

Luciano Procópio da Silva

**Estudos fisiológicos e moleculares da estirpe degradadora
de hidrocarbonetos do petróleo *Dietzia cinnamea* P4**

Orientadora: Dra. Lucy Seldin



TESE APRESENTADA AO INSTITUTO DE
MICROBIOLOGIA PAULO DE GÓES –
UNIVERSIDADE FEDERAL DO RIO DE
JANEIRO, VISANDO A OBTENÇÃO DO GRAU
DE DOUTOR EM CIÊNCIAS BIOLÓGICAS
(MICROBIOLOGIA)

**UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE MICROBIOLOGIA PAULO DE GÓES
RIO DE JANEIRO**

2010

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

PROCÓPIO da SILVA, LUCIANO

ESTUDOS FISIOLÓGICOS E MOLECULARES DA ESTIRPE DEGRADADORA DE
HIDROCARBONETOS DO PETRÓLEO *DIETZIA CINNAMEA* P4/ Luciano Procópio da
Silva – Rio de Janeiro, 2010

xvii, 135

Tese: Doutorado em Ciências (Microbiologia)

Universidade Federal do Rio de Janeiro/ Instituto de Microbiologia Paulo de Góes, 2010.

Orientador: Lucy Seldin

Referências bibliográficas: 1773

1. Microbiologia do petróleo 2. Biorremediação 3. *Dietzia cinnamea* P4 4. Genoma 5. PCR
em tempo real 6. Alcanos 7. Biorreator

Procópio da Silva, Luciano

UFRJ, Instituto de Microbiologia Paulo de Góes, Doutorado em Ciências (Microbiologia).

LUCIANO PROCÓPIO DA SILVA

Rio de Janeiro, 26 de julho de 2010

(Profa. Dra. Lucy Seldin, IMPPG/UFRJ)

(Prof. Dr. Andrew Macrae, IMPPG/UFRJ)

(Prof. Dr. Orlando Martins, IB/UFRJ)

(Profa. Dra. Magali Christe Cammarota, EQ/UFRJ)

(Prof. Dr. Alexandre Rosado, IMPPG/UFRJ)

Este trabalho foi realizado no Laboratório de Genética Microbiana, Departamento de Microbiologia Geral, Instituto de Microbiologia Prof. Paulo de Góes da UFRJ, Rio de Janeiro, Brasil, sob orientação da Profa. Dra. Lucy Seldin com colaboração do Prof. Dr. Jan Dirk van Elsas da Universidade de Groningen - Holanda

*Em memória à minha mãe Yara.
(Eis que eu estarei com vocês todos os dias, até o fim do mundo. Mateus 28,18-20)*

Agradecimento especial

A aproximadamente um ano atrás minha mãe foi atropelada às 7:00 horas da manhã por um motorista bêbado, em alta velocidade sem habilitação e com um carro emprestado de terceiro e o pior: fugiu sem prestar socorro. Após um resgate surpreendentemente rápido pela equipe do SAMU, minha mãe foi esquecida em uma enfermaria de emergência por mais de oito horas, amparada apenas pelos braços de meu irmão. A partir daquele momento uma legião de pessoas desconhecidas, amigas e familiares mobilizou em uma operação para tentar salvá-la. Minha esposa grávida de nove meses teve que gerenciar todo meu trajeto entre a Europa e a cidade onde minha mãe estava hospitalizada. Meu irmão ficou do lado de minha mãe o tempo todo tomando decisões difíceis como transferência entre hospitais e recorrendo ao poder judiciário e policial para que ela recebesse os cuidados necessários. Após três semanas desacordada em uma UTI, minha mãe faleceu sem eu poder me despedir, sem que ela pudesse conhecer a sua neta e deixando ainda sua última lição: que ainda devemos acreditar nas pessoas.

*Hoje eu sei que não teria forças para continuar com este trabalho se não fosse o amparo de todas as pessoas que nos ajudaram, em especial minha esposa **Joana** que me ajudou o tempo todo e me presenteou com a minha filha **Maria Flor**, a maior felicidade de minha vida; ao meu irmão **Michael** que teve a coragem e a força que eu nunca conseguiria e a minha cunhada **Claudiane** que foi uma “rocha” de serenidade onde encontramos apoio.*

Por isso eu também dedico este trabalho a todas as pessoas de boa fé.

Agradecimentos

Agradeço a Deus por me iluminar em todos os momentos de dúvida.

Agradeço imensamente a minha orientadora Lucy Seldin, por ter me recebido de braços abertos em seu laboratório, acreditado em mim e em meu trabalho e pela amizade, apoio e orientação em todas as etapas de meu desenvolvimento profissional e pessoal. Ao meu orientador Jan Dirk van Elsas, por ter me recebido em seu laboratório e em sua casa, pela orientação científica e futebolística!! Agradeço também pela imensa amizade. Eu realmente espero poder continuar trabalhando com vocês com muito tempo ainda.

Aos participantes da banca examinadora por ter aceitado meu convite.

Aos meus sempre e eternos (e distantes) amigos Thiago, Michele, Sasha e Joana Salles. Espero e acredito que ainda vamos encontramos sempre.

A minha minha querida prima Eugênia e sua mãe, minha tia Edith por terem me ajudado no início de minha vida aqui no Rio de Janeiro.

Aos meus amigos e colegas de laboratório Márcia, Elisa, Vanessa, Dju, Diobs, Simone, Joana, Juliana, Sílvia, Renata e Natália pelo apoio e ajuda que foram imprescindíveis durante estes quatro anos. A técnica Janinha sempre prestativa e presente.

Ao Instituto de Microbiologia Prof. Paulo de Góes, aos professores e funcionários pelas condições de trabalho oferecidas.

A todos que não foram citados nominalmente, mas que certamente contribuíram para a realização deste trabalho.

Resumo

Em um estudo prévio sobre o efeito da bioestimulação na diversidade bacteriana em solo contaminado com petróleo em uma região da Mata Atlântica, Brasil, foi isolada uma estirpe que foi denominada P4 e identificada como pertencente à espécie *Dietzia cinnamea*. Essa estirpe é Gram-positiva, aeróbia estrita, imóvel e não formadora de esporos, capaz de crescer até 45°C, pH 10 e meio contendo 10% de NaCl e capaz de degradar tanto hidrocarbonetos alifáticos quanto hidrocarbonetos aromáticos do petróleo. No presente trabalho, a estirpe P4 teve o seu genoma pirosequenciado, dando origem a 435 *supercontigs*. A anotação parcial do genoma mostrou a presença de 3.557.916 pb arranjados em um único cromossoma com um conteúdo de G+C igual a 70,9%. *D. cinnamea* P4 é a primeira espécie do gênero *Dietzia* a ter seu genoma sequenciado. A análise do genoma sequenciado demonstrou a presença de todos os genes para a utilização da glicose como fonte de carbono, através da via glicolítica, bem como a presença dos genes envolvidos nos ciclos de ácido cítrico, gliconeogênese, metabolismo de ácidos graxos e glioxilato. O sequenciamento também confirmou a capacidade da estirpe P4 em utilizar diversos tipos de hidrocarbonetos alifáticos, cíclicos e aromáticos como fontes de carbono. Foi demonstrada também a presença de genes necessários para autotrofia por meio da oxidação de monóxido de carbono, fixação de CO₂ através dos ciclos de regeneração do glioxilato e β-hidroxiopropionato. O seