 dias, utilizando-se hexano bidestilado como solvente extrator. Após extração da fase orgânica das amostras, as mesmas foram submetidas à secagem sob fluxo de N₂, e suspensas em solução de fenantreno (0,5 mg/mL), o qual foi utilizado como padrão interno de quantificação nas análises de CG-EM. Foram utilizadas como controle positivo a bactéria *Micrococcus* sp (CBMAI 636) e como controle negativo a cepa *Escherichia coli* EPI300. Os clones do *pool* que mostrou elevada degradação de hexadecano foram a seguir avaliados individualmente utilizando os mesmos protocolos descritos acima.

As condições para o ensaio de degradação de fenantreno foram similares ao ensaio com hexadecano. Cada clone foi crescido em meio LB + cloranfenicol 12,5 µg/mL durante 24 horas a 37°C e, após esse período, as células foram centrifugadas e suspensas em água Milli-Q esterilizada até a concentração aproximada de 10⁸ NMP/mL. Alíquotas de 5 mL da suspensão de células de cada clone foram inoculadas em frascos erlenmeyers de 125 mL contendo 50 mL de Meio Mineral Modificado (14 g de NaCl, 2,8 g de K₂HPO₄, 2,8 g de KH₂PO₄, 2,8 g de (NH₄)₂SO₄, 0,56 g de MgSO₄ ·H₂O e 8,4 g de NaNO₃, em 1000 mL de água destilada), acrescidos de 50 µL da solução de vitaminas (Nakayama, 2005), cloranfenicol na concentração final de 12,5 µg/mL e 50 mg de fenantreno como única fonte de carbono. Os ensaios foram incubados a 37°C e 100 rpm por 14 dias. Após esse período, os metabólitos produzidos pelos clones nos tempos 0, 7 e 14 dias foram extraídos com acetato de etila, rota-evaporados e suspensos em solução de nonadecano (0,5 mg), como padrão interno. Foram utilizadas como controle positivo a bactéria *Micrococcus* sp (CBMAI 636) e como controle negativo a cepa *Escherichia coli* EPI300.

Para as análises espectrométricas utilizou-se o cromatógrafo a gás HP5890II acoplado a detector de massas HP 5970-MSD, injetor tipo split (10:1), He como gás de arraste, fluxo de 1 mL / min, coluna capilar de sílica fundida do tipo HP5 (30 m x 0,25 mm x 0,25 µm), cuja fase estacionária consiste de 5% de fenil metil silicone. O forno foi operado na seguinte programação: 60 a 290°C, 20°C / min, durante 5 minutos. O injetor foi utilizado à temperatura de 240°C. Os dados foram adquiridos utilizando-se o modo SCAN e os valores de massa variaram entre 50 e 700 u.m.a (unidades de massa atômica).

A extensão da biodegradação foi medida a partir dos dados cromatográficos, aplicando as equações e cálculos descritos em Aldas et al. (2007) e Vasconcellos et al. (2009).

5.3. Construção de Sub-bibliotecas

Os clones fosmidiais com elevada capacidade de biodegradação dos hidrocarbonetos testados foram selecionados para serem sequenciados. A partir de cada fosmídeo (com inserto de aproximadamente 40 kb) foi construída uma sub-biblioteca em vetor plasmidial (insertos de aproximadamente 1 kb), utilizando uma abordagem de sequenciamento aleatório denominada “shotgun” (Green, 1997). Esses experimentos foram realizados e conduzidos em colaboração e supervisão da Profa. Dra. Anete Pereira de Souza e o grupo do Laboratório de Análise Genética e Molecular do CBMEG/UNICAMP.

5.3.1. Extração de DNA fosmidial

O DNA fosmidial de cada clone foi extraído com o kit de isolamento e purificação de plasmídeos *Invisorb® Plasmid Maxi Kit* (Invitex). Devido às características do fosmideo (comprimento maior que um plasmídeo padrão) algumas modificações precisaram ser feitas no protocolo fornecido pelo fabricante. A solução de eluição foi pré-aquecida a 65°C e foram feitas duas eluições com metade do volume recomendado, a fim de incrementar o rendimento das extrações. O emprego deste kit permite a remoção do DNA genômico contaminante, que ocorre nas purificações de DNA com métodos caseiros, para obter um DNA suficientemente limpo para as subseqüentes aplicações. O uso deste kit evita também a super-estimativa da quantidade de DNA fosmidial de interesse. Posteriormente, procedeu-se à visualização do DNA obtido em gel de agarose 0.8%.

5.3.2. Produção de fragmentos aleatórios de DNA fosmidial

O DNA extraído dos clones fosmidiais foi submetido à fragmentação randômica por sonicação no equipamento Microson Ultrasonic Cell disruptor XL. Aproximadamente 10 µg de DNA foram distribuídos em alíquotas em microtubos de 1,5 mL e mantidos no gelo em uma câmara fria. As soluções contendo o DNA fosmidial foram sonicadas com pulsos de 5 segundos durante 10 e 15 segundos com potencia 2. Após uma rápida centrifugação, as alíquotas de cada tempo de sonicação foram observadas em gel de agarose 1% para determinar as condições que produziram os fragmentos de tamanho na faixa desejada (600 – 2000 pb). Esses fragmentos foram recuperados do gel, purificados utilizando-se mini-colunas (*GFX PCR DNA and gel band kit*, GE Healthcare) e estocados a -20°C para serem posteriormente clonados.

5.3.3. Reação de reparo das extremidades dos fragmentos de DNA

Os fragmentos purificados foram tratados em suas extremidades com a enzima T4 DNA Polimerase e Klenow. Nessa reação foram usados 8 µg de DNA em um volume de 50 µL, 16 µL de Tampão de T4 polimerase 5X (Invitrogen), 2 µL BSA (10 mg/mL), 8 µL dNTP's 10 mM e 4 µL de T4 DNA polimerase 5 U/µL (Invitrogen). A reação foi incubada a 12°C por 15 min e, a seguir, foi adicionada 1 µL da enzima Large Fragment of DNA Polimerase I –Klenow (6 U/µL, Invitrogen) e novamente incubada por 30 min a 37°C, e finalmente a 75°C por 20 min para inativação da enzima.

5.3.4. Reação de fosforilação

Ao volume completo de DNA reparado (80 µL) foram adicionados 10 µL de Tampão ligase + 1 mM ATP (Biolabs), 1 µL de enzima Kinase 10 U/µL (Invitrogen) e água Milli-Q para o volume final de 100 µL. A reação foi incubada a 37°C por 30 min e, a seguir, a 65°C por 20 min para

inativação da enzima, e a reação foi mantida a 4°C por aproximadamente 12 h. Posteriormente, o volume completo da reação foi purificado com mini-colunas (*GPX PCR DNA and gel band kit*, GE Healthcare), seguindo as recomendações do fabricante e, em seguida, a concentração do DNA final foi estimada.

5.3.5. Reação de ligação

O material purificado e fosforilado foi ligado ao vetor plasmidial pBluescript II SK+ (Stratagene) (Figura 12). O DNA foi clonado no sítio múltiplo de clonagem com a enzima de restrição EcoRV, interrompendo o gene LacZ do vetor. Portanto, previamente à reação de ligação, o vetor pBluescript foi digerido com a enzima EcoRV e desfosforilado com a fosfatase alcalina *Calf intestinal Alkaline Phosphatase*, CIAP (Invitrogen).

A mistura de ligação foi preparada com base na molaridade do inserto e do vetor, na proporção 6:1 (inserto:vetor) em 4 µL de Tampão T4 DNA ligase (Invitrogen), 1 µL de enzima T4 DNA ligase 5 U/µL (Invitrogen) e água Milli-Q esterilizada, em um volume total de 20 µL. A ligação foi incubada a 16°C por aproximadamente 12 horas. Após a incubação, a solução foi utilizada para a transformação de células competentes de *Escherichia coli*.

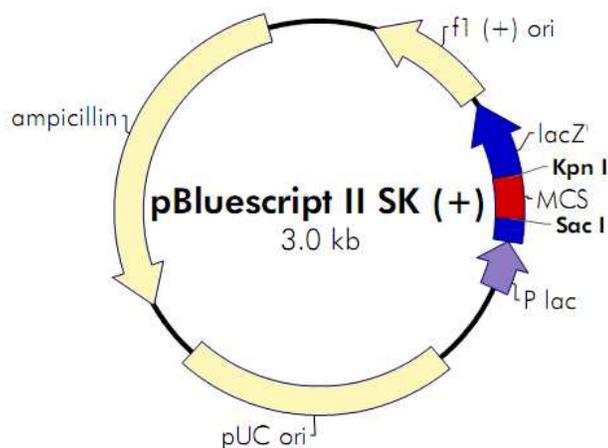


Figura 12. Esquema ilustrativo das características do vetor pBluescript II SK+ (Stratagene).

5.3.6. Produção e transformação de células competentes

A transformação de células competentes foi realizada por eletroporação.

5.3.6.1. Produção de células competentes para transformação

Para os experimentos de transformação da linhagem de *E. coli* DH10B, foi necessária a obtenção de células electrocompetentes seguindo o protocolo de Kasap et al. (2008), com algumas modificações detalhadas a seguir. A partir de um pré-inóculo de 50 mL da cepa *E. coli*

H10B em meio Circle Grow (CG), incubado durante 16 h a 37°C sob agitação (250 rpm), foi inoculado um frasco Erlenmeyer de 1 L contendo 500 mL de meio CG, com volume apropriado para começar a cultura de 500 mL em $D.O_{600nm}$ de aproximadamente 0,1. A cultura foi incubada a 37°C sob agitação com freqüentes medidas de densidade óptica, até obtenção de $D.O_{600nm}$ 0,6. Atingindo este ponto, a cultura foi resfriada em gelo por 10 min e centrifugada a 6.000 X g por 10 min a 4°C. Após a centrifugação, o sobrenadante foi descartado e os tubos mantidos invertidos por alguns minutos para eliminação do restante do sobrenadante.

Decorrido este tempo, a massa das células foi suspensa suavemente com agitação manual, mantendo sempre os tubos gelados, em 200 mL, 20 mL e 2 mL de glicerol 10% gelado, consecutivamente, e centrifugadas após cada suspensão, conforme descrito acima. A suspensão final das células foi feita em 1 mL de glicerol 10% e distribuída em alíquotas de 40 μ L em microtubos de 1,5 μ L previamente resfriados. Estas alíquotas foram congeladas com nitrogênio líquido e armazenadas a -80°C.

5.3.6.2. *Transformação de células competentes por eletroporação*

A reação de ligação foi utilizada para transformar células de *E. coli* DH10B eletrocompetentes utilizando-se o equipamento *E. coli PulserTM Transformation Apparatus* (Bio-Rad). Este método descrito por Sambrook et al. (1989) permeabiliza a membrana, por carga eléctrica, formando fendas temporárias que permitem a entrada dos plasmídeos. Primeiramente, as células eletrocompetentes foram descongeladas em gelo. Foi adicionado 1 μ L de reação de ligação a 40 μ L de células competentes e procedeu-se à sua homogeneização. Após 1 minuto, a mistura foi depositada numa cubeta (2 mm) previamente resfriada em gelo. Procedeu-se à descarga eléctrica programada para 25 μ F, 200 Ω e 2,5 kV na câmara do aparelho que resultaram em pulsos eléctricos de aproximadamente 4.8 – 5.2 milisegundos. As culturas foram recuperadas com 1 mL de meio LB o mais rapidamente possível.

As células foram crescidas durante 1 hora a 37°C, com agitação moderada de 200 rpm. Após esse período, diferentes diluições das bactérias transformadas foram inoculadas em placas de Petri contendo Luria Bertani ágar + ampicilina (100 μ g/mL) + 100 μ L de IPTG (0,1 M) e X-gal (5-bromo-4-cloro-3-indolil- β -D-galactopiranosídeo) 2%, e incubadas a 37°C por 16 h. As células transformantes positivas (colônias brancas) foram selecionadas com palitos de madeira esterilizados e repicadas tanto para uma nova placa de Petri de LB Agar + ampicilina (100 μ g/mL) como também para placas de 96 poços, contendo 150 μ L de meio 2YT-HMFM + ampicilina (100 μ g/mL). Em seguida, as placas foram incubadas a 37°C e 200 rpm por 24 h, e posteriormente estocadas a -80°C.

5.4. Verificação da presença e tamanho de inserto nos clones das sub-bibliotecas

A presença e estimativa do tamanho dos insertos de cada sub-clone foram determinadas por PCR (*Polymerase Chain Reaction*) de colônia usando os primers M13 *forward* (5'- CGC CAG GGT TTT CCC AGT CAC GAC - 3') e *reverse* (5'-TTT CAC ACA GGA AAC AGC TAT GAC - 3'). A reação de PCR foi feita em microplacas, após a reativação dos clones da sub-biblioteca (repique em 150 µL LB + Ampicilina a 100 µg/mL). O volume final da reação foi 50 µL contendo 0,2 µM de cada um dos *primers* M13F e M13R, 0,2 mM de dNTP's, 1,5 mM de MgCl₂, 2 U de Taq DNA polimerase (Invitrogen) e 0,5 µL da suspensão de células crescidas por 12 h (seguindo as recomendações de utilização de uma diluição 1:100 do cultivo na reação). O programa de amplificação esta constituído por 1 ciclo de 94°C por 3 minutos, 30 ciclos a 94°C por 20 segundos, 60°C por 20 segundos, 72°C por 1,5 minutos, e 1 ciclo de extensão final de 72°C por 3 minutos, em termociclador Mastercycler Gradient (Eppendorf). Os produtos de PCR foram visualizados em gel de agarose 1%.

5.5. Sequenciamento completo da sub-biblioteca

As reações de sequenciamento foram realizadas para todos os subclones utilizando aproximadamente 200 ng de produto de PCR e/ou 200ng do DNA plasmidial extraído usando os primers M13 *forward* e M13 *reverse*. A extração do DNA dos plasmídeos para o sequenciamento de seus insertos foi realizada em placas de 96 amostras pelo método de lise alcalina (Sambrook et al., 1989). As reações de sequenciamento e de purificação / precipitação das amostras dependeram do sistema utilizado para o sequenciamento. Na Divisão de Recursos Microbianos (CPQBA/UNICAMP) foram realizadas as reações com o kit *DYEnamic ET Dye Terminator Cycle Sequencing* para o sequenciamento no sistema MegaBace 500 (GE Healthcare), seguindo as recomendações do fabricante. As reações continham 5 µL do DNA amplificado (aproximadamente 200 ng de DNA), 0,16 µL de um único *primer* (3,2 pM) e 4 µL de DYEnamic ET Dye terminator e água qsp 10 µL. Esta reação foi submetida ao termociclador Mastercycler Gradient (Eppendorf) programado para 1 minuto de desnaturação inicial a 95 °C, seguido de 30 ciclos, sendo cada ciclo constituído das etapas de desnaturação (95°C, 20 seg), pareamento (dependendo do primer utilizado: 50°C com M13F e 55°C com o primer M13R), e extensão (60°C).

No sistema de sequenciamento com o equipamento ABI prism 377 DNA sequencer [™] (Applied Biosystems) do Laboratório de Análise Genética e Molecular (CBMEG–UNICAMP), a reação de sequenciamento foi constituída de 4 µL de DNA plasmidial (aproximadamente 200 ng de DNA), 1 µL de um único *primer* (5 pM), 2 µL tampão *Save Money* (500 µL de MgCl₂ 50mM e 1mL de Tris-HCl 1M pH 9,0), 2 µL do reagente "Bid Dye" versão 3 (Applied Biosystems) e água qsp 10 µL.

As condições da reação foram: desnaturação inicial a 96°C durante 2 minutos, seguida de 26 ciclos, sendo cada ciclo constituído das etapas de desnaturação (96°C, 45 seg), pareamento (50°C, 30 seg) e extensão (60°C, 4 minutos). As reações foram realizadas no termociclador Mastercycler Gradient (Eppendorf).

Posteriormente, as reações de sequenciamento foram precipitadas, suspensas no tampão de corrida apropriado para cada sistema (10 µL do tampão “loading” componente do kit para o MegaBace e 4 µL de BlueDextran + formamida para o sistema ABI 377) e então aplicadas no respectivo sequenciador automático. Do mesmo modo, o sequenciamento das duas extremidades de cada fosmídeo foi realizado utilizando DNA fosmidial extraído e os *primers* pCC2F *forward* (GTACAACGACACCTAGAC) e pCC2R *reverse* (CAGGAAACAGCCTAGGAA)(Invitrogen), e as condições descritas acima para sequenciamento de DNA plasmidial.

5.6. Análise das sequências geradas através de bioinformática

5.6.1. Limpeza e montagem de sequências

A avaliação da qualidade e a montagem das sequências foram realizadas com varias ferramentas disponíveis. Apenas as sequências de alta qualidade (Q20), conforme avaliado pelo Phred (Ewing e Green, 1998; Ewing et al., 1998), foram utilizadas na análise. Os softwares LUCY (Chou e Holmes 2001) e BDtrimmer (Mondego et al., 2011) foram utilizados para avaliação da qualidade de corte (*quality trimming*), remoção de bases nucleotídicas de baixa qualidade, remoção de sequencias dos vetores (pCC2FOS e pBluescript II SK+) assim como possíveis sequencias contaminantes provenientes de *E. coli*. O algoritmo de phrap (Ewing et al., 1998; Ewing e Green, 1998) foi utilizado na montagem inicial. Outros programas foram empregados para avaliar a melhor montagem como Newbler assembler (Roche) (Margulies et al., 2005) e CAP 3 (Huang e Madan, 1999). O software CONSED (Gordon et al., 1998) foi usado para auxiliar a visualização, finalização e curagem manual da montagem.

5.6.2. Identificação de quadros de leitura abertos

Quadros de leitura abertos (*Open Reading Frames* - ORFs) foram identificados usando diversas ferramentas disponíveis, como ORF-finder (Tatusov e Tatusov, 2009) <http://www.ncbi.nlm.nih.gov/projects/gorf/>), Metagene (Noguchi et al., 2006; <http://metagene.cb.k.u-tokyo.ac.jp/>), MetaGeneMark (Zhu et al., 2010; <http://exon.gatech.edu/GeneMark/metagenome/>) e Orphelia (Hoff et al., 2009; <http://orphelia.gobics.de/>), desenhados para sequências derivadas de metagenoma, e GLIMMER

(Delcher et al., 1999; disponível no site do NCBI (http://www.ncbi.nlm.nih.gov/genomes/MICROBES/glimmer_3.cgi). Vários critérios foram utilizados para a definição das ORFs a partir dos resultados que surgiram entre os diferentes programas empregados. Foram mantidas apenas as ORFs com pouca ou nenhuma sobreposição, mais longas do que 30 aminoácidos e, no caso de ter ORFs em diferentes frames para um mesmo fragmento, aquelas com homólogos conhecidos foram selecionadas. Os espaços entre ORFs foram posteriormente comparados com a base de dados não redundante (nr) utilizando BLAST (Altschul et al., 1997) para garantir que nenhum quadro de leitura fosse perdido.

5.6.3. Anotação de genes

As sequências de cada ORF predita foram analisadas com várias ferramentas de bioinformática, como a busca de sequências similares utilizando BLASTp (Altschul et al., 1997) contra as bases de dados de sequências EMBL/GeneBank e de famílias de proteínas utilizando InterProScan (combina diferentes métodos de reconhecimento de proteínas contra proteínas nos bancos de dados PROSITE, PRINTS, Pfam, ProDom, SMART e TIGRFAMMs) (Quevillon et al., 2005; <http://www.ebi.ac.uk/Tools/pfa/iprscan/>). COGNITOR (Tatusov et al., 2001; <http://www.ncbi.nlm.nih.gov/COG/old/xognitor.html>) foi usado para atribuir cada ORF a categorias funcionais denominadas Clusters de Grupos Ortólogos (*Clusters of Orthologous Groups - COG*). A partir destas predições, a anotação final das ORFs foi feita manualmente e visualizada no software Artemis 12.0 (Rutherford et al., 2000), que também foi utilizado para calcular o conteúdo de guanina/citosina (% G + C). Os sítios de ligação dos ribossomos putativos foram identificados com RBSFINDER (Suzek et al., 2001) e a presença de genes RNA transportador foi explorada com tRNAscan-SE (Lowe e Eddy, 1997). A anotação automática também foi realizada usando a plataforma (*Rapid Annotation using Subsystem Technology* (RAST) (Aziz et al., 2008).

Outro esforço feito para a anotação das proteínas foi baseado na composição dos aminoácidos utilizando o programa PROPSEARCH (Hobohm e Sander, 1995; <http://abcis.cbs.cnrs.fr/propsearch/>).

Domínios transmembrana putativos nas ORFs identificadas foram preditos usando o servidor TMHMM (Krogh et al., 2001; <http://www.cbs.dtu.dk/services/TMHMM/>), o qual usa um modelo oculto de Markov (*Hidden Markov Model - HMM*) capaz de prever hélices transmembranas.

O programa Phylopythia que utiliza a frequência de padrões de oligonucleotídeos para prever com alta exatidão a origem taxonômica de fragmentos maiores de 3kb foi utilizado para a classificação filogenética dos fragmentos metagenômicos (McHardy et al., 2007).

5.6.4. Determinação das relações evolutivas

Alguns genes envolvidos na degradação de hidrocarbonetos foram comparados e analisados com relação às sequências homologas depositadas nas bases de dados. As sequências foram alinhadas utilizando o programa CLUSTAL W (Thompson et al., 1994) e analisadas para a determinação das relações evolutivas com proteínas já conhecidas através de inferência filogenética utilizando o método de *Neighbor-joining* (Saitou e Nei, 1987) no software MEGA v.5.0 (Tamura et al., 2011).

6. RESULTADOS E DISCUSSÃO

6.1. Análises cromatográficas de hexadecano e fenantreno

Foram selecionados 72 clones no screening preliminar (tolerância à presença de hidrocarboneto no meio e crescimento em hexadecano como única fonte de carbono) (Vasconcellos et al., 2010). Ensaios de degradação de hexadecano foram realizados para esses 72 clones e analisados cromatograficamente em CG-EM, para a quantificação da degradação do hidrocarboneto.

Os 72 clones foram avaliados em consórcios contendo 9 clones cada um, ou seja, 8 *pools*, os quais foram monitorados por 28 dias através de análises cromatográficas realizadas a cada 7 dias. Dentre os 8 *pools* monitorados, um consórcio denominado como CO8 apresentou taxas de degradação de hexadecano acima de 50% e foi selecionado para a verificação da habilidade degradadora dos clones que o compõem individualmente.

Desta forma, os 9 clones foram re-inoculados e analisados individualmente com a mesma metodologia e condições descritas para os ensaios em *pools*. Após os 28 dias de monitoramento, verificou-se que cinco clones, (2B, 6A, 1A, 10A e 3B) dentre os 9 investigados, apresentaram taxas de biodegradação de hexadecano acima de 70% (Tabela 5). Estes resultados foram considerados excelentes, uma vez que se trata de um sistema de expressão heterólogo e os níveis de biodegradação foram comparáveis a dados reportados para cepas bacterianas selvagens (Whyte et al., 1998).

Tabela 5. Maiores taxas de degradação de hexadecano dos clones do consorcio C08 avaliados individualmente

Clones	Taxa de degradação de Hexadecano
10A	86%
1A	91%
6A	92%
3B	70%
2B	98%

Estes clones foram então utilizados em ensaios para avaliar seu potencial na biodegradação de hidrocarbonetos aromáticos, usando o fenantreno como composto modelo. As análises dos perfis cromatográficos em CG-EM durante os dias 0, 7 e 14 do experimento, mostraram diferentes níveis de degradação, e o cálculo dos índices de biodegradação do fenantreno foi feito para cada um dos clones, como mostra a Tabela 6.

Tabela 6. Taxas de degradação de fenantreno dos clones.

Clones	Taxa de degradação de Fenantreno
10A	49%
1A	5%
6A	15%
3B	21%
2B	44%

As Figuras 13 e 14 ilustram os cromatogramas do monitoramento dos clones 2B e 10A que apresentaram as maiores taxas de degradação para o hexadecano e o fenantreno respectivamente.

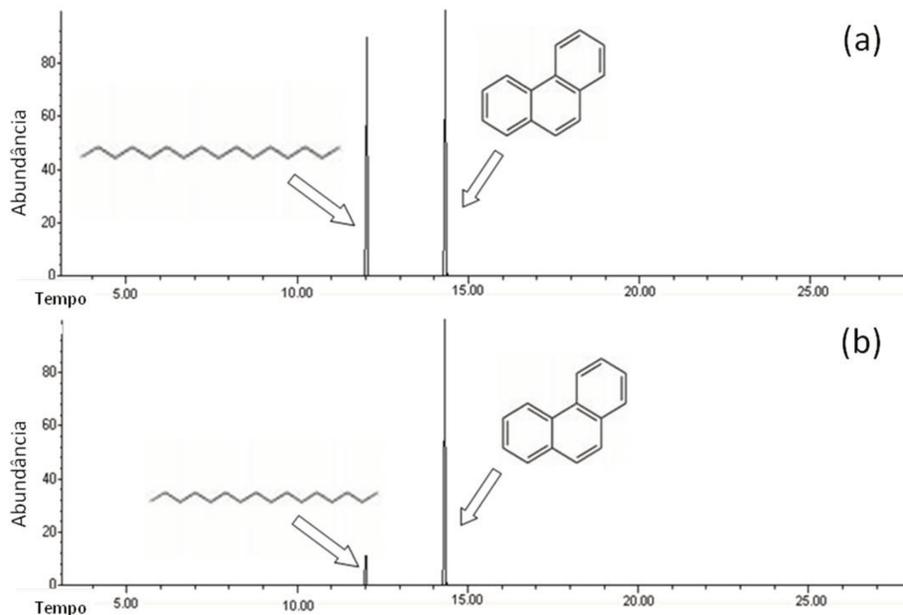


Figura 13. Cromatogramas (CG-EM) do clone 2B, o qual apresentou a maior taxa de degradação do hexadecano após 28 dias de monitoramento. (a) tempo 0; (b) após 28 dias.

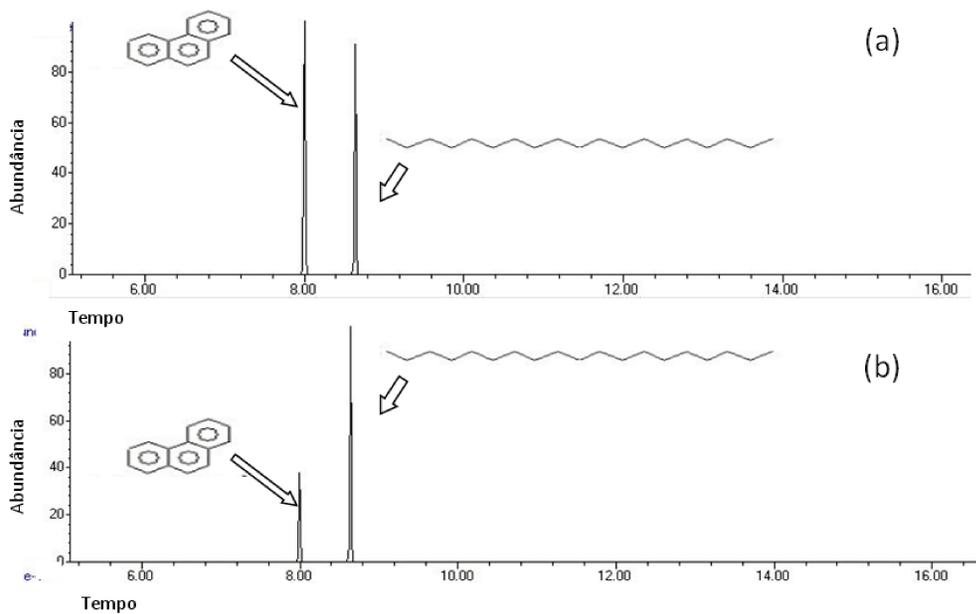


Figura 14. Cromatogramas (CG-EM) do clone 10A, o qual apresentou a maior taxa de degradação do fenantreno após 14 dias de monitoramento. (a) tempo 0; (b) após 14 dias.

Cabe salientar que, a escolha do consórcio C08 foi necessária para tornar esta dissertação exequível no prazo proposto, mas não significa que os outros consórcios não apresentem clones

com potenciais atividades biodegradadoras interessantes. É importante considerar que as taxas de degradação dos clones do consórcio C08, quando avaliados individualmente, superaram a média apresentada pelo consórcio. Este resultado poderia ser explicado pelo efeito da ação conjunta de todos os clones, possíveis interações, ou adequações biológicas à presença de outros “competidores” no meio. De qualquer forma, existe a possibilidade deste resultado ser observado também para outros *pools* de clones avaliados no trabalho de Vasconcellos et al. (2010).

6.2. Construção de sub-bibliotecas

6.2.1. Extração de DNA fosmidial

No início do projeto foi proposta a construção de 1 sub-biblioteca em vetor plasmidial (insertos de aproximadamente 1 kb) para cada um dos 5 fosmídeos (insertos de aproximadamente 40 kb) que apresentaram as melhores taxas de degradação de hexadecano. Portanto, esta primeira etapa da construção da biblioteca foi feita para os 5 fosmídeos que apresentaram as maiores taxas de degradação de hexadecano. Na figura 15 observa-se o DNA fosmidial extraído usando o kit da Invitex. As bandas visualizadas nas canaletas 1 a 5 correspondem ao DNA fosmidial dos clones selecionados da biblioteca metagenômica.

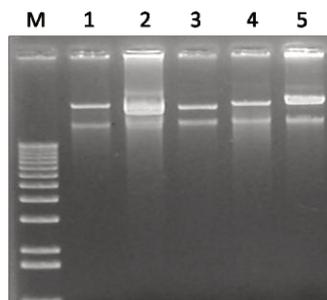


Figura 15. Eletroforese do DNA total extraído dos clones provenientes das bibliotecas metagenômicas. Canaletas: (M) Marcador de peso molecular 1 Kb Plus DNA ladder (Invitrogen); (1 a 5): DNA extraído dos clones 1A, 10A, 6G, 2B e 3B.

6.2.2. Produção de fragmentos aleatórios de DNA fosmidial (Shotgun library)

No decorrer dos experimentos, e seguindo a sugestão da banca de qualificação (Maio/2010), foram selecionados dois clones para a construção de sub-bibliotecas, sequenciamento e análises. Os clones selecionados foram o **2B**, por apresentar a maior taxa de biodegradação do hexadecano (98%) e a segunda taxa mais alta de degradação do fenantreno (44%), e o **1A**, por apresentar alta taxa de degradação do hexadecano (91%) e baixa taxa de

degradação de fenantreno. A idéia é comparar os operons contidos nestes foscídeos e identificar as vias gênicas, provavelmente distintas, que permitem a degradação de apenas 1 ou dos 2 hidrocarbonetos.

A sonicação do DNA ocorreu de acordo com o esperado, essa quebra mecânica do DNA foi necessária para a escolha do tamanho dos insertos utilizados na clonagem (Figura 16). Os fragmentos na faixa de 600 – 1600 pb foram recuperados do gel e purificados utilizando-se minicolunas (*GFX PCR DNA and gel band kit*, GE Healthcare).

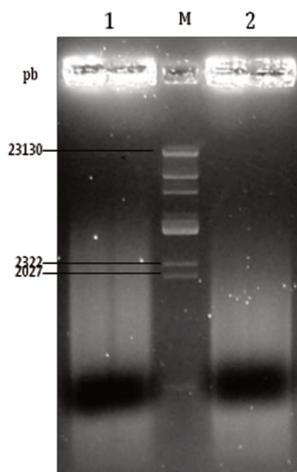


Figura 16. Eletroforese do DNA fosmidial do clone 2B fragmentado por sonicação durante tempos diferentes. Canaletas: (1): 5 segundos; (2): 10 segundos; (M): Marcador Lambda digerido com *HindIII*.

6.2.3. Reações de reparo/fosforilação/ligação e transformação de células competentes

As reações de reparo e de fosforilação dos fragmentos ocorreram de acordo com o esperado, essas reações foram necessárias para que o DNA quebrado pudesse ser ligado ao vetor. Após a quantificação dos fragmentos metagenômicos no eluato das reações, foi realizada a ligação dos insertos no vetor plasmidial. Em seguida, o DNA ligado foi transformado em células hospedeiras de *E. coli* DH10B, conforme descrito na seção de Material e Métodos.

A partir da sub-clonagem do foscídeo 1A foram coletados e armazenados 604 sub-clones, e a partir do foscídeo 2B, 960 sub-clones.

6.3. Avaliação de sub-bibliotecas e sequenciamento

A sub-biblioteca do clone fosmidial 1A (604 subclones) foi avaliada quanto à presença e tamanho dos insertos de cada um dos sub-clones por PCR direto a partir de colônias utilizando os *primers* M13F e M13R. Do total dos 604 sub-clones, 441 apresentaram inserto com tamanho médio de 1081 pb. Na figura 17 observam-se alguns dos produtos de PCR obtidos para essa sub-biblioteca.

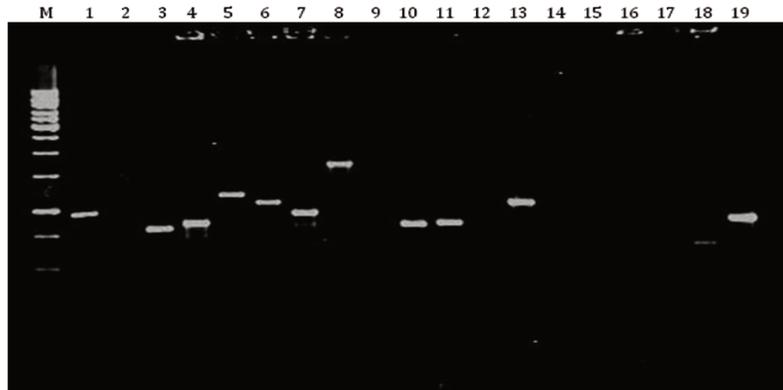


Figura 17. Eletroforese em gel de agarose 1% dos insertos dos sub-clones do fósideo 1A amplificados com os *primers* M13F e M13R. As canaletas 1 a 19 mostram a presença ou ausência de produtos de PCR nos sub-clones FOS1AC4 E8 a G2. M: Marcador de peso 1 kb DNA ladder (Fermentas).

O sequenciamento dos clones positivos foi realizado e as sequências geradas foram analisadas com a utilização dos softwares citados. Esses softwares propiciaram a verificação da qualidade obtida no sequenciamento para cada clone. Desta forma, foi possível identificar as sequências que não apresentavam qualidade satisfatória para o procedimento das análises, indicando assim, quais clones deveriam ser sequenciados novamente.

6.4. Montagem das sequências

Após sucessivos sequenciamentos dos sub-clones, as sequências geradas no sequenciador ABI 377 e as obtidas com o sequenciador MEGABACE foram reunidas e submetidas às análises citadas no Item 5.6.1. Um total de 697 sequências foram avaliadas, e após os procedimentos de “trimming” (remoção de sequências de baixa qualidade, de vetores e sequências contaminantes de *E. coli*), 507 sequências foram utilizadas na montagem (*Assembly*). Finalmente, foram obtidos 20 contigs e 129 singlets através do programa CAP 3. As características gerais da montagem são mostradas na Tabela 7.

A finalização da montagem do fósideo 1A, no intuito de se obter um único contig do tamanho total do inserto, poderia ser concluída fechando os gaps presentes entre os insertos já obtidos através da técnica de *primer walking* ou através do sequenciamento de sub-clones adicionais. Entretanto, estes resultados não foram concluídos até agora, dados alguns obstáculos no desenvolvimento do projeto, iniciando pelos problemas técnicos do equipamento sequenciador até a grande demanda de tempo e sequências de qualidade requeridas na metodologia de sequenciamento por shotgun.

A sub-biblioteca do Fosmídeo 2B foi sequenciada comercialmente (MACROGEN INC., Coréia) e foi obtido um único contig de 37.367 pb. As análises de bioinformática subsequentes foram realizadas para ambos os contigs dos fosmídeos 1A e 2B.

Tabela 7. Estatística da montagem obtida para o sequenciamento da biblioteca *shotgun* do fosmídeo metagenômico “FOS1A”.

Estatística da montagem	
Comprimento total	30408
No. de contigs	20
No. de singlets	129
Tamanho médio de contigs	1520
Contig maior	8175
N50	3368 (3 contigs)

6.5. Características gerais dos insertos fosmidiais

As sequências consenso dos diferentes contigs derivados da montagem do fosmídeo 1A e a sequência do único contig do fosmídeo 2B foram submetidas à análise utilizando os diversos programas citados no item 5.6. A predição das ORFs e outras características anotadas dos insertos metagenômicos foi manualmente refinada no programa Artemis (Figura 18). As características gerais dos insertos, como tamanho, conteúdo de G+C, total de ORFs preditas, encontram-se resumidas na Tabela 8. Os produtos dos genes preditos foram divididos em categorias funcionais de acordo com a lista fornecida pelo agrupamento de grupos ortólogos de proteínas (COGs). De um total de 63 ORFs preditas, 60 ORFs foram atribuídas a alguma categoria COG. O número total de produtos de genes atribuídos a cada categoria é mostrada na Tabela 9.

Nenhum marcador taxonômico conhecido, como RNA ribossomal 16S, estava presente nas sequências metagenômicas analisadas. Portanto, a origem microbiana dos insertos foi predita com o servidor Phylopythia, o qual sugeriu que os fragmentos fosmidiais 1A e 2B pertencem ao filo *Proteobacteria* com alguns fragmentos de 1A identificados até a classe *AlfaProteobacteria*. Nenhum dos fosmídeos apresentou sequências putativas de RNA transportador (RNAt). As informações da existência de hélices transmembrana preditas com o uso do programa TMHMM complementaram as informações da anotação. Foram preditos domínios transmembrana para 10 proteínas preditas no fosmídeo 2B e 8 ORFs com domínios transmembrana no fosmídeo 1A.



Figura 18. Esquema das características anotadas para o inserto fosmidial FOS2B proveniente da biblioteca metagenômica de reservatório de petróleo. Visualização através do programa Artemis 12.0.

Tabela 8. Resumo das características gerais dos dois fragmentos metagenômicos analisados.

	Fosmideo 2B	Fosmideo 1A
Tamanho do inserto (pb)	37367	30408*
% G + C	65,32%	62,38%
Número de ORFs preditas	40	23
Comprimento médio de gene (pb)	886	840
% total de ORFs hipotéticas	10	9
Número de ORFs atribuídas a COG	39	21
Filo representado	Proteobacteria	Proteobacteria

*Pares de bases na montagem

Tabela 9. Resumo do numero total de ORFs encontradas para cada fosmideo atribuídas em categorias COG estabelecidas.

Código COG	No. de ORFs fosmideo 2B	No. de ORFs fosmideo 1A	Categorias funcionais
Armazenamento e processamento de informações			
J	0	1	Tradução
K	3	1	Transcrição
L	0	2	Replicação, recombinação e reparo
Processos celulares e sinalização			
D	2	0	Controle do ciclo celular, divisão, ou particionamento do cromossomo
T	0	5	Mecanismos de transdução de sinal
M	2	1	Membrana/parede celular
U	1	0	Tráfego intracelular, secreção e transporte vesicular
O	2	1	Modificação pós-traducional, chaperonas
Metabolismo			
C	11	2	Produção de Energia e conversão
E	5	2	Transporte de aminoácidos e metabolism
G	2	0	Transporte e metabolismo de carboidratos
F	0	1	Transporte de nucleotídeos e metabolismo
H	1	1	Transporte de Coenzima e metabolismo
I	3	0	Transporte e metabolismo lípidico
Q	4	0	Biossíntese, transporte e metabolismo de metabólitos secundários
Pouco caracterizadas			
R	2	2	Predição da função geral
S	1	2	Função desconhecida

6.6. Anotação do fosmideo 2B

Várias ferramentas de bioinformática foram utilizadas para a identificação dos quadros de leitura abertos. As funções preditas para as 40 ORFs detectadas no fosmideo 2B são descritas na Tabela 10 com informação acerca do tamanho das regiões codificantes em aminoácidos, suas referências relacionadas às bases de dados INTERPRO e COG e os homólogos mais próximos nas bases de dados públicas. Funções foram atribuídas para 36 ORFs, enquanto 4 ORFs permaneceram com função desconhecida. A organização das ORFs presentes no inserto fosmidial encontram-se ilustradas na Figura 19.

Tabela 10. ORFs anotadas do fosmideo 2B proveniente da biblioteca metagenômica de reservatório de petróleo.

ORF	Tamanho (aa)	INTERPRO /COG Referencias	BLASTP hit usada para anotação					Observações*
			Nome do Gene	Organismo	E-value	Acceso	Phylum /clase	
ORF 1	54	IPR023606 /COG1804	L- carnitina desidratase/Proteína induzível de ácidos biliares F	<i>Pseudomonas putida</i> GB1	1 e - 18	YP_001666428.1	Proteobacteria / Gammaproteobacteria	Acil-CoA transferasa / Carnitina desidratase
ORF 2	395	IPR006090, /COG1960	Acil-CoA desidrogenase da família III (Gcdh)	<i>Thauera</i> sp. MZ1T	0.0	YP_002354094.1	Proteobacteria / Betaproteobacteria	Acil-CoA desidrogenase
ORF 3	303	IPR005119, /COG0583	Regulador transcripcional HTH/LysR (IisR)	<i>Thauera</i> sp. MZ1T	6e-120	YP_002354093.1	Proteobacteria / Betaproteobacteria	Regulador transcripcional
ORF 4	239	IPR002925 /COG0412	Dienelactona hidrolase (dlh)	<i>Methylococcus capsulatus</i> str. Bath	3e-83	YP_114241.1	Proteobacteria / Gammaproteobacteria	Dienelactona hidrolase
ORF 5	306	IPR003399 /COG1463	Transportador tipo ABC, proteína periplasmática de ligação de substrato	<i>Arcobacter butzleri</i> RM4018	2,00E-28	YP_001489060.1	Proteobacteria / Delta/epsilonproteobacteria	Componente permease de um ABC transportador
ORF 6	196	IPR005586 /COG3218	Proteína de função desconhecida	<i>Nitrosococcus halophilus</i> Nc4	5e-13	YP_003526039.1	Proteobacteria / Gammaproteobacteria	Desconhecida
ORF 7	238	IPR005122	Uracil DNA glicosilases monofuncional de fita simples (UNG)	<i>Aromatoleum aromaticum</i> EbN1	9e-92	YP_158606.1	Proteobacteria / Betaproteobacteria	Uracil DNA glicosilase
ORF 8	370	IPR000064 /COG0791	Família de proteínas NLP/P60	<i>Pseudomonas fluorescens</i> Pf-5	7e-33	YP_262894.1	Proteobacteria / Gammaproteobacteria	Putativa hidrolases da parede celular associadas (proteínas associadas a invasão)
ORF 9	201	IPR003382 IPR004507 /COG0163	3-octaprenil-4-hydroxybenzoato carboxi-liase (ubiX)	<i>Burkholderia</i> sp. CCGE1001	4e-50	YP_004229638.1	Proteobacteria / Betaproteobacteria	3-octaprenyl-4-hydroxybenzoate carboxy-lyase
ORF 10	583	IPR006318 /COG1080	Proteína fosfoenol piruvato fosfotransferase (ptsI)	<i>Thauera</i> sp. MZ1T	0.0	YP_935284.1	Proteobacteria / Betaproteobacteria	Fosfoenol piruvato fosfotransferase
ORF 11	89	IPR000032 /COG1925	Sistema fosfotransferase, proteína fosfo-carrier HPr (ptsH)	<i>Thauera</i> sp. MZ1T	2e-32	YP_002354142.1	Proteobacteria / Betaproteobacteria	Sistema fosfotransferase, HPr.
ORF 12	266	IPR000073 /COG0596	Hidrolase de clivagem do anel benzoato	<i>Dechloromonas aromatica</i> RCB	2e-59	YP_286030.1	Proteobacteria / Betaproteobacteria	Alfa/beta hidrolase (bamA)
ORF 13	152	IPR002810	Proteína de função desconhecida DUF107	<i>Geobacter</i> sp. FRC-32	2e-28	YP_002538157.1	Proteobacteria / Delta/epsilonproteobacteria	Atividade putativa de regulador de protease de membrana YbbK

ORF 14	285	IPR001107 /COG0330	Proteína contendo domínio SPFH / família banda 7	<i>Dechloromonas aromatica</i> RCB	6e-124	YP_283489.1	Proteobacteria / Betaproteobacteria	Família de protease de membrana YbbK
ORF 15	640	COG0445	tRNA uridine 5-carboxymethylaminomethyl modification enzyme (GidA)	<i>Azoarcus</i> sp. BH72	0.0	YP_931646.1	Proteobacteria / Betaproteobacteria	Proteínas NAD de ligação envolvidas na adição de um grupo carboximetilaminometil no tRNA
ORF 16	222	IPR003682 /COG0357	rRNA metiltransferase (GidB)	<i>Azoarcus</i> sp. BH72	3e-45	YP_931647.1	Proteobacteria / Betaproteobacteria	Subunidade de rRNA pequeno metiltransferase
ORF 17	263	IPR002586 /COG1193	chromosome partitioning (parA)	<i>Cupriavidus taiwanensis</i> LMG 19424	7e-109	YP_002007082.1	Proteobacteria / Betaproteobacteria	PROPSEARCH: 1, 2 naftaleno dioxigenase (DIST: 7.71)
ORF 18	282	IPR003115 /COG1475	Proteína de partição parB-like (parB)	<i>Thauera</i> sp. MZ1T	2e-87	YP_002353898.1	Proteobacteria / Betaproteobacteria	Ligação ao DNA
ORF 19	559	IPR002078 /COG3284	Regulador transcricional sigma 54	<i>Rhodopseudomonas palustris</i> BisB18	1e-144	YP_531839.1	Proteobacteria / Alphaproteobacteria	Regulação da transcrição
ORF 20	393	IPR001670 /COG1464	Álcool desidrogenase que contém ferro (adh)	<i>Burkholderia gladioli</i> BSR3	0.0	AEA63300.1	Proteobacteria / Betaproteobacteria	Processo de oxidação e redução
ORF 21	506	IPR015590 /COG1012	Aldeído desidrogenase (aldh)	<i>Geobacillus</i> sp. G11MC16	0.0	ZP_03148146.1	Firmicutes	Processo de oxidação e redução
ORF 22	351	IPR002085 /COG1064	Álcool desidrogenase com domínio de ligação de zinco. (adh)	<i>Thermaerobacter marianensis</i> DSM 12885	5e-126	YP_004102555.1	Firmicutes	Processo de oxidação e redução Zinco- álcool desidrogenase de cadeia longa
ORF 23	91		Proteína hipotética MDMS009_82	<i>Methylophaga thiooxidans</i> DMS010	9e-34	ZP_05102946.1	Proteobacteria / Gammaproteobacteria	PROPSEARCH: 3-fenilpropionato dioxigenase (DIST: 12,09)
ORF 24	136	IPR008497 /COG3564	Proteína hipotética XCV0106	<i>Xanthomonas campestris</i> pv. vesicatoria str. 85-10	2e-53	YP_361837.1	Proteobacteria / Gammaproteobacteria	Desconhecida
ORF 25	272	IPR003439 /COG0411	Transportador ABC de aminoácidos de cadeia ramificada, proteína de ligação ao ATP	<i>Polaromonas</i> sp. JS666	2e-87	YP_552043.1	Proteobacteria / Betaproteobacteria	Transportador tipo ABC
ORF 26	654	IPR000873 /COG1022	Sintetase e ligase dependente de AMP	<i>Polaromonas</i> sp. JS666	0.0	YP_552044.1	Proteobacteria / Betaproteobacteria	Atividade ligase Família COG: Sintetase de acil-CoA de cadeia

								longa
ORF 27	295	IPR001851 /COG0559	Transportador tipo ABC de aminoácidos de cadeia ramificada, proteína permease	<i>Roseovarius</i> sp. 217	2e-79	ZP_01037671.1	Proteobacteria / Alfabroteobacteria	Sistema de transporte
ORF 28	357	IPR001851 /COG0559	Transportador tipo ABC de aminoácidos de cadeia ramificada, proteína permease	<i>Polaromonas</i> sp. JS666	3e-97	ZP_01037670.1	Proteobacteria / Betaproteobacteria	Sistema de transporte
ORF 29	439	COG0683	Transportador tipo ABC de aminoácidos, subunidade periplasmica de ligação ao substrato	<i>Roseovarius</i> sp. 217	2e-100	ZP_01037667.1	Proteobacteria / Alfabroteobacteria	Sistema de transporte
ORF 30	273	IPR003439 /COG0410	Transportador ABC relacionado	<i>Polaromonas</i> sp. JS666	1e-95	YP_552047.1	Proteobacteria / Betaproteobacteria	Sistema de transporte
ORF 31	146	IPR003736 /COG2050	Proteína de degradação do ácido fenilacético (PaaD)	<i>Magnetospirillum gryphiswaldense</i> MSR-1	1e-33	CAM75444.1	Proteobacteria / Alfabroteobacteria	Proteína não caracterizada, possivelmente envolvida no catabolismo de aromáticos
ORF 33	113	IPR005598	ATP sintetase, protein I (atpi)	<i>Polaromonas</i> sp. JS666	3e-08	YP_547183.1	Proteobacteria / Betaproteobacteria	Geração de energia
ORF 34	283	IPR000568 /COG0356	Subunidade A do complexo F0F1 ATP sintase (atpl)	<i>Thauera</i> sp. MZ1T	3e-129	YP_002353900.1	Proteobacteria / Betaproteobacteria	Geração de energia
ORF 35	81	IPR002379 /COG0636	Subunidade C do complex F0F1 ATP sintetase (atpE)	<i>Aromatoleum aromaticum</i> EbN1	3e-34	YP_158717.1	Proteobacteria / Betaproteobacteria	Geração de energia
ORF 36	157	IPR005864 /COG0711	Subunidade B do complex F0F1 ATP sintetase (atpF)	<i>Thauera</i> sp. MZ1T	3e-45	YP_002353902.1	Proteobacteria / Betaproteobacteria	Geração de energia
ORF 37	180	IPR000711 /COG0712	Subunidade delta do complexo F1-ATP sintetase (atpH)	<i>Aromatoleum aromaticum</i> EbN1	2e-46	YP_158719.1	Proteobacteria / Betaproteobacteria	Geração de energia
ORF 38	512	IPR000194 /COG0056	Subunidade alfa do complexo F0F1 ATP sintetase (atpA)	<i>Azoarcus</i> sp. BH72	0.0	YP_931662.1	Proteobacteria / Betaproteobacteria	Geração de energia
ORF 39	289	IPR000131 /COG0224	Subunidade gama do complexo F1 ATP sintetase (atpG)	<i>Thauera</i> sp. MZ1T	7e-135	YP_002353905.1	Proteobacteria / Betaproteobacteria	Geração de energia
ORF 40	465	IPR005722 /COG0055	Subunidade beta do complexo f0f1 ATP sintetase (atpB)	<i>Cupriavidus taiwanensis</i> LMG 19424	0.0	YP_002007073.1	Proteobacteria / Betaproteobacteria	Geração de energia
ORF 41	140	IPR001469 /COG0355	Subunidade epsilon do complexo F0F1 ATP sintetase (atpC)	<i>Thauera</i> sp. MZ1T	3e-53	YP_002353907.1	Proteobacteria / Betaproteobacteria	Geração de energia

*Observações sobre a função predita ou de anotação por outro método.

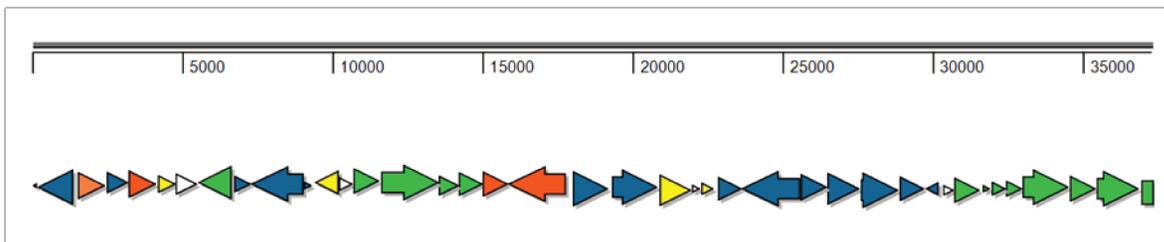


Figura 19. Representação esquemática do inserto totalmente sequenciado do fósideo 2B derivado da biblioteca metagenômica. As diferentes cores indicam as categorias funcionais dos genes codificando proteínas putativas. Setas azuis: ORFs relacionadas com metabolismo; setas verdes: ORFs relacionadas com processos celulares; setas cor de laranja: ORFs relacionadas com armazenamento da informação. Proteínas pouco caracterizadas e genes sem uma categoria COG relacionada são mostrados em amarelo e branco, respectivamente.

6.6.1. Organização genética do fósideo 2B

Nesta etapa do presente estudo, acreditava-se que o sequenciamento completo do fósideo 2B, que apresentou a habilidade de degradar hidrocarbonetos alifáticos (hexadecano) e aromáticos (fenantreno), permitiria encontrar os clusters de genes responsáveis pela degradação observada na expressão heteróloga do inserto metagenômico em *E. coli*. Um dos achados mais interessantes deste trabalho é a ausência de clusters de enzimas chave que mostrassem suficiente similaridade com as enzimas bem estudadas de degradação de hidrocarbonetos sob condições aeróbias ou anaeróbias. Diferentemente, os produtos de genes preditos para esse fósideo mostraram similaridade, em nível de família e superfamília, com grupos de enzimas que apresentam uma grande diversidade metabólica de funções, que podem incluir algumas etapas dos processos de degradação de xenobióticos em micro-organismos anaeróbios e aeróbios assim como sequências envolvidas em outros tipos de funções e reações celulares.

Na revisão bibliográfica deste trabalho se fez referência a um dos mecanismos considerados como o mais comum da degradação anaeróbia de compostos aromáticos, assim como de alifáticos em micro-organismos filogeneticamente diversos (Egland et al., 1997; Chakraborty e Coates 2004). O mecanismo consiste na adição de fumarato graças à ação de uma enzima glicil radical, sendo a mais conhecida a benzilsuccinato sintase, a qual tem sido descrita para as bactérias *Aromatoleum aromaticum* EbN1, *Thauera aromática* e *Geobacter metallireducens*. Entretanto, esta enzima não foi achada dentre as ORFs preditas para o fósideo 2B. A ausência desta enzima já foi considerada na análise do genoma completo de *Dechloromonas aromática* RCB por Salineiro e colaboradores (2009). Estes autores notaram que todas as enzimas chave na degradação anaeróbia de aromáticos, como a benzoilCoA redutase e a benzilsuccinato sintase, previamente denotadas como “centrais” para o catabolismo anaeróbio de aromáticos,

estavam ausentes no genoma de *D. aromática* e a maioria também ausentes em *Azoarcus* BH72. A carência de sobreposição dos genes que codificam estas enzimas entre estas espécies foi totalmente inesperada, uma vez que ambos *A. aromaticum* EbN1 e *D. aromática* são degradadores de compostos aromáticos (Salineiro et al., 2009).

Em outro estudo, Suenaga e colaboradores (2007) usaram uma estratégia de screening funcional para investigar o DNA metagenômico de lodos ativados utilizados no tratamento de uma planta de águas residuárias. O sequenciamento subsequente revelou que apenas um subconjunto de clones fosmidiais continha vias de degradação completas similares às encontradas em bactérias degradadoras de aromáticos. A maioria dos clones apresentou um subconjunto de genes de vias de degradação em novos rearranjos genéticos (Suenaga et al., 2009). Por outro lado, o fato de que compostos aromáticos no ambiente possam ser degradados através de uma ação combinada de várias vias fragmentadas tem sido também reportado para culturas puras. *P. putida* GJ31 é um exemplo de um organismo capaz de utilizar clorobenzeno como substrato de crescimento graças à expressão de um agrupamento de genes no plasmídeo pKW1 que foi descrito por constituir um mosaico de 4 clusters genéticos diferentes (Kunze et al., 2009).

Abu Laban e colaboradores (2010) utilizaram uma abordagem combinada de genômica e proteômica para o estudo da degradação anaeróbia de benzeno, utilizando um cultivo redutor de ferro altamente enriquecido com benzeno, fenol e benzoato. O DNA total do cultivo ferro redutor foi sequenciado e foram identificadas as enzimas e os reguladores transcricionais envolvidos na degradação anaeróbia de hidrocarbonetos aromáticos. Clusters de genes para algumas proteínas degradadoras de aromáticos foram identificados em posições adjacentes. No entanto, a maioria dos genes mostrou-se espalhada em 63 contigs diferentes do metagenoma.

Neste contexto, não foi totalmente surpreendente que no mesmo fragmento metagenômico, o fosmídeo 2B, foram observados componentes das vias de degradação aeróbia e anaeróbia, de diferentes compostos aromáticos, cíclicos e alifáticos e que serão descritos a seguir. Com o fim de explorar a aparente carência das vias de degradação amplamente conhecidas que se esperava encontrar no fosmídeo, todas as proteínas preditas e as regiões flanqueadoras de cada uma, foram comparadas com genes similares e, na medida do possível, também foram comparadas com regiões similares em organismos similares.

6.6.2. ORFs relacionadas às funções metabólicas de degradação aeróbia de aromáticos

A ORF 4 codifica para a enzima dienolactona hidrolase (DLH), a qual pertence à família dienolactone hidrolases. Estas enzimas apresentam um dobramento α/β e desempenham um papel fundamental na via de degradação do clorocatecol. Essa via envolve a clivagem do anel

aromático e é constituída pelas enzimas: catecol 1,2-dioxigenase, cloromuconato cicloisomerase, dienolactone hidrolase (DLH) e uma maleacetato redutase (MAR) (Marin e Pieper, 2009). Segundo Pieper (2005), os genes que codificam as enzimas da via clorocatecol têm sido evidenciados formando clusters e com as estruturas dos operons correspondentes altamente conservados, apesar das diferentes origens geográficas das bactérias ou diferenças na filogenia, como é o caso das bactérias isoladas *Pseudomonas putida* AC858, *Burkholderia* sp. NK8, *Wautersia eutropha* JMP134 (pJP4), *Burkholderia cepacia* 2a (pIJB1), *Delftia acidovorans* P4a, e *Rhodococcus opacus* 1CP (Figura 20). Diferentemente destas considerações, o fosmideo 2B apresentou uma ORF codificando a enzima DH que não se encontra flanqueada pelas outras enzimas presentes na via de degradação do clorocatecol, e até onde foi pesquisado, nenhum cluster similar tem sido reportado. No entanto, elementos regulatórios como da família LysR precedendo os genes catabólicos, são comuns nos clusters descritos por Pieper (2005) e também foram observados na região anexa à DH do fosmideo 2B, atuando provavelmente como regulador positivo (Figura 21).

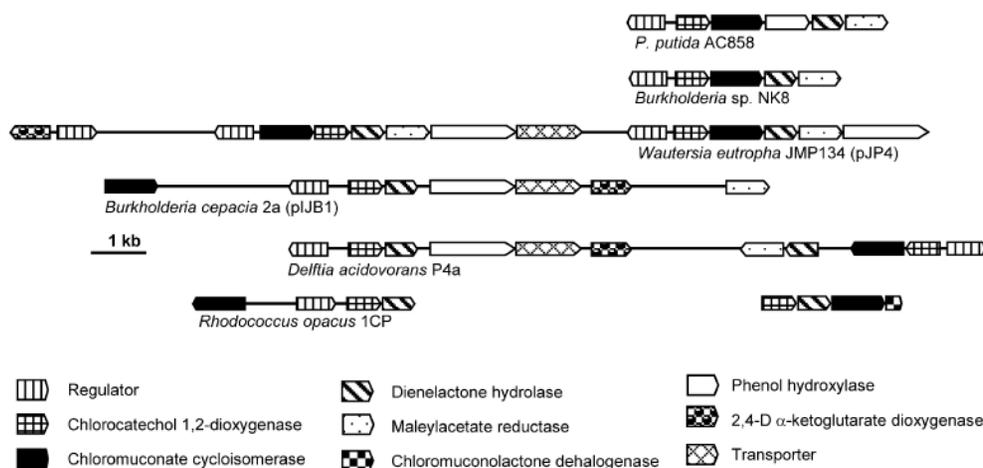


Figura 20. Organização genética do cluster de genes para degradação do clorocatecol *clcR-ABCDE* em *P. putida* AC858. Genes similares são observados em outras cepas de *Pseudomonas* (arranjos não mostrados), em *Burkholderia* sp. NK8, *Wautersia eutropha* JMP134 (pJP4), *B. cepacia* 2a (pIJB1), *Delftia acidovorans* P4a e em *Rhodococcus opacus* 1CP. Fonte: Pieper (2005).

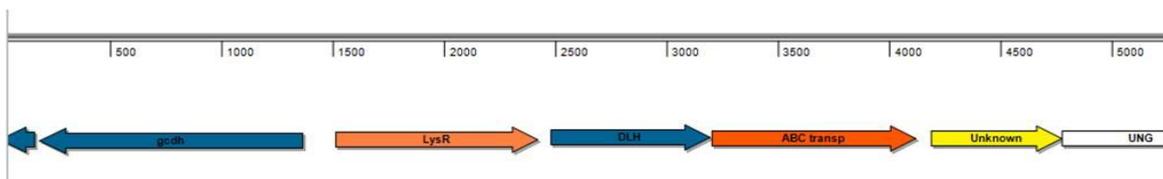


Figura 21. Mapa genético da ORF 4 e dos produtos da região flanqueadora encontrada no fósideo 2B. *gcdh*: glutarilCoA desidrogenase; *LysR*: regulador transcricional da família *LysR*; *DLH*: diene lactona hidrolase; *ABC transp*: proteína transportadora tipo ABC; *Unknown*: proteína desconhecida. As diferentes cores indicam as categorias funcionais dos genes. Azul: ORFs relacionadas com metabolismo, laranja: ORFs relacionadas com armazenamento da informação. Proteínas pouco caracterizadas e genes sem uma categoria COG relacionada são mostrados em amarelo e branco, respectivamente

O produto codificado pela ORF **31** foi similar à proteína de degradação do ácido fenilacético (*PaaD*). Fenilacetato deriva principalmente a partir da fenilalanina, mas é também formado nas vias de degradação de uma variedade de compostos aromáticos de origem ambiental e sintética (Song et al., 2006). A ORF classificou na categoria Q de COG e descrita como uma “proteína não caracterizada envolvida possivelmente no catabolismo de compostos aromáticos”. No banco de dados de domínios InterProScan, apresentou um domínio envolvido na degradação aeróbia de fenilacetato. Vários genomas microbianos contêm agrupamentos de genes codificando os componentes da via catabólica do fenilacetato (*PA*), no entanto, a evidência experimental da via funcional está disponível apenas para algumas bactérias: *Escherichia coli*, *Azoarcus evansii*, *Pseudomonas putida*, *A. aromaticum* EbN1 e *Rhodococcus* sp. (Law et al., 2008). Na maioria destes micro-organismos, o agrupamento de genes se dá no operon único *paaABCDEFGHIJKXYZ*, que codifica genes estruturais, regulatórios e de transporte (Figura 22).

O gene codificado pela ORF 31 mostrou alta similaridade com a região *Paal* tioesterase que pertence a subfamília *acilCoA* tioesterase. *AcilCoA* tioesterases geralmente funcionam dentro da célula para liberar a unidade de carboxilato para a degradação, a exportação, regulação e ou para possibilitar a liberação do *CoA* em metabolitos tioester e a incorporação em vias metabólicas alternativas. A família inclui a 4-hidroxibenzoilCoA tioesterase, que catalisa a etapa final da biossíntese de 4-hidroxibenzoato a partir de 4-clorobenzoato em *Pseudomonas* sp. CBS-3 por dehalogenação hidrolítica (Carmona et al., 2009) (Figura 23). A via consiste de três enzimas cujos genes que as codificam apresentam-se estruturalmente relacionados em clusters (Dunaway-Mariano e Babbit, 1994). O gene presente na ORF 31 não parece estar flanqueado por genes relacionados com as proteínas presentes na via de degradação do fenilacetato nem por nenhum dos genes envolvidos na via de degradação do 4-clorobenzoato (Figura 24). Portanto, esta ORF poderia codificar uma proteína envolvida no metabolismo de aromáticos que ainda não foi descrita para outro substrato quimicamente relacionado com os anteriores.

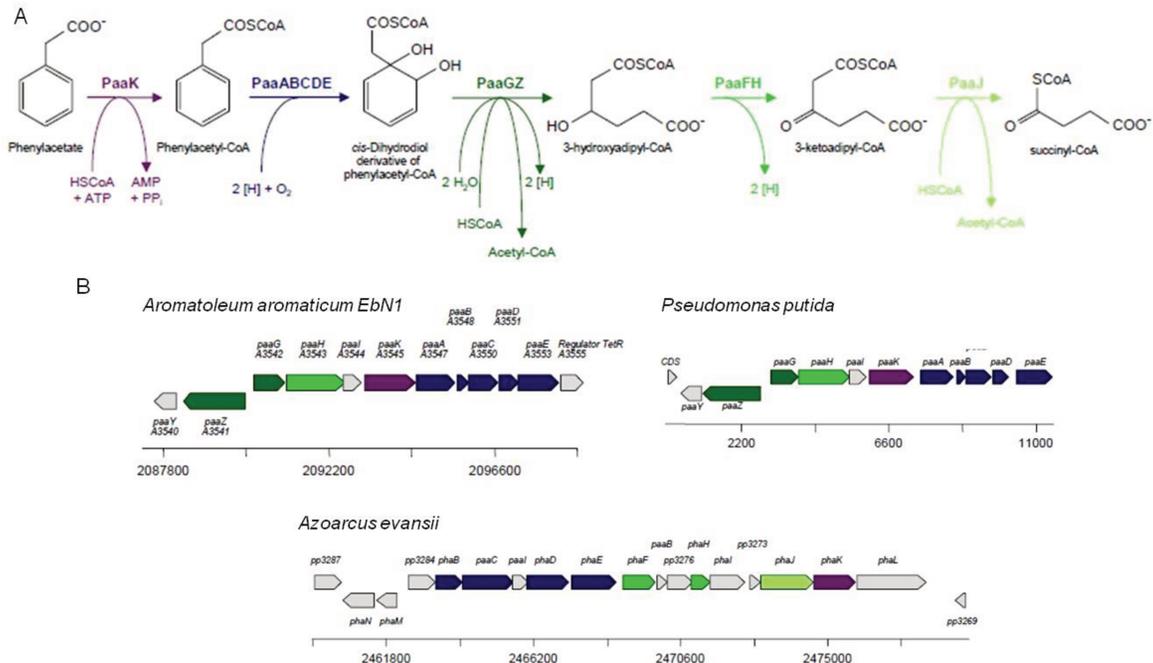


Figura 22. Via de degradação aeróbia do fenilacetato em *A. aromatoleum* EbN1. (A) via; (B) cluster de genes da via em diferentes organismos. Paak, fenilacetatoCoA ligase; PaaABCDE, fenilacetilCoA dioxigenase; PaaGZ, abertura do anel (PaaG, enoilCoA hidratase; PaaZ, aldeído dehidrogenase); PaaFHJ, envolvido na β -oxidação. Fonte: Rabus et al. (2005).

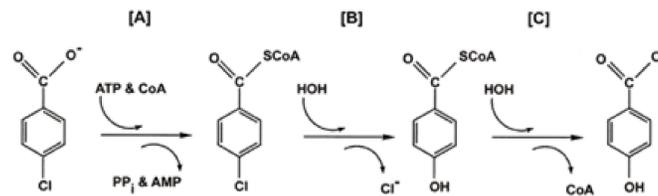


Figura 23. Via de dehalogenação hidrolítica de 4-clorobenzoato. A, 4-clorobenzoilCoA ligase; B, 4-clorobenzoilCoA dehalogenase; C, 4-hidroxi benzoilCoA tioesterase. Fonte: Thoden et al. (2003).

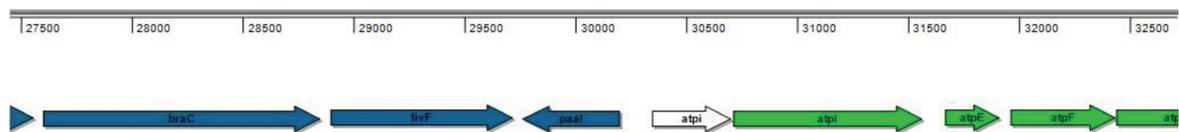


Figura 24. Mapa genético da ORF 31 e genes da região flanqueadora encontrada no fosmideo 2B. braC, Transportador tipo ABC (proteína de ligação a Leucina-, isoleucina-, valina-, treonina-, e alanina); livF, Transportador tipo ABC (Transporte de aminoácidos de cadeia ramificada – proteína de ligação ao ATP); PaaI, proteína de degradação do ácido fenilacético; atpi, ATP sintase putativa; atpl, subunidade A de ATP sintase; atpE, subunidade C de ATP sintase; atpF subunidade B de ATP sintase. As cores representam as categorias COG dos genes. Azul: ORFs relacionadas com o metabolismo, verdes: ORFs relacionadas com processos celulares e sinalização, branco: ORFs sem nenhuma categoria COG relacionada.

As ORFs **20**, **21** e **22** codificam para as enzimas que oxidam os alcoóis primários gerados na oxidação terminal de alcanos. Álcool desidrogenases (ADHs) (ORFs 20 e 22) desempenham uma grande variedade de funções fisiológicas e apresentam uma ampla série de substratos. No contexto da biodegradação de alcanos, ADHs catalisam a oxidação reversível de alcoóis a aldeídos, a segunda etapa da via de oxidação terminal de alcanos (Wentzel et al., 2007). Muitas bactérias contêm diferentes ADHs que podem ser utilizadas para a assimilação de diferentes classes de alcoóis (Rojo, 2010). ADHs podem ser divididas em três grandes grupos com base em suas propriedades moleculares e teor de metais. O grupo I inclui ADHs de cadeia longa e requer zinco como cofator (aproximadamente 350 aminoácidos por subunidade). O grupo II inclui ADHs de cadeia curta livres de metal (aproximadamente 250 aminoácidos por subunidade) e o grupo III de cadeia longa que requer do cofator Fe (385-900 aminoácidos por subunidade) (Liu et al., 2009). A ORF 20 contém 394 aminoácidos e mostrou 82% de identidade com uma álcool desidrogenase que contém ferro de *Burkholderia gladioli* BSR3. A ORF 22 contém 352 aminoácidos e mostrou 63% de identidade com uma álcool desidrogenase com domínio de ligação ao zinco de *Thermaerobacter marianensis* DSM 12885. Com base nos dados sequenciados, a ORF 20 pertence à família de ADHs que contem Ferro e a ORF 22 à ADHs que contêm zinco como cofator.

O terceiro passo da via de oxidação de alcanos é catalisado por enzimas aldeído desidrogenase (ALDHs), que catalisam a oxidação de aldeídos a ácidos carboxílicos usando NAD ou NADP como coenzima. A ORF **21** mostrou a maior similaridade (82%) com uma aldeído desidrogenase de *Geobacillus* sp. G11MC16. ALDHs estão amplamente distribuídas entre os organismos vivos e atuam como enzimas de desintoxicação de aldeídos tóxicos gerados em vários metabolismos celulares, e também são reconhecidas como uma das enzimas essenciais para o processo de degradação de vários compostos de hidrocarbonetos.

Genes que codificam para ADHs e ADLHs estão presentes no plasmídeo OCT de *P. putida* GPo1 (Figura 25). O plasmídeo contém todos os componentes necessários para a degradação de alcanos. Apesar de que essa via foi claramente transferida horizontalmente para muita bactérias (van Beilen et al., 2001), a ligação de genes foi perdida em diferentes graus para varias cepas. Além disto, muitas bactérias contem diferentes ADH que podem ser utilizadas para a oxidação de álcoois distintos. Por exemplo, no plasmídeo OCT a álcool desidrogenase denominada como AlkJ é necessária só se a álcool desidrogenase cromossomal é inativada por mutação, sugerindo novamente a redundância destas enzimas em um mesmo organismo (Rojo, 2010). Similarmente, a aldeído desidrogenase codificada pelo plasmídeo AlkH não é necessária para o crescimento na presença de alcanos (Rojo, 2009)

P. putida (oleovorans) GPo1

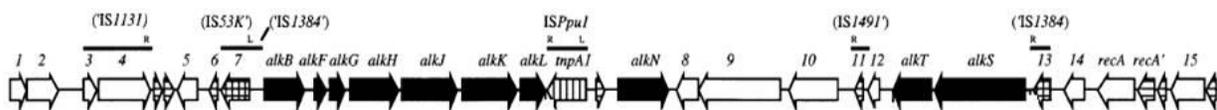


Figura 25. Mapa genético do plasmídeo OCT de *P. putida* GPo1 mostra a presença no cluster dos genes *alkJ* e *alkH* codificando para uma ADH e uma ALDH, respectivamente no operon *alkBFGHJKL*. Tomada e modificada de van Beilen, et al., 2001.

Os ácidos carboxílicos gerados na oxidação de aldeídos são a seguir metabolizados por β -oxidação, gerando AcilCoA que entra no ciclo do Ácido Tricarboxílico. Uma parte da via de β -oxidação foi encontrada no fosmídeo 2B e se encontra descrita posteriormente no item 6.6.4. No entanto, as enzimas mais conhecidas codificantes do primeiro passo da degradação aeróbia de alcanos (monooxigenases) não foram identificadas no fragmento metagenômico estudado.

Interessantemente, a aldeído desidrogenase do fosmídeo 2B faz parte de um operon que mostra similaridade com o operon de *Rhodoferrax ferriducens* DSM 15236, o qual codifica para uma putativa álcool desidrogenase e uma aldeído desidrogenase reguladas à montante por um regulador transcricional da família Fis (Figura 26).

Fos 2B



Rhodoferrax ferriducens DSM 15236

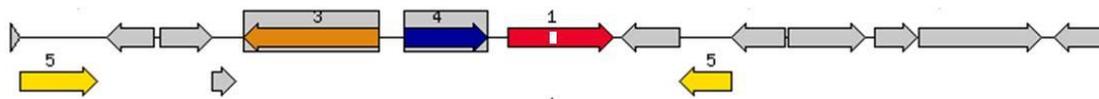


Figura 26. Organização genética dos clusters do fosmídeo 2B (acima) e o cluster presente na bactéria *R. ferriducens* DSM 15236 (abaixo). Nas cores, laranjas, azuis e vermelhas, apresentam-se as ORFs similares. Em amarelo, ORFs presentes na bactéria *R. ferriducens* DSM 15236, codificando proteínas de transporte com ligação ao ATP, com função similar à proteína LivG no fosmídeo 2B envolvida no transporte de aminoácidos de cadeia ramificada. Nas tonalidades cinza se mostram as ORFs que não apresentaram similaridade entre os dois clusters. Genômica comparativa gerada no programa RAST (Aziz et al., 2008).

Rhodoferrax ferriducens DSM 15236 é um anaeróbio facultativo que pertence ao phylum Proteobacteria e que acopla a oxidação de várias moléculas orgânicas (entre elas, acetato, piruvato, succinato e benzoato) à redução de ferro, usando ferro insolúvel como receptor de elétrons. O genoma completo deste micro-organismo já foi sequenciado pelo DOE Joint Genome Institute e se encontra sob estudo, dado o potencial biotecnológico deste micro-organismo que,

adicionalmente, é capaz de gerar eletricidade em células de combustível microbianas (Finneran et al., 2003).

6.6.3. ORFs relacionadas às funções metabólicas de degradação anaeróbia de aromáticos

A ORF **12** foi anotada como uma hidrolase de clivagem do anel benzoato e pertence à família de enzimas com um dobramento estrutural $\alpha\beta$ hidrolase, um dos maiores grupos de enzimas estruturalmente relacionadas com diversas funções catalíticas. Membros desta família incluem lipases, acetilcolinestases, esterases, lactonases, éóxido hidrolases e outros.

A hidrolase de clivagem do anel benzoato é uma enzima envolvida no metabolismo anaeróbio do benzoato. Especificamente, a hidrolase faz parte da terceira etapa da via, denominada como β oxidação modificada (Wischgoll et al., 2005).

Em anaeróbios facultativos, o catabolismo anaeróbio do benzoato tem sido estudado em nível molecular, como é o caso das bactérias desnitrificantes *T. aromatica*, *Azoarcus* sp. (betaproteobacterias), e *Magnetospirillum* sp. (alfaproteobacteria); da bactéria fotossintética *R. palustris* (alfaproteobacteria); a estritamente anaeróbia redutora de ferro, *G. metallireducens* e da fermentativa *S. aciditrophicus* (deltaproteobacteria) (Carmona et al., 2009). Em todas essas bactérias a via de degradação do benzoato envolve várias etapas e os genes codificando cada uma dessas etapas se encontram agrupados em clusters. A primeira etapa consiste na ativação do benzoato para benzoilCoA pela ação de uma benzoilCoA ligase dependente de ATP (Egland et al., 1995). A segunda etapa, é a redução do anel de benzeno pela enzima benzoil-CoA redutase, que catalisa a desaromatização do benzoilCoA a um dieno (Boll e fuchs 1995; Boll et al., 2000). A terceira etapa é a β oxidação modificada do dieno desaromatizado, resultando na clivagem do anel. O produto da desaromatização é um alifático C7-dicarboxilCoA, pela ação de hidratases, desidrogenases e hidrolases. A quarta e última etapa, consiste na β oxidação convencional a 3 acetilCoA e CO_2 (Figura 27).

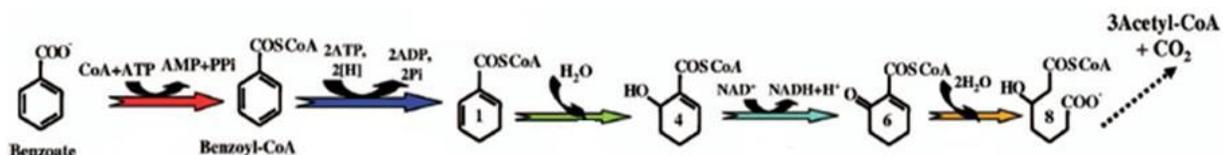


Figura 27. Via enzimática de degradação anaeróbia do benzoato. Em vermelho se mostra a ativação do benzoato a benzoilCoA; em azul escuro, a desaromatização do benzoilCoA; em verde, as diferentes etapas da β oxidação modificada (hidratação); em azul claro, a desidrogenação e em laranja, a hidroxilação. Adaptado de Carmona et al., 2009.

Como foi mencionado na revisão deste trabalho, a ativação de hidrocarbonetos aromáticos por meio de CoA ligases é um mecanismo alternativo e, no caso referido acima, a degradação anaeróbia do benzoato começa com a ativação do benzoato via uma benzoilCoA dependente de ATP. Os genes que codificam para estas benzoato CoA ligases têm sido descritos com diferentes nomes em várias bactérias, *badA* em *R. palustris*, *bclA* em *T. aromática* e *Magnetospirillum* sp. e *bzdA* em *Azoarcus* sp. A ORF **26** predita no inserto metagenômico 2B objeto do presente estudo, corresponde a um domínio de ligase dependente de AMP, a presença desta ORF adjacente à ORF **12** que codifica para outras etapas da degradação do benzoilCoA indica que estas ORFs poderiam estar relacionadas, sendo que a ORF 26 atuaria na ativação de compostos aromáticos e alicíclicos (Figura 28).

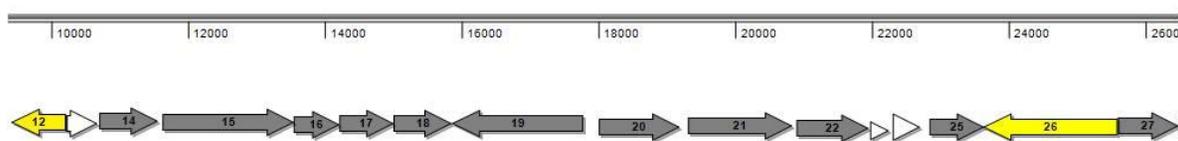


Figura 28. Organização genética do fósideo 2B. As ORFs se encontram enumeradas de acordo com a Tabela 6. Os genes codificando putativos componentes enzimáticos da via de degradação anaeróbia do benzoato estão em amarelo. Em branco são mostradas as ORFs de função desconhecida e em cinza as ORFs envolvidas em outras vias metabólicas.

Notavelmente, não se encontraram homólogos aos genes que codificam para as 4 subunidades da benzoilCoA redutase, o que poderia sugerir o envolvimento de um mecanismo enzimático diferente para a desaromatização do anel de benzeno. Em Wischgoll e colaboradores (2005) descreve-se o estudo do metabolismo aromático do micro-organismo estritamente anaeróbio, *G. metallireducens*. Os resultados obtidos nesse trabalho sugerem que a ativação do benzoato nas reações de β oxidação modificada são altamente similares às descritas para outros organismos anaeróbios facultativos e no organismo modelo estritamente anaeróbio *G. metallireducens*. No entanto, também foi observada a ausência da enzima benzoilCoA redutase, similarmente ao fósideo 2B sob estudo. Estes dados sugerem que os processos de desaromatização para *G. metallireducens* e provavelmente também para o fósideo 2B permanecem desconhecidos.

Fuchs (2008) considera que o metabolismo de compostos aromáticos pode ser classificado em 4 tipos: (1) Metabolismo aeróbico caracterizado pelo uso de oxigênio como co-substrato de oxigenases que hidroxilam e clivam os anéis aromáticos; (2) Na presença de oxigênio, os aeróbios facultativos utilizam outro tipo de metabolismo aeróbio denominado como tipo híbrido, onde a coenzima A é usada como substrato e não há a necessidade de oxigênio para a clivagem do anel, ao contrário, eles usam uma oxigenase/redutase para desaromatizar o anel; (3) Na ausência de

oxigênio, aeróbios facultativos e fototróficos usam a redução do anel aromático de benzoil-CoA catalisada por uma benzoil-CoA redutase, conduzida pela hidrólise de 2 ATP; e (4) Uma via completamente diferente ainda pouco caracterizada de benzoil-CoA que opera em anaeróbios estritos, não dependente de ATP.

As vias que envolvem a formação de tioésteres de CoA (2, 3, e 4) podem ser vantajosas, pois permitem uma flexibilidade e rápida adaptação à flutuação dos níveis de oxigênio, uma vez que ambas as condições (aeróbia e anaeróbia) requerem substratos tioésteres de CoA. Uma mudança de condição anaeróbia à aeróbia inativaria a enzima benzoil-CoA redutase e provocaria a acumulação do tioéster de CoA, o qual poderia ser consumido pela benzoil-CoA oxigenase/redutase uma vez que essa enzima é ativa em condições microaerofilicas. Além disso, a formação de intermediários éster CoA podem ser menos tóxicos que algumas rotas de clivagem aeróbias. Adicionalmente, a formação de tioéster facilita indiretamente o transporte de ácidos aromáticos.

A ORF 9 predita mostrou a maior similaridade com a enzima 3-octaprenil-4-hidroxi benzoato carboxi-liase (genes *ubi*), uma enzima da biossíntese de ubiquinona (Breinig et al., 2000). Esta enzima é considerada semelhante à fenilfosfato carboxilase (genes *ppc*), envolvida na degradação anaeróbia do fenol em *T. aromática* (Kube et al., 2005). Recentemente, a presença da enzima 3-octaprenil-4-hidroxi benzoato carboxi-liase tem sido evidenciada na expressão diferencial da cepa NaphS2 uma deltaproteobacteria quando crescida em naftaleno (Didonato et al., 2010; Bergmann et al., 2010). Em 2005, Kube e colegas utilizaram a abordagem metagenômica para estudar a comunidade microbiana presente no Mar Negro. O sequenciamento completo de um fosmídeo da biblioteca que exibiu genes de degradação anaeróbica do benzoato revelou a presença à jusante de genes *ubi*, o que foi considerado pelo grupo como um indicativo da possibilidade que o organismo correspondente à sequência metagenômica poderia ser capaz de utilizar o fenol anaerobicamente. Kube sugere a entrada do fenol na via central do benzoilCoA através da enzima 4-hidroxi benzoil-CoA redutase, levando à degradação do fenol pela via do benzoilCoA utilizada para a degradação do benzoato.

Outra abordagem para o estudo dos genes *ubi* foi feita por Abu Laban e colaboradores (2010), que investigaram os genes putativos para o metabolismo do benzeno através de uma análise combinada de metagenômica e proteômica em um cultivo ferro redutor enriquecido. Estes autores consideraram os 3 mecanismos iniciais para a degradação de benzeno, o primeiro pela metilação direta a tolueno e subsequente degradação via benzilsuccinato, o segundo, pela hidroxilação a fenol e subsequente degradação pela via benzoilCoA, e o terceiro, pela carboxilação direta a benzoato com a subsequente ativação com a benzoilCoA ligase e degradação via benzoilCoA. Neste mesmo estudo, as proteínas carboxilases UbiD/UbiX (3-octaprenil-4-

hidroxibenzoato carboxi-liase) foram diferencialmente expressas em cultivo em fenol e benzeno, e as ORFs que as codificam foram encontradas agrupadas junto com as enzimas fenilfosfato carboxilases (genes *ppc*) (Abu Laban et al., 2010). No entanto, outros contigs que apresentaram os genes *ubiD/ubiX* não apresentaram ORFs similares a *ppcB* e *ppcC*, necessárias para a carboxilação e degradação do fenol, o que levou os autores a sugerir um novo mecanismo inicial de ativação anaeróbia do benzeno por carboxilação direta mediada por uma enzima carboxilase candidata que provavelmente pertence à família da UbiX, e que leva à entrada do benzeno na via de degradação de benzoato e benzoilCoA. (Figura 29). Neste contexto, a presença da enzima carboxilase UbiD/UbiX (ORF 12) no mesmo contig que a putativa CoA ligase (ORF 26), também poderia sugerir uma carboxilação direta do benzeno pelo produto da ORF 12 e entrada na via de degradação de benzoilCoA pela enzima da ORF 26.

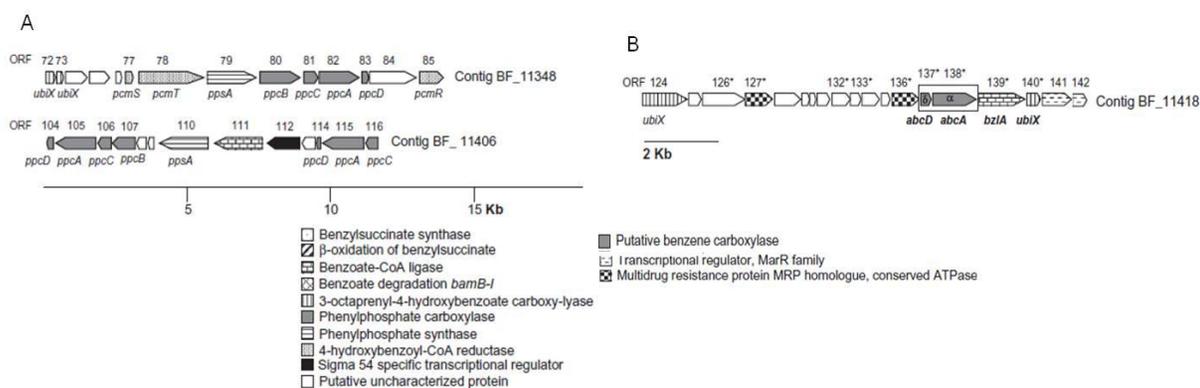


Figura 29. Genes identificados no metagenoma do cultivo redutor de ferro enriquecido em Abu Laban (2010). As ORFs codificam para proteínas putativas envolvidas com a degradação anaeróbia de aromáticos. (A) Diferentes contigs no metagenoma envolvidos com a degradação de fenol e hidroxibenzoato; (B) Cluster de genes putativo envolvido com a degradação de benzeno proposto para o cultivo redutor de ferro.

6.6.4. ORFs relacionadas a outras características metabólicas

Os genes transportadores estão freqüentemente presentes em agrupamentos de genes do catabolismo de aromáticos. As **ORFs 5, 25, 27, 28, 29 e 30** codificam para proteínas componentes do sistema transportador do tipo ABC. Os transportadores ABC utilizam a hidrólise de ATP para energizar diversos sistemas biológicos. Estão envolvidos no transporte de vários substratos, abrangendo desde pequenos íons a macromoléculas e, desta maneira, provêem as bactérias de nutrientes.

A captação de benzoato em bactérias aeróbias tem sido sugerida através do transportador Benk. Embora nada se saiba sobre a captação de benzoato em degradadores anaeróbios, os

transportadores tipo ABC foram encontrados no sequenciamento completo da bactéria anaeróbia desnitrificante degradadora de aromáticos *A. aromatoleum* EbN1, onde, se presume, são usados para a captação de benzoato a baixas concentrações (Rabus et al., 2005). Adicionalmente, à jusante do cluster de genes para o metabolismo de benzoato em *Azoarcus* sp. CIB (Barragan et al., 2004) e *R. palustris* (Egland et al., 1997), foram encontrados 5 genes que codificam para proteínas da família de transportadores ABC de aminoácidos de cadeia longa e diferentes subunidades de ABC transportadores. Embora esses genes que codificam para transportadores tipo ABC tenham sido encontrados na vizinhança dos clusters para a degradação de compostos aromáticos, o papel de tais transportadores na captação dos compostos aromáticos ainda não foi demonstrado (Barragan et al., 2004).

As ORFs **1**, **2** e **3** apresentam genes que codificam para enzimas da β oxidação (Figura 30). No entanto, devido à diversidade metabólica de muitas destas enzimas, é esperado que alguns dos produtos destes genes participem em outras reações (Rabus et al., 2005). A ORF 1 codifica para uma CoA transferase da família III, que catalisa reações de transferência de grupos da Coenzima A para ácidos livres. A ORF 1 apresenta maior similaridade com uma L-carnitina desidratase/ Proteína induzível de ácidos biliares. Esta enzima pertence à família CoA transferase III, descrita por Heider em 2001 precisamente por ser uma CoA transferase envolvida nas vias metabólicas anaeróbias de degradação de tolueno, carnitina e ácidos biliares, as quais não apresentavam semelhança com as CoA transferases das famílias I e II.

A ORF **2** contém os genes para uma Acyl-CoA desidrogenase, enzima claramente correlacionada com o metabolismo de substratos aromáticos e alifáticos. A enzima glutaryl-CoA desidrogenase (GDH) já tem sido caracterizada nas espécies de *Azoarcus evansii* e *Thauera aromática*, onde se encontram envolvidas na via metabólica do catabolismo anaeróbio de aromáticos (Rabus et al., 2005). Wischgoll e colaboradores (2009) estudaram as GDHs dos micro-organismos anaeróbios estritos degradadores de aromáticos *Geobacter metallireducens* (reduzidor de Fe[III]) e *Desulfococcus multivorans* (reduzidor de sulfato). Estes autores demonstraram que as GDHs eram responsáveis pela descarboxilação oxidativa de glutarilCoA a crotonilCoA

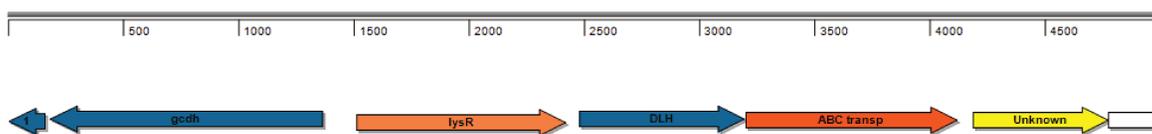


Figura 30. Esquema de parte da via da β -oxidação presente em um primeiro segmento do fosmideo 2B. 1, CoA transferase III; gcdh, glutaryl-CoA desidrogenase; lysR, regulador transcricional LysR; DLH, dienolactona hidrolase; ABC transp, transportador do tipo ABC; Unknown, proteína de função desconhecida.

A ORF **3** foi anotada como proteína LysR, a qual pertence a uma subfamília de reguladores transcricionais que se ligam ao DNA através do motivo hélice-volta-hélice (helix-turn-helix), e atuam, na sua maioria, como ativadores da transcrição. Na literatura, já existe evidência experimental de que a expressão do gene *gcdH* que codifica para a enzima glutarilCoA desidrogenase (ORF 2) se encontra sob o controle transcricional de LysR (Carmona et al., 2009).

As ORFs **10** e **11** contêm genes que codificam proteínas para o sistema fosfotransferase fosfoenol piruvato dependente de açúcar (PTS), que constituem o maior sistema de transporte de carboidratos em bactérias (Figura 31). O sistema fosfoenol piruvato fosfotransferase (PTS) inclui uma fosfotransferase (ORF 10) que catalisa a fosforilação dos substratos de açúcar que entram na célula concomitantemente com sua translocação através da membrana celular. A proteína portadora de fosfatos que contem histidina (Histidine-containing phosphocarrier system - HPr) (ORF 11) é o componente do sistema PTS que transfere carboidratos através da membrana celular.

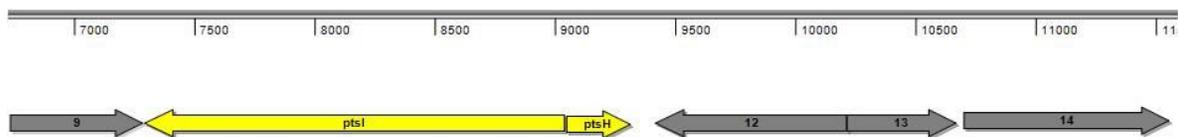


Figura 31. Esquema do sistema PTS e regiões flanqueadoras no fósideo 2B. Em amarelo se mostram os genes que codificam os componentes do sistema. Em cinza, as regiões flanqueadoras designadas com o numero descrito para cada ORF na Tabela 6.

6.6.5. ORFs relacionadas à respiração e metabolismo energético

Todos os componentes da ATP sintase estão presentes no inserto metagenômico 2B e estão organizados em uma estrutura única do tipo operon (Figura 32). As ATPases são complexos de enzimas ligados à membrana que combinam a síntese de ATP com o transporte de prótons através da membrana. As ATPases podem aproveitar a energia gerada a partir de um gradiente de prótons, utilizando o efluxo de íons através da membrana via o canal de próton ATPase para conduzir a síntese de ATP.

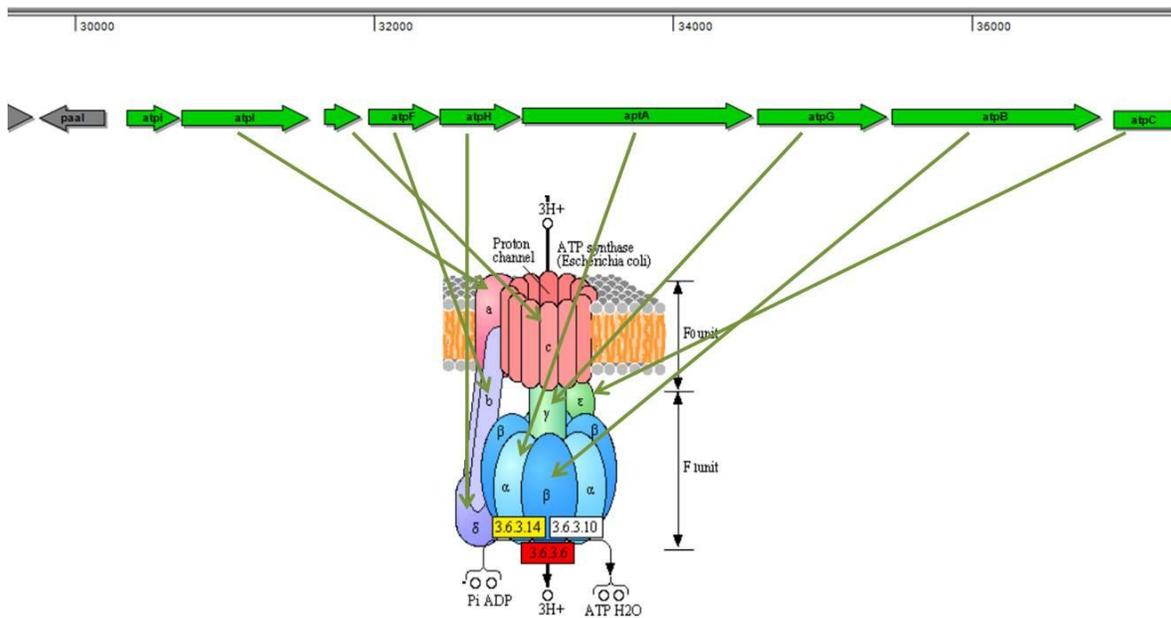


Figura 32. Organização do agrupamento de genes que codificam para o complexo da ATP sintase no fosmideo 2B e a conformação final dos seus respectivos produtos. *AtpI*, Proteína I ATP sintase; *atpI*, subunidade A do complexo F₀; *atpE*, subunidade C do complexo F₀; *atpF*, subunidade B do complexo F₀; *atpH*, subunidade delta do complexo F₁; *atpA*, subunidade alfa do complexo F₁; *atpG*, subunidade gamma do complexo F₁; *atpB*, subunidade beta do complexo F₁; *atpC*, subunidade delta do complexo F₁; α , subunidade alfa; β , subunidade beta; γ , subunidade gamma; ϵ , subunidade epsilon; δ , subunidade delta.

6.6.6. Genes relacionados a outras características e funções celulares

A ORF 7 codifica para a proteína Uracil DNA glicosilase, componente importante nas reações de reparo do DNA. A sua função é remover resíduos de uracil por hidrólise, iniciando a via de reparo por excisão de bases.

Na ausência de marcadores filogenéticos estabelecidos, como o RNA ribossomal 16S, outros genes, principalmente aqueles associados ao processamento da informação, têm sido sugeridos para a reconstrução da evolução de um organismo, dado que esses genes são pouco afetados por transferência horizontal (Jain et al., 1999). A afiliação do organismo que deu origem ao inserto metagenômico sob estudo foi baseada no filo a que pertence o melhor hit atribuído para cada ORF. Posteriormente, o inserto metagenômico inteiro foi designado ao filo mais comumente encontrado (Vide Tabela 10). De acordo com isto, o fosmideo 2B mostrou-se mais afiliado ao filo Proteobacteria, com 95% das ORFs atribuídas em este filo, e o restante (5%) ao filo Firmicutes.

Em estudos prévios, a sequência da subunidade β de ATP sintase tem sido estudada como uma ferramenta alternativa de marcador filogenético (Wolfgang et al., 1998). Este marcador foi usado por Pope e Patel (2007) para a afiliação filogenética de diversos clones derivados do metagenoma de um *bloom* de cianobactérias. No presente estudo, foi observado que a ORF 40 codifica para uma proteína de 465 aminoácidos com 91% de identidade de sequência com a subunidade β de ATP sintase da espécie *Cupriavidus taiwanensis*, pertencente à classe Betaproteobacteria. A sequência dessa proteína foi utilizada no presente estudo com o intuito de esclarecer a afiliação do inserto fosmidial 2B dentro do filo Proteobacteria. Uma árvore filogenética foi gerada usando o alinhamento de sequências homólogas à sequência da subunidade β da ATP sintase de alguns representantes de outras classes do filo Proteobacteria. Na Figura 33 observa-se a reconstrução filogenética para esse gene que mostrou ser consistente com a classificação filogenética em nível de classe no filo Proteobacteria. A sequência nucleotídica da β ATP sintase do inserto metagenômico (FOS_2B_ATP) mostrou-se mais similar às sequências da classe Betaproteobacteria. Estes resultados corroboraram o que já tinha sido deduzido pela porcentagem de ORFs designadas a esse grupo, ou seja, que o inserto metagenômico poderia ser originado de um organismo muito relacionado com a classe Betaproteobacteria.

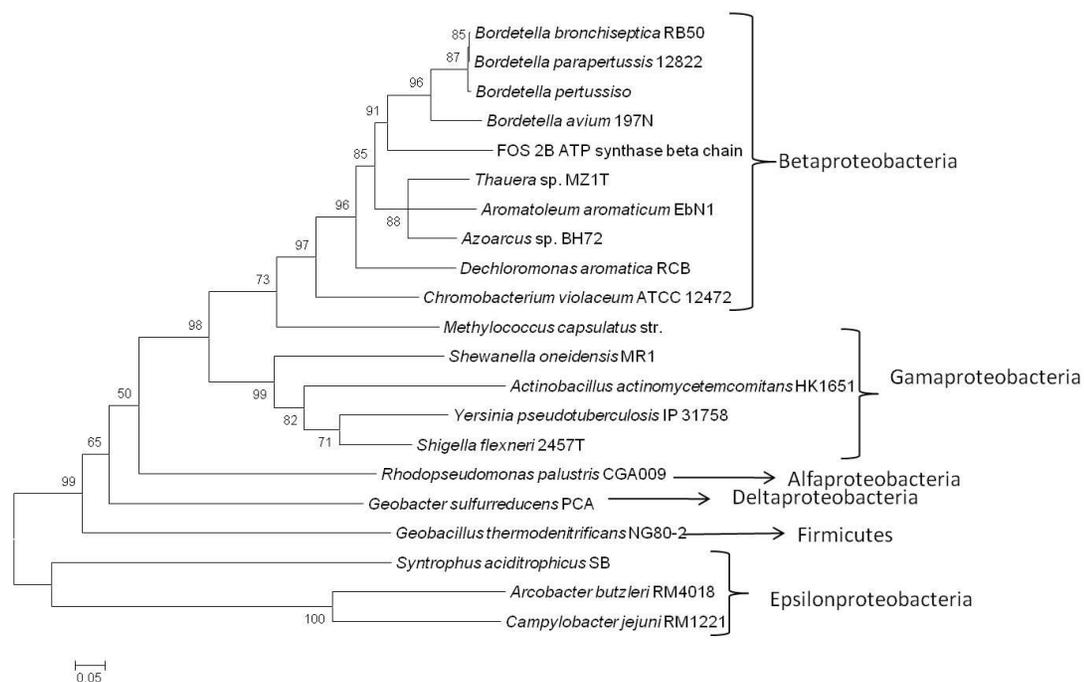


Figura 33. Análise filogenética baseada na sequência completa do gene da subunidade β da ATP sintase presente no inserto metagenômico do fosmideo 2B e de genes homólogos de organismos conhecidos contidos nas bases de dados.

6.7. Anotação do fosmídeo 1A

Similarmente ao fosmídeo 2B, diversas ferramentas foram utilizadas para a predição de proteínas no inserto fosmidial 1A. Um total de 23 ORFs foram detectadas, e foram atribuídas funções para 21 delas, enquanto 2 permaneceram com função desconhecida. A Tabela 11 apresenta a lista completa de cada ORF predita, a categoria COG à que pertence, assim como a informação do tamanho da proteína predita e referências das proteínas mais similares nas bases de dados. A organização de algumas ORFs presentes no contig de maior tamanho montado para o inserto fosmidial 1A está esquematizada na Figura 34.



Figura 34. Representação esquemática do maior contig do fosmídeo 1A. As diferentes cores indicam as categorias funcionais dos genes que codificam proteínas putativas. Setas azuis indicam ORFs relacionadas com metabolismo; setas verdes, ORFs relacionadas com processos celulares; e setas brancas, ORFs sem categoria COG relacionada.

Tabela 11. ORFs anotadas a partir dos dados de montagem do fosmídeo 1A proveniente da biblioteca metagenômica de reservatório de petróleo.

ORF	Tamanho (aa)	Categoria COG	BLASTP hit usada para anotação					Observações
			Nome do Gene	Organismo	E-value	Acceso	Phylum /clase	
C_1.1	351	R - COG1064	Dominio de álcool desidrogenase contendo zinco	<i>Thauera sp.</i> MZ1T	3,E-147	YP_002355505.1	Betaproteobacteria	
C_1.2	137	O - COG1764	Proteína de resistencia a hidroperóxidos orgânicos	<i>Ketogulonicigenium vulgare</i> Y25	6,E-52	YP_003962763.1	Alfaproteobacteria	
C_13	501	E - COG1982	Arginina descarboxilase; Ornitina descarboxilase; Lisina descarboxilase	<i>Azoarcus sp.</i> BH72	0.0	YP_934701	Betaproteobacteria	
C_14	154	R - COG1741	Proteína tipo Pirin YhaK	<i>Escherichia coli</i> O157:H7 str. 1044	7,E-57	EGD71100.1	Gammaproteobacteria	Proteína da superfamília Cupin
C_16	302	C - COG1012	Succinato-semialdeido desidrogenase [NAD]; [NADP+]	<i>Shewanella putrefaciens</i> 200	2,E-119	ADV56194.1	Gammaproteobacteria	
C_18.1	231	TK - COG0745	Regulador de resposta, de ligação ao DNA CoIR	<i>Azotobacter vinelandii</i> DJ	2,E-64	YP_002798968.1	Gammaproteobacteria	
C_18.2	419	T - COG0642	Proteína sistema sensor de dois componentes	<i>Sideroxydans lithotrophicus</i> ES-1	6,E-36	YP_003522891.1	Betaproteobacteria	
C_18.3	149	-	Proteína hipotética	<i>Dechloromonas aromatica</i> RCB	3,E-29	YP_286388.1	Betaproteobacteria	
C_21.1	126	M - COG0463	Manosil transferase dolicol-fosfato	<i>Thiobacillus denitrificans</i> ATCC 25259	8,E-38	YP_315294.1	Betaproteobacteria	
C_21.2	99	-	Sintase de lipídio A dissacarideo (Lipid-A-disaccharide synthase)	<i>Methylobacter tundripaludum</i> SV96	6,E-29	ZP_07653004.1	Gammaproteobacteria	
C_21.3	649	T - COG2200	Diguanilato ciclase / fosfodiesterase (domínios GGDEF e EAL), com sensor PAS/PAC	<i>Methylotenera mobilis</i> JLW8	0.0	YP_003047568.1	Betaproteobacteria	
C_21.4	265	F - COG0207	Timidilato sintase	<i>Aromatoleum aromaticum</i> EbN1	3,E-125	YP_159532.1	Betaproteobacteria	
C_21.5	165	H - COG0262	Diidrofolato redutase	<i>Thiomonas intermedia</i> K12	8,E-43	YP_003643103.1	Betaproteobacteria	
C_21.6	699	T - COG2200	Diguanilato ciclase / fosfodiesterase (domínios GGDEF e EAL), com sensor PAS/PAC	<i>Dechloromonas aromatica</i> RCB	2,E-149	YP_284581.1	Betaproteobacteria	

6.7.1. Organização genética do fosmideo 1A

Apesar dos diversos esforços realizados para conseguir um único contig para o fosmideo 1A que refletisse a organização genética do inserto metagenômico inteiro, ainda permaneceram regiões que provavelmente não foram suficientemente sequenciadas para gerar uma superposição de sequências (*reads*) e a montagem em um único fragmento. Não obstante, as predições de ORFs nos contigs montados foram realizadas com o fim de conhecer a estrutura genômica deste fragmento que demonstrou uma habilidade destacada para degradar o hidrocarboneto alifático modelo, o hexadecano. As características metabólicas de cada proteína predita foram analisadas segundo a similaridade da sequência e, quando possível, através de comparações de agrupamentos similares de genes em outros organismos.

Similarmente ao que foi encontrado para o fosmideo 2B, os genes que codificam as enzimas das vias de degradação de hidrocarbonetos não foram encontrados formando clusters no fosmideo 1A. Neste último caso, isto poderia ser explicado por duas razões principais. Primeiro, pela falta de conhecimento do fragmento original inteiro, o que leva a pensar que provavelmente os genes que codificam as outras etapas nas vias de degradação permaneceram nas regiões designadas como GAPS entre os contigs. A segunda (que também pode ser aplicada para o fosmideo 2B), tem a ver com a baixa similaridade entre as proteínas preditas quando são comparadas com as enzimas amplamente descritas na literatura que pertencem a organismos tipo ou isolados, extensivamente estudados.

Em particular, os genes que codificam para as enzimas álcool desidrogenase e aldeído desidrogenase não se apresentaram em agrupamentos similares aos clusters de genes aeróbios descritos primeiramente para *P. putida* GPo1 e posteriormente para a maioria dos organismos com habilidade de degradar aerobicamente os alcanos. A ausência de um sistema de alcano hidroxilase nos dois insertos metagenômicos, os quais se esperava que estivessem presentes próximos às regiões que codificam as enzimas álcool e aldeído desidrogenases, foi uma das características mais importantes compartilhadas entre os dois fosmideos. No entanto, a presença de genes codificando álcool desidrogenase e aldeído desidrogenase nos insertos fosmidiais também poderia estar acompanhada de outro sistema de oxidação de alcanos ou de outros compostos, que ainda não foram descritos, ou poderia ainda ser explicada pela já reportada redundância das enzimas álcool e aldeído desidrogenases nos genomas (Rojo, 2010).

6.7.2. ORFs relacionadas às funções metabólicas de degradação aeróbia

A ORF **C_14**, anotada como proteína tipo Pirin YhaK, pertence à superfamília Cupin, entre cujos membros se encontram várias enzimas de clivagem do anel extradiol (extradiol-ring-cleavage

- EDO) tipo III. Por definição, EDOs são enzimas extradiol dioxigenases, chaves nas etapas de degradação subsequentes à oxidação inicial de compostos aromáticos para formar ácidos carboxílicos, futuramente degradados a intermediários de vias catabólicas centrais (Vilchez-Vargas et al., 2010). Diversos esforços têm sido feitos para ampliar o conhecimento da genética deste grupo de enzimas do metabolismo aromático usando métodos independentes de cultivo. Em um estudo reportado por Junca e Pieper (2004) foi avaliada a diversidade de genes EDO em solos contaminados. No entanto, é bem sabido que as aproximações baseadas em similaridade de sequência ainda sofrem de viés, já que se baseiam na estreita janela de genes idênticos ou similares àqueles observados em bactérias conhecidas para o desenho de *primers*. Dada a diversidade metabólica, dificilmente um par de primers permitirá cobrir a diversidade catabólica de uma família de genes (Vilchez-Vargas et al., 2010). Essas dificuldades foram superadas com outras aproximações baseadas na expressão da atividade EDO em Suenaga e colaboradores (2009). O sequenciamento completo de vários insertos fosmidiais revelaram que mais da metade das sequências de EDOs encontradas pertenciam a novas subfamílias e que só alguns subconjuntos de sequências exibiram as vias de degradação descritas na literatura, ao passo que a maioria dos genes apresentaram novos arranjos (Suenaga et al., 2009).

A ORF **C_14**, que codifica para uma das famílias das enzimas EDO tipo III, se encontra em um contig com uma única ORF predita de 153 aminoácidos (Figura 35) e apresenta a maior similaridade (84%) com uma proteína de *E. coli*. A ausência de informação da região flanqueadora desta proteína não permitiu estabelecer comparações com os clusters de ORFs preditas para a mesma família de proteínas descritas em Suenaga et al. (2009) e nem por outros autores. No entanto, a maioria das ORFs preditas para o fosmideo 1A, em geral, também foram encontradas nos insertos metagenômicos de Suenaga et al. (2009), como por exemplo, álcool desidrogenases, semialdeído desidrogenases, transportadores tipo ABC, proteínas do sistema sensor de dois componentes, histidina quinases de transdução sensorial, entre outras. Contudo, a similaridade da ORF **C_14** foi maior com uma proteína da bactéria *E. coli* que com outras bactérias biodegradadoras. Esta baixa similaridade indica que essa proteína poderia constituir uma nova enzima da superfamília cupin ainda não descrita.

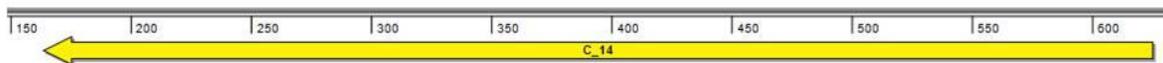


Figura 35. Esquema do contig contendo a única ORF C_14, que codifica a proteína da superfamília Cupin, presente no inserto metagenômico do fosmideo 1A.

Similarmente ao fosmideo 2B, no fosmideo 1A também foram encontrados genes codificando para as enzimas álcool desidrogenase (ORF **C_1.1**) e aldeído desidrogenase (ORF **C_16**). O papel destas enzimas no metabolismo aeróbio de alcanos já foi explicado previamente (Item 6.11.2). A ADH codificada pela ORF **C_1.1** apresenta 351 aminoácidos, o que a faz corresponder às ADHs de cadeia longa que usam zinco como cofactor. Esta proteína apresentou 75% de similaridade com a proteína do domínio álcool desidrogenase de ligação ao zinco de *Thauera sp.* MZ1T. A etapa seguinte da via de oxidação de alcanos é catalisada por enzimas aldeído desidrogenase (ADLHs). Uma sequência codificadora para este tipo de enzima foi observada na ORF **C_16** do fosmideo 1A, a qual mostrou a maior similaridade (70%) da sua sequência de aminoácidos com uma succinato aldeído desidrogenase dependente de NAD, de *Shewanella putrefaciens* 200.

No genoma completo de *A. aromatoleum* (Rabus et al., 2005), um micro-organismo que apresenta a maioria das vias de degradação de hidrocarbonetos, foram identificados 12 genes diferentes para álcool desidrogenases (11 genes pertencendo à subfamília de cadeia longa contendo zinco e um gene da subfamília de álcool desidrogenase contendo ferro), assim como 17 genes para aldeído desidrogenases.

A aldeído desidrogenase (ORF **C_16**) do fosmideo 1A encontra-se presente em um contig com uma única ORF e que, portanto, não permitiu estabelecer as regiões flangeadoras dentro do inserto metagenômico util numa análise de genômica comparativa. A Figura 36 exhibe a organização genética do contig 1. O gene da enzima álcool desidrogenase - ORF **C_1.1** (em amarelo) se encontra flanqueado por uma região codificante para o regulador transcripcional da família Fis (em laranja) e a proteína de resistência a hidroperóxidos orgânicos - ORF **C_1.2** (em verde).



Figura 36. Esquema do contig 1 contendo as ORFs **C_1.1** e **C_1.2** que codificam para uma álcool desidrogenase e uma proteína de resistência a hidroperóxido orgânicos, respectivamente. Em laranja, região parcial codificante do regulador transcripcional da família Fis; em amarelo, o gene da álcool desidrogenase; em verde, a proteína de resistência a hidroperóxido orgânicos.

6.7.3. ORFs relacionadas a outros processos celulares e metabólicos

O contig de maior tamanho do fosmideo 1A apresentou várias ORFs, com diferentes funções (Figura 37). A ORF **C_21.1** codifica para uma proteína glicosiltransferase envolvida na

biogênese da parede celular, a ORF seguinte à jusante codifica uma proteína sem qualquer categoria COG relacionada, mas que apresentou 67% de similaridade com uma proteína da síntese de lipídio A (o componente que envolve a membrana de bactérias Gram negativas) de *Methylobacter tundripaludum* SV96. As ORFs **C_21.3** e **C_21.6** à jusante apresentaram significativa similaridade com a proteína implicada na sinalização diguanilato ciclase (GGDEF) (52 e 61% de identidade, respectivamente). Um alto nível de capacidade de sinalização por diguanilato ciclase está implícito no genoma de *Dechloromonas aromática* pela presença de 57 proteínas codificantes de domínio GGDEF, em comparação a 19 codificadas por *E. coli*, e mais 10 domínios adicionais de reguladores da resposta de GGDEF. Essa família de genes de GGDEF parece estar sob expansão recente na história evolutiva microbiana (Salineiro et al., 2009). Os domínios GGDEF/EAL presentes nestas sequências são reguladores de resposta implicados na colonização de raízes em *P. putida*, e em *E. coli* o domínio GGDEF regula a transcrição de enzimas da parede celular (Salineiro et al., 2009).



Figura 37. Esquema do contig de maior tamanho do fosmideo 1A. As diferentes cores indicam as categorias funcionais dos genes que codificam proteínas putativas. Setas azuis indicam ORFs relacionadas com metabolismo; setas verdes, ORFs relacionadas com processos celulares; e setas brancas, ORFs sem categoria COG relacionada.

Outra ORF putativa no meio deste contig, sem filiação a COG, apresentou algum grau de similaridade com a flavoproteína em *Azoarcus* sp. BH72 (56%) e, em geral, com o domínio beta-lactamase. As ORFs subseguintes à jusante, **C_21.4** e **C_21.5**, codificam para duas enzimas relacionadas com a biossíntese de pirimidinas e folato, respectivamente. Por último, a ORF **C_21.7** apresentou 66% de identidade com a proteína que codifica um ABC transportador em *Dechloromonas aromática*.

Similarmente ao fosmideo 2B, no fosmideo 1A também não foram encontradas sequências do marcador filogenético RNA ribossomal 16S. Foram consideradas as mesmas análises utilizadas para investigar a origem microbiana do inserto fosmidial 2B. Primeiramente, utilizando o número de ORFs atribuídas para cada filo e classe taxonômica e, posteriormente, com a reconstrução filogenética de algum gene selecionado. No fosmideo 1A deu-se preferência para se utilizar um gene relacionado com as funções de processamento da informação, os denominados genes “housekeeping”, definidos como a categoria J do COG (Salineiro et al., 2009). O único gene representante da categoria J do COG no fosmideo 1A foi a ORF **C_9.1**, que codifica para uma metiltransferase da subunidade pequena do RNA ribossomal (Ribosomal RNA small subunit

methyltransferase B). Todas as ORFs preditas para o fosmideo 1A apresentaram homólogos no filo Proteobacteria. Mais de 63% das ORFs preditas foram afiliadas à classe Betaproteobacteria, e, em segundo lugar, Gamaproteobacteria (22%). Para conhecer a afiliação mais próxima do fosmideo 1A dentro das classes do filo Proteobacteria, foi construída uma árvore filogenética com varias sequências da proteína metiltransferase da subunidade pequena do RNA ribossomal do fosmideo 1A e diferentes representantes da mesma proteína nas classes dentro do filo Proteobacteria. A Figura 38 exibe a árvore gerada usando homólogos alinhados com o método de Neighbor-joining. A sequência do gene mostrou ser relacionada com as classes do filo Proteobacteria, e sugere-se uma relação evolutiva mais próxima do fosmideo 1A com as sequências da classe Betaproteobacteria, reforçando a hipótese de que a origem microbiana do inserto metagenômico poderia pertencer a algum organismo dentro da classe Betaproteobacteria.

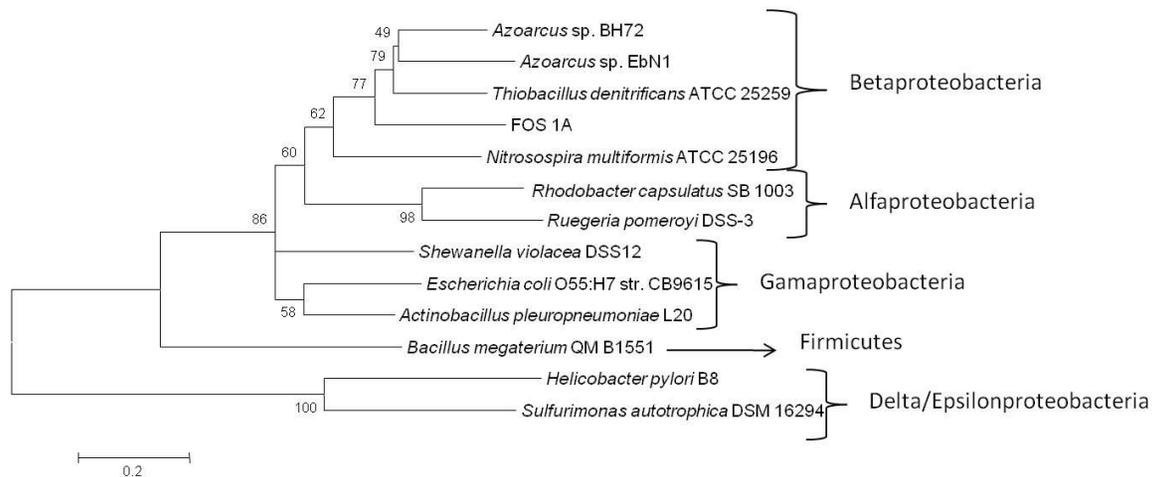


Figura 38. Análise filogenética baseada na sequência completa do gene da subunidade pequena de RNA ribossomal metiltransferase presente no inserto metagenômico do fosmideo 1A.

6.8. Considerações sobre a afiliação filogenética dos fosmídeos

A afiliação filogenética dos fragmentos metagenômicos constitui uma etapa importante na análise do metagenoma, pois representa a oportunidade de ligar a filogenia de micro-organismos não caracterizados ou não cultivados às atividades funcionais codificadas pelos fragmentos de DNA capturados na biblioteca metagenômica. Os dois fragmentos metagenômicos mostraram maior afiliação com a classe Betaproteobacteria do filo Proteobacteria. Os membros desta abundante classe são fenotipicamente, metabolicamente e ecologicamente bastante diversificados. Muitos isolados degradam aerobicamente hidrocarbonetos alifáticos e aromáticos,

assim como compostos cloroaromáticos, nitroaromáticos e aminoaromáticos (Parales, 2010). Os gêneros mais comuns incluem: *Acidovorax* (Kimbara et al., 1989; Haigler et al., 1994; Fahy et al., 2006), *Burkholderia* (Haigler et al., 1992; Laurie e Lloyd-Jones, 1999; Yuste et al., 2000), *Polaromonas* (Coleman et al., 2002), entre outros. A classe Betaproteobacteria também tem demonstrado habilidade para biodegradar aromáticos anaerobicamente, principalmente sob condições nitrato redutoras (Parales, 2010). Tal degradação é principalmente realizada por espécies de *Azoarcus* (Krieger et al., 1999) e *Thauera* (Shinoda et al., 2004).

Embora as betaproteobactérias não cresçam tão rápido quanto alguns isolados de *Pseudomonas* (classe Gamaproteobacteria) sobre hidrocarbonetos, vem sendo considerado que essas bactérias poderiam ser mais abundantes ou mais eficientes degradadoras em ambientes naturais e, portanto, suas contribuições específicas no ambiente precisam ser determinadas (Parales, 2010). Parales (2010) acredita que o sequenciamento dos genomas de outras cepas da classe Betaproteobacteria revelaria a presença de genes e vias de degradação de hidrocarbonetos que ainda não foram detectados até hoje.

De acordo com a literatura, a classe Betaproteobacteria já foi encontrada em ambientes petrolíferos. Em um estudo independente de cultivo desenvolvido por outro grupo de pesquisa brasileiro (Von der Weid et al., 2008), utilizando amostras de rocha reservatório de subsuperfície profunda proveniente da Bacia dos Campos, foi observada a predominância de clones identificados como pertencentes à classe Betaproteobacteria (53%), seguida da classe Alfaproteobacteria (25%). No nosso grupo de pesquisa, Vasconcellos et al. (2010) e Silva (2010) empregaram métodos dependentes e independentes de cultivo na tentativa de caracterizar a diversidade procariótica presente no reservatório de petróleo GMR75 da Bacia Portiguar, o mesmo poço que deu origem à amostra utilizada na construção da biblioteca metagenômica objeto deste estudo. A partir destas pesquisas pode ser concluído que a classe Betaproteobacteria constitui uma fração representativa da diversidade total da microbiota presente neste reservatório, com espécies relacionadas a *Petrobacter*, *Thauera*, *Dechlorosoma* e *Hydrogenophilus*.

De acordo com os resultados encontrados no presente estudo, sequências responsáveis por elevadas taxas de biodegradação de hidrocarbonetos alifáticos e aromáticos estão relacionadas com micro-organismos da classe Betaproteobacteria, o que é um indicativo da presença e provável metabolismo ativo destas bactérias nos reservatórios de petróleo.

C_21.7	189	E - COG0683	Transportador tipo ABC de ligação ao substrato.	<i>Dechloromonas aromatica</i> RCB	2,E-70	YP_284659.1	Betaproteobacteria	
C_4.1	43	CR- COG0604	Álcool desidrogenase	<i>Novosphingobium aromaticivorans</i> DSM 12444	9,E-10	YP_496709.1	Alfaproteobacteria	
C_4.2	98	L - COG2801	Regulador transcriptional da familia Fis	<i>Nitrococcus mobilis</i> Nb-231	6,E-25	ZP_01127243.1	Gammaaproteobacteria	
C_5.1	347	L - COG1198	Helicase PriA esencial para a replicação do DNA, independente de dnaA	<i>Thauera</i> sp. MZ1T	2,E-112	YP_002891033.1	Betaproteobacteria	
C_5.2	51	-	Proteína hipotética					
C_5.3	400	S - COG2855	Proteina putativa de membrana YeiH	<i>Nitratiruptor</i> sp. SB155-2	1,E-102	YP_001355639.1	Delta/Epsilon proteobacteria	
C_9.1	179	J - COG0144	Metiltransferase da subunidade pequena do RNA ribossomal.	<i>Aromatoleum aromaticum</i> EbN1	8,E-53	YP_158615.1	Betaproteobacteria	
C_9.2	174	No related COG	Peptido sinal rico em prolina (proline rich signal peptide)	<i>Comamonas testosteroni</i> CNB-2	3,E-11	YP_003276403.1	Betaproteobacteria	
C_9.3	713	T - COG0642	Histidina quinase de transdução sensorial (sensory transduction histidine kinase)	<i>Aromatoleum aromaticum</i> EbN1	1,E-171	YP_158613.1	Betaproteobacteria	

6.9. Considerações sobre as relações evolutivas dos genes de degradação de hidrocarbonetos

Numerosos genes relevantes para a degradação de hidrocarbonetos foram identificados no fosmideo 2B e alguns no fosmideo 1A. A maioria deles mostrou-se relacionada com o metabolismo de hidrocarbonetos aromáticos. Esses genes estão envolvidos em processos muito dinâmicos de intercâmbio de material genético (Vilchez-Vargas et al., 2010). Isso ocorre em função da pressão seletiva gerada pela presença de compostos externos, levando à seleção de bactérias mutantes que são capazes de utilizá-los. Portanto, os genes catabólicos que são freqüentemente abrigados em elementos genéticos móveis (como plasmídeos, transposons e elementos integrativos e conjugativos) disseminam efetivamente as características a bactérias filogeneticamente diversas (Salineiro, et al., 2009).

No presente trabalho, foram encontradas várias proteínas putativas codificando para as mesmas enzimas ou famílias de proteínas, como por exemplo, elementos regulatórios e transportadores tipo ABC. Por outro lado, as únicas proteínas relacionadas com a degradação presentes nos dois insertos fosmidiais foram as álcool e aldeído desidrogenases. Recentemente, diversas aldeído-desidrogenases envolvidas na degradação de hidrocarbonetos foram caracterizadas na bactéria *Geobacillus thermodenitrificans*, isolada de reservatório de petróleo na China e que mostrou atividade degradadora de alcanos (Li et al., 2010).

As sequências de aminoácidos das aldeído desidrogenases codificadas pelos dois fosmideos foram comparadas com sequências homólogas depositadas nas bases de dados. As sequências foram alinhadas utilizando o programa CLUSTAL W e as árvores filogenéticas foram reconstruídas com o programa MEGA usando o algoritmo Neighbor-joining com correção p-distancia (Figura 39).

A árvore filogenética gerada foi congruente com a classificação de ALDHs descrita por Perozich e colaboradores (1999) e sugere que as duas ALDHs encontradas pertencem a duas famílias diferentes. As famílias de ALDHs representam grupos de sequências ortológicas, as quais são claramente parálogas às sequências de outras famílias de ALDHs (Perozich et al., 1999). O gene da ALDH do fosmideo "FOS2B" pertence ao grupo de ALDHs não específicas (14th family), as quais utilizam uma ampla gama de aldeídos aromáticos e alifáticos; e o gene da ALDH do fosmideo "FOS1A" pertence ao grupo de ALDHs substrato-específicas (SSALDHs).

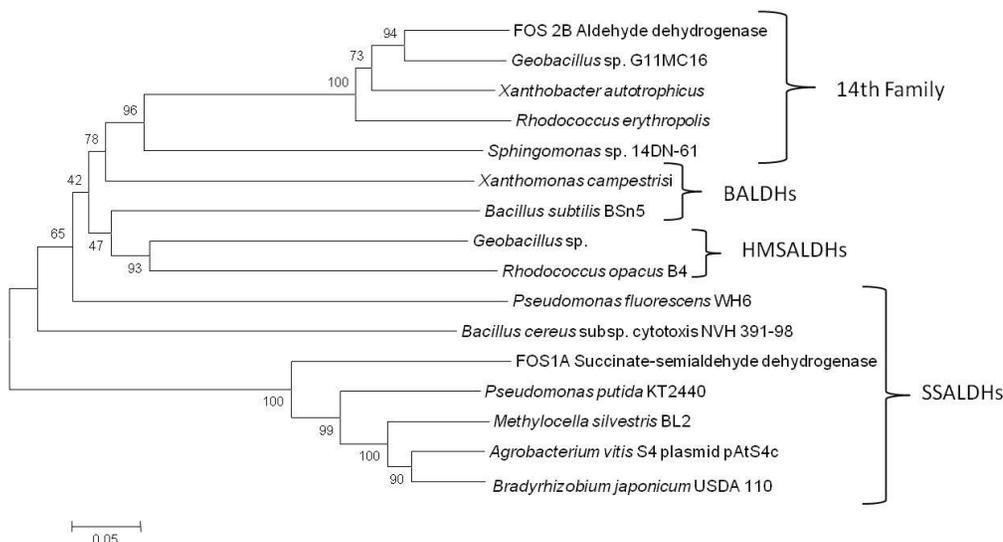


Figura 39. Árvore filogenética das seqüências de aldeído desidrogenase preditas nos insertos metagenômicos e seus homólogos bacterianos. 14th ALDH family, família de ADLHs não específicas; BALDHs, família de betaina desidrogenase; HMSALDs, família 2-hidroximuconico semialdeido desidrogenase; SSALDHs, família succinato semialdeido desidrogenases.

O tipo de atividade esperada para as famílias às quais pertencem as seqüências de ALDHs dos dois fósmeos é coerente com os resultados obtidos nos ensaios de biodegradação. O fósmeo 2B demonstrou habilidade para degradar ambos os compostos alifáticos e aromáticos, sendo que, neste caso, a ALDH presente apresenta uma ampla gama de substratos incluindo alifáticos e aromáticos. Por outro lado, o fósmeo 1A apresentou habilidade para degradar um composto alifático, mas não o aromático e, neste caso, a ALDH encontrada pertence à família semialdeído succínico desidrogenase - SSALDH substrato-específica (EC 1.2.1.16). SSALDHs catalizam a oxidação de semialdeído succínico em bactérias (composto alifático) (Figura 40).

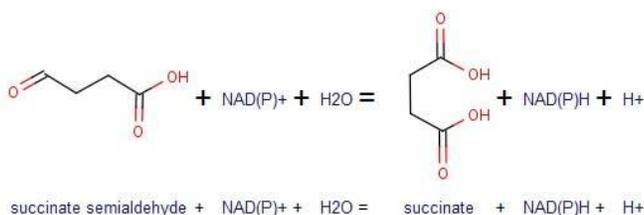


Figura 40. Reação catalisada pela enzima semialdeído succínico desidrogenase. Fonte: BRENDA (Barthelmes et al., 2007).

A presença dessas ALDHs com diferentes especificidades de substratos no metagenoma sugere que as bactérias presentes no reservatório desempenhem um papel na degradação não só de alcanos, mas também de hidrocarbonetos aromáticos presentes no petróleo.

7. CONCLUSÕES

Os reservatórios de petróleo representam um ecossistema pouco compreendido, onde os micro-organismos e mecanismos envolvidos nos processos de biodegradação natural do óleo ainda precisam ser elucidados. Este estudo utilizou uma abordagem funcional para a prospecção e caracterização de genes com elevadas capacidades de degradação de hidrocarbonetos originados no metagenoma microbiano de reservatório de petróleo.

A partir dos resultados obtidos pode ser concluído que:

- A avaliação da capacidade de degradação dos clones por cromatografia foi eficiente na seleção daqueles com maiores habilidades catalíticas, permitindo o estudo subsequente das vias de degradação com detalhe.
- Foi possível a otimização das técnicas de construção de sub-bibliotecas de plasmídeos a partir de fosmídeos para sequenciamento completo pela estratégia “*shotgun*”.
- A análise e montagem das sequências derivadas da biblioteca *shotgun* permitiu a reconstituição dos insertos metagenômicos e o acesso às informações genéticas nos fosmídeos selecionados.
- A análise das sequências contidas nos fosmídeos revelou a presença de genes que codificam várias etapas de diferentes vias de degradação, que vão desde a degradação de compostos alifáticos a aromáticos, sob condições aeróbias e anaeróbias. A detecção destes genes catabólicos amplia o nosso conhecimento sobre a diversidade metabólica do reservatório de petróleo.
- A ausência nos fosmídeos de agrupamentos de genes que codificam as vias de degradação de hidrocarbonetos amplamente descritas na literatura pode ser explicada pela prevalência de estudos e informações sobre os processos de degradação em bactérias isoladas. Isto sugere que existem diferenças marcantes entre os genes de degradação encontrados na comunidade microbiana presente nos reservatórios de petróleo e aqueles que foram previamente identificados em bactérias isoladas de ambientes contaminados ou não.
- A afiliação filogenética das sequências metagenômicas com membros da classe Betaproteobactera indica que este grupo estaria desempenhando um papel importante nos processos de biodegradação dos componentes do óleo nos reservatórios.
- Esses resultados reforçam o potencial da abordagem metagenômica para a prospecção e elucidação de novos genes e vias metabólicas que, no caso deste estudo, constituem novos arranjos gênicos não descritos para a degradação de hidrocarbonetos até hoje.

- Atualmente, e pela primeira vez no Brasil estão sendo identificadas sequências gênicas responsáveis pelos processos de degradação do óleo nos reservatórios a partir de dados de metagenoma microbiano.
- Uma melhor caracterização destes genes quanto à especificidade de substratos e a verdadeira contribuição de cada um no metabolismo de compostos hidrocarbonetos só poderá ser identificada com testes funcionais das proteínas expressas. Adicionalmente, a descoberta destes novos genes ou vias de degradação abrem perspectivas para a sua aplicação futura processos biotecnológicos.

8. REFERÊNCIAS

- Abu Laban, N., Selesi, D., Rattei, T., Tischler, P., Meckenstock, R. U. (2010) Identification of enzymes involved in anaerobic benzene degradation by a strictly anaerobic iron-reducing enrichment culture. *Environmental Microbiology* 12: 2783-2796.
- Aitken, C. M., Jones, D. M., Larter, S. R. (2004) Anaerobic hydrocarbon biodegradation in deep subsurface oil reservoirs. *Nature* 431: 291-294.
- Aldas, R. E., Jenkins, B. M., VanderGheynst, J. S. (2007) Degradation potential and soil carbon mineralization of biomass gasification tars. (ASABE). Annual International Meeting of the American Society of Agricultural and Biological Engineers, Minnesota, pp. 1–16.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25: 3389–3402.
- Amann, R. I., Ludwig, W. and Schleifer, K. H. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews*. 59:143-169.
- Amouric, A., Quéméneur, M., Grossi, V., Liebgott, P. P., Auria, R., Casalot, L. (2009) Identification of different alkane hydroxylase systems in *Rhodococcus ruber* strain SP2B, an hexane-degrading actinomycete. *Journal of applied microbiology* 1-14.
- Aziz, R. K., Bartels, D., Best, A. A., Dejongh, M., Disz, T., et al. (2008) The RAST Server: Rapid Annotation using Subsystems Technology. *BMC Genomics* 9: 75
- Badger, J. H., e Olsen, G. J. (1999) CRITICA: coding region identification tool invoking comparative analysis. *Molecular Biology and Evolution* 16: 512-524.
- Barragan, M. J. L., Carmona, M., Zamarro, M. T., Thiele, B., Boll, M., et al. (2004) The bzd gene cluster, coding for anaerobic benzoate catabolism, in *Azoarcus* sp strain CIB. *Journal of Bacteriology* 186: 5762–5774.
- Barthelmes, J., Ebeling, C., Chang, A., Schomburg, I., Schomburg, D. (2007) BRENDA, AMENDA and FRENDA: the enzyme information system in 2007. *Nucleic Acids Research* 35: D511–D514.
- Bastin, E. S. (1926) The problem of the natural reduction of sulphates. *Bulletin of American Association of Petroleum Geologist* 10: 1270–1299.
- Beloqui, A., de María, P. D., Golyshin, P. N., Ferrer, M. (2008) Recent trends in industrial microbiology. *Current Opinion in Microbiology* 11: 240–248.
- Beloqui, A., Pita, M., Polaina, J., Martinez-Arias, A., Golyshina, O. V., et al. (2006) Novel polyphenol oxidase mined from a metagenome expression library of bovine rumen: biochemical properties, structural analysis and phylogenetic relationships. *Journal of Chemical Biology* 281: 22933–22942.
- Bergmann, F. D., Selesi, D., Weinmaier, T., Tischler, P., Rattei, T., Meckenstock, R. U. (2010) Genomic insights into the metabolic potential of the polycyclic aromatic hydrocarbon degrading sulphate-reducing Deltaproteobacterium N47. *Environmental Microbiology* 13: 1125-1137
- Bernard, F. P., Connan, J., Magot, M. (1992) SPE 24811 Indigenous microorganisms in connate water of many oil fields: a new tool in exploration and production techniques. In Proceedings of the 67th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers, Washington, DC, pp. 467–476.
- Besemer, J., e Borodovsky, M. (1999) Heuristic approach to deriving models for gene finding. *Nucleic Acids Research* 27: 3911-3920.
- Birkeland, N. K. (2004) The microbial diversity of deep subsurface oil reservoirs. *Studies in Surface Science and Catalyses* 151: 385–403.
- Boll, M., e Fuchs, G. (1995) Benzoyl-CoA reductase (dearomatizing), a key enzyme of anaerobic aromatic metabolism. ATP-dependence of the reaction, purification and some properties of the enzyme from *Thauera aromatica* strain K 172. *European Journal of Biochemistry* 234: 921–933.

- Boll, M., Fuchs, G., Heider, J. (2002) Anaerobic oxidation of aromatic compounds and hydrocarbons. *Current Opinion in Chemical Biology* 6: 604– 611
- Boll, M., e Heider, J. (2010) Anaerobic degradation of hydrocarbons: mechanisms of C-H-Bond activation in the absence of oxygen. In: Handbook of hydrocarbon and lipid Microbiology. Timmis K. N. (ed.). Berlin Heidelberg. Springer-Verlag, 1012-1024
- Boll, M., Laempe, D., Eisenreich, W., Bacher, A., Mittelberger, T., Heinze, J. and Fuchs, G. (2000) Nonaromatic products from anoxic conversion of benzoyl-CoA with benzoyl-CoA reductase and cyclohexa-1,5-diene-1-carbonyl-CoA hydratase. *Journal of Biological Chemistry*, 275, 21889-21895.
- Bown, D. P., Gatehouse, J. A. (2004) Characterization of a digestive carboxypeptidase from the insect pest corn earworm (*Helicoverpa armigera*) with novel specificity towards C-terminal glutamate residues. *European Journal of Biochemistry*. 271: 2000–2011.
- Breinig, S., Schlitz, E., Fuchs, G. (2000) Genes involved in anaerobic metabolism of phenol in the bacterium *Thauera aromatic.*, *Journal of Bacteriology*. 182: 5849–5863.
- Brennerova, M. V., Josefiova, J., Brenner, V., Pieper, D. H., and Junca, H. (2009) Metagenomics reveals diversity and abundance of meta-cleavage pathways in microbial communities from soil highly contaminated with jet fuel under air-sparging bioremediation. *Environmental Microbiology* 11: 2216– 2227.
- Bressler, D. C. e Fedorak, P. M. (2001) Identification of disulfides from the biodegradation of dibenzothiophene. *Applied and Environmental Microbiology*. 67: 5084–5093.
- Carmona, M., Zamarro, M. T., Blazquez, B., Durante-Rodríguez, G., Juárez, J. F., Valderrama, J. A., Barragán, M. J. L. García, J. L., e Díaz, E. (2009) Anaerobic Catabolism of Aromatic Compounds: a Genetic and Genomic View. *Microbiology and Molecular Biology Reviews* 73: 71–133.
- Chakraborty, R., e Coates, J. (2004) Anaerobic degradation of monoaromatic hydrocarbons. *Applied Microbiology and Biotechnology* 71: 5427–5432.
- Chandler, D. P., Shu-Mei, L., Spadoni, C. M., Drake, G. R., Balkwill, D. L., Fredrickson, J. K., Brockman, F.J. (1997) A molecular comparison of culturable aerobic heterotrophic bacteria and 16S rDNA clones derived from a deep subsurface sediment. *FEMS Microbiology Ecology* 23: 131–144
- Chen K, e Pachter L. (2005) Bioinformatics for whole-genome shotgun sequencing of microbial communities. *PLoS Computational Biology* 1:106-112
- Chou, H. H., e Holmes, M. H. (2001) DNA sequence quality trimming and vector removal. *Bioinformatics*. 17: 1093–1104.
- Coates, J. D., Woodward, J., Allen, J., Philip, P., Lovley, D. R. (1997) Anaerobic degradation of polycyclic aromatic hydrocarbons and alkanes in petroleum-contaminated marine harbor sediments. *Applied Environmental Microbiology* 63: 3589-3593.
- Coleman, N. V., Bui, N. B., Holmes, A. J. (2006) Soluble di-iron monooxygenase gene diversity in soils, sediments and ethene enrichments. *Environmental Microbiology* 8:1228–1239
- Coleman, N. V., Mattes, T. E., Gossett, J. M., Spain, J. C. (2002) Biodegradation of cis-dichloroethene as the sole carbon source by a beta-proteobacterium. *Applied Environmental Microbiology* 68: 2726–2730.
- Cunha, C. D., Rosado, A. S. Sebastián G. V. Seldin, L. and Von der Weid, I. (2006). Oil biodegradation by *Bacillus* strains isolated from the rock of an oil reservoir located in a deep-water production basin in Brazil. *Applied Microbiology and Biotechnology* 73:949–959
- Daniel, R. (2005) The metagenomics of soil. *Nature Reviews Microbiology* 3: 470-478.
- Dassa, E., e Bouige, P. (2001) The ABC of ABCS: a phylogenetic and functional classification of ABC systems in living organisms. *Research in Microbiology*. 152 211-29
- Delcher, A. L., Harmon, D., Kasif, S., White, O., Salzberg, S. L. (1999) Improved microbial gene identification with GLIMMER. *Nucleic Acids Research* 27:4636-4641.

- Diaz-Torres, M. L., McNab, R., Spratt, D. A., Villedieu, A., Hunt, N., Wilson, M., Mullany, P. (2003) Novel tetracycline resistance determinant from the oral metagenome. *Antimicrobial Agents and Chemotherapy* 47: 1430–1432.
- DiDonato, R. J., Young, N. D., Butler, J. E., Chin, K., Hixson, K. K., Mouser, P., Lipton, M. S. DeBoy, R., Methé, B. (2010) A Genome sequence of the deltaproteobacterial strain NaphS2 and analysis of differential gene expression during anaerobic growth on naphthalene. *PLoS one*. 5: 14072
- Dunaway-Mariano, D., Babbitt, P. C. (1994) On the origins and functions of the enzymes of the 4-chlorobenzoate to 4-hydroxybenzoate converting pathway. *Biodegradation* 5: 259–276
- Egland, P. G., Pelletier, D. A., Dispensa, M., Gibson, J., Harwood, C. S. (1997) A cluster of bacterial genes for anaerobic benzene ring biodegradation. *The National Academic of Sciences* 94, 6484–6489.
- Erwin, D. P., Erickson, I. K., Delwiche, M. E., Colwell, F. S., Strap, J. L., Crawford, R. L. (2005) Diversity of oxygenase genes from methane- and ammonia- oxidizing bacteria in the Eastern Snake River Plain aquifer. *Applied Environmental Microbiology*. 71: 2016–2025
- Ewing, B., Green, P. (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Research* 8: 186–194
- Ewing, B., Hiller, L., Wendl, M.C., Green P. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Research* 8: 175–185
- Eyers, L., George, I., Schuler, L., Stenuit, B., Agathos, S. N. and Said El Fantroussi. (2004) Environmental genomics: exploring the unmined richness of microbes to degrade xenobiotics. *Applied Microbiology and Biotechnology* 66: 123–130.
- Fahy, A., McGenity, T.J., Timmis, K. N., Ball, A. S. (2006) Heterogeneous aerobic benzene-degrading communities in oxygen-depleted groundwaters. *FEMS Microbiology Ecology* 58: 260–270.
- Feng, L., Wang, W., Cheng, J., Ren, Y., Zhao, G., Gao, C. (2007) Genome and proteome of long-chain alkane degrading *Geobacillus thermodenitrificans* NG80-2 isolated from a deep-subsurface oil reservoir. *PNAS National Academy of Science of the USA* 104: 5602–5607.
- Fernandez-Lima, F. A., Becker, C., McKenna, A. M., Rodgers, R. P., Marshall, A. G., Russel, D. H. (2009) Petroleum crude oil characterization by IMS-MS and FTICR MS. *Analytical Chemistry* 81: 9941–9947
- Ferrer, M., Golyshina, O.V., Chernikova, T.N., Khachane, A.N., Reyes-Duarte, D., Santos, V. A., et al. (2005) Novel hydrolase diversity retrieved from a metagenome library of bovine rumen microflora. *Environmental Microbiology* 7: 1996–2010.
- Fierer, N., Breitbart, M., Nulton, J., Salamon, P., Lozupone, C., Jones, R., et al. (2007). Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of bacteria, archaea, fungi, and viruses in soil. *Applied Environmental Microbiology* 73: 7059–7066.
- Finn, R. D., Tate, J. Mistry, P. C., Coghill, S. J., Sammut, H. R., et al. (2008) The Pfam protein families database. *Nucleic Acids Research* 36:D281–D288.
- Finneran, K. T., Johnsen, C. V. e Lovley, D. R. (2003) *Rhodoferrax ferrireducens* sp. nov., a psychrotolerant, facultatively anaerobic bacterium that oxidizes acetate with the reduction of Fe(III). *International Journal of Systematic and Evolutionary Microbiology* 53: 669-673.
- Foght, J. M. (2010) Microbial communities in oil shales, biodegraded and heavy oil reservoirs, and bitumen deposits. In: Handbook of Hydrocarbon and Lipid Microbiology. Timmis K. N. (ed.). Berlin Heidelberg. Springer-Verlag, 2160-2172.
- Friedberg, I. (2006) Automated protein function prediction - The genomic challenge. *Briefings in Bioinformatics*, 7: 225-242.
- Fuchs, G. (2008). Anaerobic metabolism of aromatic compounds. *Annals of the New York Academy of Sciences* 1125: 82–99.
- Galvão, T., Mohn, W., de Lorenzo, V. (2005) Exploring the microbial biodegradation and biotransformation gene pool. *Trends in Biotechnology* 23: 497–506.

- Gill, S. R., Breitbart, M., Nulton, J., Salamon, P., Lozupone, C., Jones, R., et al. (2006) Metagenomic analysis of the human distal gut microbiome. *Science* 312: 1355–1359.
- Giovannoni, S. J., Britschgi, T. B., Moyer, C. L., Field, K. G. (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* 345:60–63.
- Gordon, D., Abajian, C., Green, P. (1998) "Consed: A Graphical Tool for Sequence Finishing". *Genome Research* 8: 195–202.
- Grassia, G. S., McLean, K. M., Glenat, P., Bauld, J., Sheehy, A. J. (1996) A systematic survey for thermophilic fermentative bacteria and archaea in high temperature petroleum reservoirs. *FEMS Microbiology Ecology* 21: 47–58.
- Green P: Phrap (www.phrap.org).
- Green, P. (1997) Against a whole genome shotgun. *Genome research*. 7: 410-417.
- Grossi, V., Cravo-Laureau, C., Guyoneaud, R., Ranchou-Peyruse, A., and Hirschler-Rea, A. (2008) Metabolism of n-alkanes and n-alkenes by anaerobic bacteria: a summary. *Organic Geochemistry* 39: 1197–1203.
- Guazzaroni, M. E., Belouqui, A., Vieites, J. M., Al-ramahi, Y., Cortés, N. L., Ghazi, A., Golyshin, P. N., Ferrer, M. (2010) Metagenomic mining of enzyme diversity. In Handbook of Hydrocarbon and Lipid Microbiology. Timmis, K.N. (ed.). Berlin, Heidelberg, Germany: Springer-Verlag, p. 2911–2927.
- Haigler, B. E., Pettigrew, C. A., Spain, J. C. (1992) Biodegradation of mixtures of substituted benzenes by *Pseudomonas* sp. strain JS150. *Applied Environmental Microbiology* 58: 2237–2244.
- Haigler, B. E., Wallace, W. H., Spain, J. C. (1994) Biodegradation of 2-nitrotoluene by *Pseudomonas* sp. strain JS42. *Applied Environmental Microbiology* 60: 3466–3469.
- Handelsman, J. (2004) Metagenomics: application of genomics to uncultured microorganisms. *Microbiology and molecular biology reviews*. 68: 669-685.
- Handelsman, J. (2007) The New Science of Metagenomics: Revealing the Secrets of Our Microbial Planet. The National Academies Press. Washington D.C.
- Hårdeman, F., e Sjöling, S. (2007) Metagenomic approach for the isolation of a novel low-temperature active lipase from uncultured bacteria of marine sediment. *FEMS Microbiology Ecology*. 59: 524-534.
- Head, I. M., Jones, D. M., Larter, S. R. (2003) Biological activity in the deep subsurface and the origin of heavy oil. *Nature* 426: 344–352.
- Head, I. M., Larter, S. R., Gray, N. D., Sherry, A., Adams, J. J., Aitken, C. M., Jones, D. M., Rowan, A. K., Huang, H. Röling, W. F. M. (2010) Hydrocarbon degradation in petroleum reservoirs In Timmis K. N. (ed.), Handbook of Hydrocarbon and Lipid Microbiology. Springer-Verlag Berlin Heidelberg. p. 4339-4347
- Head, I.M., Jones, D.M., Larter, S.R. (2006) Marine microorganisms make a meal of oil. *Nature* 4:173-182.
- Heider, J. (2001) A new family of CoA-transferases. *FEBS Letters*, 509: 345–349.
- Heider, J., e Rabus, R. (2008) Genomic insights in the anaerobic biodegradation of organic pollutants. In: Microbial biodegradation: genomics and molecular biology. Diaz, E. (ed.). Norfolk, UK Caister academic press, p. 25-54.
- Heiss-Blanquet, S., Benoit, Y., Marechaux, C., Monot, F. (2005) Assessing the role of alkane hydroxylase genotypes in environmental samples by competitive PCR. *Journal Applied Microbiology* 99:1392–1403.
- Hobohm, U., e Sander, C. (1995) A sequence property approach to searching protein database. *Journal of Molecular Biology*. 251: 390-399.
- Hoff, K. J., Lingner, T., Meinicke, P., Tech, M. (2009) Orphelia: predicting genes in metagenomic sequencing reads. *Nucleic Acids Research* 37:W101-105
- Huang, X. e Madan, A. (1999) CAP3: A DNA sequence assembly program. *Genome Research* 9: 868-877.
- Hugenholtz, P., Goebel, B. M., Pace, N. R. (1998) Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *Journal of Bacteriology* 180: 4765–4774

- Huson, D. H., Auch, A. F., Qi, J., Schuster, S. C. (2007) MEGAN analysis of metagenomic data. *Genome Research* 17:377-386.
- Jain, R., Rivera, M. C. e Lake, J. A. (1999) Horizontal gene transfer among genomes: the complexity hypothesis. *PNAS National Academy of Sciences of the USA* 96: 3801-3806.
- Jogler, C., Lin, W., Meyerdierks, A., Kube, M., Katzmann, E., Flies, C., Pan, Y., Amann, R., Reinhardt, R., Schüler, D. (2009) Towards cloning the magnetotactic metagenome: Identification of magnetosome island gene clusters in uncultivated magnetotactic bacteria from different aquatic sediments. *Applied and Environmental Microbiology*. 75: 3972–3979.
- Jouanneau, Y. (2010) Oxidative inactivation of ring-cleavage extradiol dioxygenases: mechanism and ferredoxin-mediated reactivation. In: *Handbook of hydrocarbon and lipid Microbiology*. Timmis, K.N. (ed.). Berlin, Heidelberg, Germany: Springer-Verlag, p. 1072-1079.
- Junca, H., e Pieper, D.H. (2004) Functional gene diversity analysis in BTEX contaminated soils by means of PCR-SSCP DNA fingerprinting: comparative diversity assessment against bacterial isolates and PCR-DNA clone libraries. *Environmental Microbiology* 6: 95–110.
- Kasap, M., Torol, S. e Gacar, G. (2008) Comparison of Electrocompetencies of Some Commonly Used Laboratory Strains of *Escherichia coli*. *Journal of Applied Biological Sciences* 2: 12-22
- Kellenberger, E. (2001) Exploring the unknown: the silent revolution of microbiology. *European Molecular Biology Organization Reports*. 2: 5-7.
- Kimbara, K., Hashimoto, T., Fukuda, M., Koana, T., Takagi, M., Oishi, M., Yano, K. (1989) Cloning and sequencing of two tandem genes involved in degradation of 2,3- dihydroxybiphenyl to benzoic acid in the polychlorinated biphenyl-degrading soil bacterium *Pseudomonas* sp. strain KKS102. *Journal of Bacteriology* 171: 2740–2747.
- Knietsch, A., Bowien, S., Whited, G., Gottschalk, G., Daniel R. (2003a) Identification and characterization of coenzyme B 12 -dependent glycerol dehydratase- and diol dehydratase encoding genes from metagenomic DNA libraries derived from enrichment cultures. *Applied Environmental Microbiology* 69: 3048–3060.
- Knietsch A, Waschkowitz T, Bowien S, Henne A, Daniel R. (2003b) Construction and screening of metagenomic libraries derived from enrichment cultures: generation of a gene bank for genes conferring alcohol oxidoreductase activity on *Escherichia coli*. *Applied Environmental Microbiology* 69: 1408–1416.
- Korda, A., Santas, P., Tenente, A., Santas, R. (1997) Petroleum hydrocarbon bioremediation: sampling and analytical techniques, in situ treatments and commercial microorganisms currently used *Applied Microbiology and Biotechnology*. 48: 677-686.
- Krieger, C. J., Beller, H. R., Reinhard, M., Spormann, A. M. (1999) Initial reactions in anaerobic oxidation of m-xylene by the denitrifying bacterium *Azoarcus* sp. strain T. *Journal of Bacteriology* 181: 6403–6410.
- Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* 305: 567–580.
- Kube, M., Heider, J., Amann, J., Hufnagel, P., Kühner, S., Beck, A., Reinhardt, R., Rabus, R. (2004) Genes involved in the anaerobic degradation of toluene in a denitrifying bacterium, strain EbN1. *Archives of Microbiology* 181: 182–194.
- Kube, M., Beck, A., Meyerdierks, R., Amann, R., Reinhardt, e Rabus, R. (2005) A catabolic gene cluster for anaerobic benzoate degradation in methanotrophic microbial Black Sea mats. *Systematic and Applied Microbiology* 28:287–294.
- Kunin, V., Copeland, A., Lapidus, A., Mavromatis, K., Hugenholtz, P. (2008) A Bioinformatician's Guide to Metagenomics. *Microbiology and Molecular Biology Reviews* 72: 557-578.
- Kunze, M., Zerlin, K. F., Retzlaff, A., Pohl, J. O., Schmidt, E., Janssen, D. B., et al. (2009) Degradation of chloroaromatics by *Pseudomonas putida* GJ31: assembled route for chlorobenzene degradation encoded by clusters on plasmid pKW1 and the chromosome. *Microbiology* 155: 4069–4083.

- L'Haridon, S., Reysenbach, A. L., Glenat, P., Prieur, D., Jeanthon, C. (1995) Hot subterranean biosphere in a continental oil reservoir. *Nature* 377: 223–224.
- Laurie, A. D., Lloyd-Jones, G. (1999) The phn genes of *Burkholderia* sp. strain RP007 constitute a divergent gene cluster for polycyclic aromatic hydrocarbon catabolism. *Journal of Bacteriology* 181: 531–540.
- Law, R. J., Hamlin, J. N., Sivro, A., McCorrister, S. J., Cardama, G. A., Cardona, S. T. (2008) A functional phenylacetic acid catabolic pathway is required for full pathogenicity of *Burkholderia cenocepacia* in the *Caenorhabditis elegans* host model. *Journal of Bacteriology* 190:7209-7218.
- Leahy, J. G., e Colwell, R. R. (1990) Microbial degradation of hydrocarbons in the environment. *Microbiology and Molecular Biology Reviews* 54: 305-315.
- LeCleir, G. R., Buchan, A. e Hollibaugh, J. T. (2004) Chitinase gene sequences retrieved from diverse aquatic habitats reveal environment-specific distributions. *Applied Environmental Microbiology* 70: 6977-6983.
- Li, W. (2009). Analysis and comparison of very large metagenomes with fast clustering and functional annotation. *BMC Bioinformatics*. 10: 359
- Li, H., Yang, S. Z., Mu, B. Z., Rong, Z. F., Zhang, J. (2007) Molecular phylogenetic diversity of the microbial community associated with a high-temperature petroleum reservoir at an offshore oilfield. *FEMS Microbiology Ecology* 60: 74–84.
- Li, X., Li, Y., Wei, D., Li, P., Wang, L., Feng, L. (2010) Characterization of a broad-range aldehyde dehydrogenase involved in alkane degradation in *Geobacillus thermodenitrificans* NG80-2. *Microbiological Research* 165: 706-712.
- Li, G., Wang, K, Liu, Y. H. (2008) Molecular cloning and characterization of a novel pyrethroid-hydrolyzing esterase originating from the metagenome. *Microbial Cell Factories* 7:38
- Liolios, K., Mavromatis, K., Tavernarakis, N., Kyrpides, N. C. (2008) The Genomes On Line Database (GOLD) in 2007: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Research* 36: D475–D479.
- Liu, J. F., Ma, L. J., Mu, B. Z., Liu, R. L., Ni, F. T., Zhou, J. X. (2005) The field pilot of microbial enhanced oil recovery in a high temperature petroleum reservoir. *Journal of Petroleum Science and Engineering*. 48: 265–271.
- Liu, X., Dong, Y., Zhang, J., Zhang, A., Wang, L., Feng, L. (2009) Two novel metal-independent long-chain alkyl alcohol dehydrogenases from *Geobacillus thermodenitrificans* NG80-2. *Microbiology*. 155: 2078–2085.
- Lorenz, P., Liebeton, K., Niehaus, F., Jurgen, E. (2002) Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. *Current Opinion in Biotechnology*. 13: 572–577.
- Lowe, T. M. e Eddy, S. R. (1997) tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*. 25: 955-964.
- Magot, M. (2005) Indigenous microbial communities in oil fields. In: Ollivier B, Magot M (eds) *Petroleum microbiology*. ASM Press, Washington, pp 21–34
- Magot, M., Ollivier, B., Patel, B. K. C. (2000) Microbiology of petroleum reservoirs. *Antonie Van Leeuwenhoek* 77: 103–116.
- Marchant, R., Sharkey, F.H., Banat, I. M., Rahman, T. J., Perfumo A. (2006) The degradation of n-hexadecane in soil by thermophilic geobacilli. *FEMS Microbiology Ecology* 56: 44–54.
- Margesin, R., Labbe, D., Schinner, F., Greer, C. W., Whyte, L. G. (2003). Characterization of hydrocarbon-degrading microbial populations in contaminated and pristine Alpine soils. *Applied Environmental Microbiology* 69: 3085–3092.
- Margesin, R., e Schinner, F. (2001) Biodegradation and bioremediation of hydrocarbons in extreme environments. *Applied Microbiology and Biotechnology* 56: 650–663.
- Margulies, M., Egholm, M., Altman, W. E. et al. (2005) Genome sequencing in open micro fabricated high density picoliter reactors. *Nature* 437: 376-380.

- Marin, P. e Pieper D. H. (2009) Novel metal-binding site of *Pseudomonas reinekei* MT1 trans-dienelactone hydrolase. *Biochemical and Biophysical Research Communications* 390: 1345–1348.
- Markowitz, V. M., Ivanova, N. N., Szeto, E., Palaniappan, K., Chu, K., Dalevi, D., Chen, I. M., Grechkin, Y., Dubchak, I., Anderson, I., et al. (2008) IMG/M: a data management and analysis system for metagenomes. *Nucleic Acids Research* D534-538.
- Marshall, A. G., e Rodgers, R. P. (2004) Petroleomics: The next grand challenge for chemical analysis. *Accounts of Chemical Research* 37: 53–59.
- Mchardy, A., Martin, H., Tsirigos, A., Hugenholtz, P., Igoutsos, I. (2007) Accurate phylogenetic classification of variable-length DNA fragments. *Nature Methods*. 4: 63–72.
- McLeod, M. P., e Eltis, L. D. (2008) Genomic insights into the aerobic pathways for degradation of organic pollutants. In E. Díaz (ed.), *Microbial biodegradation: genomics and molecular biology*. p. 1–23. Norfolk, United Kingdom, Caister Academic Press.
- Mehboob, F., Junca, H., Schraa, G., Stams, A. J. M. (2009) Growth of *Pseudomonas chloritidismutans* AW-1(T) on n-alkanes with chlorate as electron acceptor. *Applied Microbiology and Biotechnology* 83: 739-747.
- Mondego, J. M. C., Vidal, R. O., Carazzolle, M. F., Tokuda, E. K., Parizzi, et al. (2011) An EST-based analysis identifies new genes and reveals distinctive gene expression features of *Coffea arabica* and *Coffea canephora*. *BMC Plant Biology* 11:30
- Myers, E. W., Sutton, G. G., Delcher, A. L., Dew, I. M., Fasulo, D. P., et al. (2000) A whole-genome assembly of *Drosophila*. *Science*, 287, 2196–2204.
- Nakayama, C. R. (2005) Degradação anaeróbia de pentaclorofenol (PCP), 2,3,4-triclorofenol (2,3,4-TCP) e 2,6-diclorofenol (2,6-DCP) por culturas enriquecidas a partir do sedimento contaminado do Sistema Estuarino Santos e São Vicente, Tese de Doutorado, Instituto de Ciências biomédicas - Universidade de São Paulo, Brasil.
- Noguchi, H., Park, J., Takagi, T. (2006) MetaGene: prokaryotic gene finding from environmental shotgun sequences. *Nucleic Acids Research* 34: 5623–5630.
- Ollivier, B., e Magot, M. (2005) *Petroleum Microbiology*. Washington DC: ASM Press, p. 365.
- Orphan, V. J., Taylor, L. T., Hafenbradl, D., Delong, E. F. (2000) Culture-dependent and culture-independent characterization of microbial assemblages associated with high-temperature petroleum reservoirs. *Applied Environmental Microbiology* 66: 700–711.
- Parales, R. E. (2010) Hydrocarbon degradation by Betaproteobacteria. In: *Handbook of Hydrocarbon and Lipid Microbiology*. Timmis K. N. (ed.). Berlin Heidelberg. Springer-Verlag, p. 1716-1724.
- Perez-Pantoja, D., Donoso, R., Junca, H., Gonzalez, B., and Pieper, D. H. (2009) Phylogenomics of aerobic bacterial degradation of aromatics. In *Handbook of Hydrocarbon and Lipid Microbiology*. Timmis, K.N. (ed.). Berlin, Heidelberg, Germany. Springer-Verlag, p. 1356–1397.
- Pérez-Pantoja, D., González, B., Pieper, D. H. (2010) Aerobic degradation of aromatic hydrocarbons. In: *Handbook of hydrocarbon and lipid Microbiology*. Timmis, K.N. (ed.). Berlin, Heidelberg, Germany. Springer-Verlag, p. 799-837.
- Perozich, J., Nicholas, H., Wang, B. C., Lindahl, R., Hempel, J. (1999) Relationships within the aldehyde dehydrogenase extended family. *Protein Science* 8: 137–46.
- Pieper, D. H. (2005) Aerobic degradation of polychlorinated biphenyls. *Applied Microbiology and Biotechnology* 67: 170-191.
- Pieper, D. H., Martins dos Santos, V. A. P., Golyshin, P. N. (2004) Genomic and mechanistic insights into the biodegradation of organic pollutants. *Current Opinion in Biotechnology*. 15: 215-224.
- Pope, P. B., e Patel, B. K. C. (2008) Metagenomic analysis of a freshwater toxic cyanobacteria bloom. *Fems Microbiology Ecology*. 64: 9-27
- Prince, R. C. (2010) Bioremediation of marine oil spills. In: *Handbook of Hydrocarbon and Lipid Microbiology*. Timmis, K.N. (ed.). Berlin, Heidelberg, Germany. Springer-Verlag, p. 2617–2630

- Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R. e Lopez R. (2005) InterProScan: protein domains identifier. *Nucleic Acids Research* 33: W116-120.
- Rabus, R., Kube, M., Heider, J., Beck, A., Heitmann, K., Widdel, F., Reinhardt, R. (2005) The genome sequence of an anaerobic aromatic-degrading denitrifying bacterium, strain EbN1. *Archives of Microbiology* 183: 27–36.
- Rees, H. C., Grant, S., Jones, B., Grant W. D., Heaphy, S. (2003) Detecting cellulose an esterase enzyme activities encoded by novel genes present in environmental DNA libraries. *Extremophiles* 7: 415-421.
- Riesenfeld, C., Schloss, P., Handelsman, J. (2004) Metagenomics: Genomic analysis of microbial communities. *Reviews in Advance* 38: 525-552.
- Rodriguez-Brito, B., Rohwer, F., Edwards, R. A. (2006) An application of statistics to comparative metagenomics. *BMC Bioinformatics* 7:162.
- Rojo, F. (2009). Degradation of alkanes by bacteria. *Environmental Microbiology* 11: 2477-2490.
- Rojo, F. (2010) Genetic features and regulation of n-alkane metabolism. In: Handbook of hydrocarbon and lipid Microbiology. Timmis, K.N. (ed.). Berlin, Heidelberg, Germany. Springer-Verlag, p. 1141-1154.
- Röling, W. F. M., Head, I. M., Larter, S. R. (2003) The microbiology of hydrocarbon degradation in subsurface petroleum reservoirs: perspectives and prospects. *Research in Microbiology*. 154: 321-328.
- Rosenberg, E., e Ron, E. Z. (1996) Bioremediation of petroleum contamination. In: Bioremediation: principles and applications. R. L. Crawford and D. L. Crawford (eds.), Cambridge, England. Cambridge University Press. p. 100–124.
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., Barrell, B. G. (2000) Artemis: sequence visualization and annotation. *Bioinformatics* 16: 944–945.
- Saitou, N. e Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4: 406–425.
- Salineiro, K., Keller, K., Feil, W., Feil, H., Trong, S., Di Bartolo, G., Lapidus, A. (2009) Metabolic analysis of the soil microbe *Dechloromonas aromatica* str. RCB: indications of a surprisingly complex life-style and cryptic anaerobic pathways for aromatic degradation. *BMC Genomics* 10: 351.
- Sambrook, J., Fritsch, E. F., Maniatis, T. (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Satyanarayana, T., Raghukumar, C., Shivaji. S. (2005) Extremophilic microbes: Diversity and perspectives. *Current Science* 89: 78-90.
- Schloss P. D., e Handelsman, J. (2008) A statistical toolbox for metagenomics: assessing functional diversity in microbial communities. *BMC Bioinformatics* 9:34.
- Schloss, P. D., e Handelsman, J. (2003) Biotechnological prospects from metagenomics. *Current opinion in biotechnology*. 14: 303-310
- Schmeisser, C., Steele, H., e Streit, W. (2007) Metagenomics, biotechnology with non-culturable microbes. *Applied Microbiology and Biotechnology*. 75: 955–962.
- Selengut, J. D., Haft, D. H., Davidsen, T., Ganapathy, A., Gwinn-Giglio, M., Nelson, W., Richter, C. A., White, R. O. (2007). TIGRFAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. *Nucleic Acids Research* 35:D260–D264.
- Seshadri, R., Kravitz, S. A., Smarr, L., Gilna, P., Frazier, M. (2007) CAMERA: A Community Resource for Metagenomics. *PLoS Biology* 5: e75.
- Shinoda, Y., Sakai, Y., Uenishi, H., Uchihashi, Y., Hiraishi, A., Yukawa, H., Yurimoto, H., Kato, N. (2004) Aerobic and anaerobic toluene degradation by a newly isolated denitrifying bacterium, *Thauera* sp. strain DNT-1. *Applied Environmental Microbiology* 70: 1385–1392.
- Simon, C., e Daniel, R. (2009) Achievements and new knowledge unraveled by metagenomic approaches. *Applied Microbiology and Biotechnology*. 85: 265–276.

- Sjöling, S., Sttaford, W., Cowan, D. (2007) Soil Metagenomic: Exploring and exploiting the soil Microbial Gen Pool. In: Van Elsas, J.D., Janson, J., Trevors, J. (eds). *Modern Soil Microbiology*, Second Edition. CRC Press, New York.
- Smits, T. H. M., Balada, S. B., Witholt, B., van Beilen, J. B. (2002) Functional analysis of alkane hydroxylases from Gram-negative and Grampositive bacteria. *Journal of Bacteriology* 184: 1733–1742
- So, C. H., Phelps, C. D., Young, L. Y. (2003) Anaerobic transformation of alkanes to fatty acids by a sulfate reducing bacterium, strain Hxd3. *Applied and environmental microbiology* 69: 3892-3900.
- Song, F., Zhuang, Z., Finci, L. et al. (2006) Structure, function, and mechanism of the phenylacetate pathway hot dog-fold thioesterase Paal. *Journal of biological chemistry* 281: 11028-11038
- Steele, H., e Streit, W. (2005) Metagenomics: advances in ecology and biotechnology. *FEMS Microbiology Letters* 247: 105-111.
- Suenaga, H., Koyama, Y., Miyakoshi, M., Miyazaki, R., Yano, H., Sota, M., et al. (2009) Novel organization of aromatic degradation pathway genes in a microbial community as revealed by metagenomic analysis. *ISME Journal* 3: 1335–1348.
- Suenaga, H., Ohnuki, T., and Miyazaki, K. (2007) Functional screening of a metagenomic library for genes involved in microbial degradation of aromatic compounds. *Environmental Microbiology* 9: 2289–2297.
- Sun, Y., Chen Z., Xu S., Cai P. (2005) Stable carbon and hydrogen isotopic fraction of individual n-alkanes accompanying biodegradation: evidence from group of progressively biodegraded oils. *Organical Geochemistry* 36: 225–238.
- Suzek, B. E., Ermolaeva, M. D., Schreiber, M. e Salzberg, S. L. (2001) A probabilistic method for identifying start codons in bacterial genomes. *Bioinformatics* 17: 1123–1130
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., e Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* (accepted).
- Tatusov, R. L., Natale, D. A., Garkavtsev, I. V., Tatusova, T. A., Shankavaram, U. T., Rao, B. S., Kiryutin, B., Galperin, M. Y., Fedorova, N. D., Koonin, E. V. (2001) The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Research* 29: 22–28.
- Tatusov, R. L., Koonin, E. V., Lipman, D. J. (1997) A genomic perspective on protein families. *Science* 278:631–637.
- Tatusov, T. e Tatusov, R. (2009) ORF Finder (Open Reading Frame Finder). [<http://www.ncbi.nlm.nih.gov/projects/gorf/>]. National Center for Biotechnology Information
- Thoden, J. B., Zhuang, Z., Dunaway-Mariano, D. e Holden, H. M. (2003) The structure of 4-hydroxybenzoyl-CoA thioesterase from *Arthrobacter* sp strain SU. *Journal of biological chemistry*. 278: 43709-43716
- Thompson, J. D., Higgins, D. G., Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673–4680.
- Throne-Host, M. Markussen, S. Winnberg, A., Ellingsen, T. E., Kotlar, H. K e Zotchev, S. B. (2006) Utilization of n-alkanes by a newly isolated strain of *Acinetobacter venetianus*: the role of two AlkB-type alkane hydroxylases. *Applied Microbiology and Biotechnology* 72: 353–360.
- Torsvik, V., Daae, F. L., Sandaa, R-A. and Øvreås, L. (1998) Novel techniques for analyzing microbial diversity in natural and perturbed environments. *Journal Biotechnology* 64: 53-62.
- Torsvik, V., Goksoyr, J., Daae, F. L. (1990) High diversity in DNA of soil bacteria. *Applied Environmental Microbiology* 56: 782-787.
- van Beilen, J. B., e Funhoff, E. G. (2005) Expanding the alkane oxygenase tool box: new enzymes and applications. *Current opinion in biotechnology* 16: 308–314.
- van Beilen, J. B., e Funhoff, E. G. (2007) Alkane hydroxylases involved in microbial alkane degradation. *Applied Microbiology and Biotechnology* 74:13–21.

- van Beilen, J. B., Li, Z., Duetz, W. A., Smits, T. H. M., Witholt, B. (2003) Diversity of alkane hydroxylase systems in the environment. *Oil and Gas Science and Technology* 58: 427–440
- van Beilen, J. B., Marin M. M., Smits, T. H., Röthlisberger, M., Franchini, A. G., Witholt, B., Rojo, F. (2004) Characterization of two alkane hydroxylase genes from the marine hydrocarbonoclastic bacterium *Alcanivorax borkumensis*. *Environmental Microbiology* 6: 264–273.
- van Beilen, J. B., Panke, S., Lucchini, S., Franchini, A. G., Röthlisberger, M., Witholt, B. (2001) Analysis of *Pseudomonas putida* alkane degradation gene clusters and flanking insertion sequences: evolution and regulation of the alk-genes. *Microbiology* 147: 1621–1630.
- van Beilen, J. B., Wubbolts, M. G.; Witholt, B. (1994) Genetics of alkane oxidation by *Pseudomonas oleovorans*. *Biodegradation* 5: 161–174.
- van Hamme, J. D., Singh, A., Ward, O. P. (2003) Recent advances in petroleum microbiology. *Microbiology and Molecular Biology Reviews* 67: 503-549.
- van Hellemond, E. W., Janssen, D. B. e Fraaije, M. W. (2007) Discovery of a novel styrene monooxygenase originating from the metagenome. *Applied Environmental Microbiology* 73: 5832–5839.
- Vasconcellos, S. P, Crespim, E., Cruz, G. F., Senatore, D. B., Simioni, K. C. M., Santos Neto, E. V., Marsaioli, A. J., Oliveira, V. M. (2009) Isolation, biodegradation ability and molecular detection of hydrocarbon degrading bacteria in petroleum samples from a Brazilian offshore basin. *Organic Geochemistry*. 40: 574-588.
- Vasconcellos, S. P., Angolini, C. F. F., García, I. N. S., Dellagnezze, B. M., Silva, S. C., Marsaioli, A. J. Santos, E. V. N., Oliveira, V. M. (2010) Screening for hydrocarbon biodegraders in a metagenomic clone library derived from Brazilian petroleum reservoirs. *Organic Geochemistry* 41: 675-681.
- Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., et al. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304: 66–74.
- Vilchez-Vargas, R., Junca, H., Pieper, D. H. (2010) Metabolic networks, microbial ecology and 'omics' technologies: towards understanding in situ biodegradation processes. *Environmental Microbiology* 12: 3089-3104.
- Voget, S., Leggewie, C., Uesbeck, A., Raasch, C., Jaeger, K., e Streit, W. (2003) Prospecting for Novel Biocatalysts in a Soil Metagenome. *Applied and Environmental Microbiology* 69: 6235–6242
- Voget, S., Steele. H. L., Streit, W. R. (2006) Characterization of a metagenome-derived halotolerant cellulose. *Biotechnology Journal* 126: 26–36.
- Von der Weid, I., Korenblum, E., Jurelevicius, D., Rosado, A. S., Dino, R., Sebastián, G. V., Seldin, L. (2008) Molecular diversity of bacterial communities from seafloor rock samples in a deep-water production basin in Brazil. *Journal of Microbiology and Biotechnology* 18, 5-14.
- Watanabe, K. e Kasai, Y. (2008) Emerging technologies to analyze natural attenuation and bioremediation. In: Microbial biodegradation: Genomics and molecular biology. Diaz, E. (ed.). Norfolk, UK Caister academic press, p. 298-318.
- Weis, V. A. Estratégias de finalização da montagem do genoma da bactéria diazotrófica endofítica *Herbaspirillum seropedicae* SmR1. (2010) 71p. Dissertação (Mestrado) - Programa de Pós-Graduação em Ciências – Bioquímica, do Setor de Ciências Biológicas da Universidade Federal do Paraná, Curitiba, 2010.
- Weizhong, L. (2009) Analysis and comparison of very large metagenomes with fast clustering and functional annotation. *BMC Bioinformatics*. 10: 359-368.
- Wentzel, A., Ellingsen, T. E., Kotlar, H. K., Zotchev, S. B., Throne-Holst, M. (2007) Bacterial metabolism of long-chain n-alkanes. *Applied Microbiology and Biotechnology* 76: 1209–1221.
- Whyte, L. G., Hawari, J., Zhou, E., Bourbonnière, L., Innis, W. E., Greer, C. W. (1998) Biodegradation of variable-chain-length alkanes at low temperatures by a psychrotrophic *Rhodococcus* sp. *Applied and Environmental Microbiology* 64: 2578–2584.
- Widdel, F., e Rabus, R. (2001) Anaerobic biodegradation of saturated and aromatic hydrocarbons. *Current Opinion Biotechnology* 12: 259–276.

- Widdel, F., e Grundmann, O. (2010) Biochemistry of the anaerobic degradation of non-methane alkanes. In: Handbook of hydrocarbon and lipid Microbiology. Timmis, K.N. (ed.). Berlin, Heidelberg, Germany: Springer-Verlag, p. 909-924
- Widdel, F., e Musat, F. (2010) Diversity and common principles in enzymatic activation of hydrocarbons. In Handbook of hydrocarbon and lipid microbiology. Timmis, K.N. (ed). Berlin: Springer Verlag. p. 983-1009.
- Wischgoll, S., Heintz, D., Peters, F., Erxleben, A., Sarnighausen, E., Reski, R., Van Dorselaer, A., Boll, M. (2005) Gene clusters involved in anaerobic benzoate degradation of *Geobacter metallireducens*. *Molecular Microbiology* 58: 1238–1252.
- Wischgoll, S., Taubert, M., Peters, F., Jehmlich, N., von Bergen, M., Boll, M. (2009) Decarboxylating and nondecarboxylating glutaryl-coenzyme A dehydrogenases in the aromatic metabolism of obligately anaerobic bacteria. *Journal of Bacteriology*. 191: 4401-4409.
- Wöhlbrand, L., Kallerhoff, B., Lange, D., Hufnagel, P., Thiermann, J., Reinhardt, R., Rabus R. (2007) Functional proteomic view of metabolic regulation in “*Aromatoleum aromaticum*” strain EbN1. *Proteomics* 7: 2222-2239
- Wolfgang, L., Strunk, O., Klugbauer, S. et al. (1998) Bacterial phylogeny based on comparative sequence analysis. *Electrophoresis* 19: 554–568.
- Wu, C. e Sun, B. (2009) Identification of novel esterase from metagenomic library of Yangtze river. *Journal of Microbiology Biotechnology* 19:187-193.
- Xu, M., Xiao, X., Wang, F. (2008) Isolation and characterization of alkane hydroxylases from a metagenomic library of Pacific deep-sea sediment. *Extremophiles* 12:255–262
- Yuste, L., Corbella, M. E., Turiegano, M. J., Karlson, U., Puyet, A., Rojo, F. (2000) Characterization of bacterial strains able to grow on high molecular mass residues from crude oil processing. *FEMS Microbiology Ecology* 32: 69–75.
- Zhang, T., Breitbart, M., Lee, W. H., Run, J. Q., Wei, C. L., Soh, S. W., et al. (2006) RNA viral community in human feces: prevalence of plant pathogenic viruses. *PLoS Biology* 4: e3.
- Zhu, W., Lomsadze, A., Borodovsky, M. (2010) Ab initio gene identification in metagenomic sequences. *Nucleic Acids Research* 38: e132
- Zinder, S. H., Cardwell, S. C., Anguish, T., Lee, M., Koch, M. (1984) Methanogenesis in a thermophilic (58 °C) anaerobic digester: Methanotrix sp. as an important acetoclastic methanogen. *Applied and Environmental Microbiology* 47: 796–807.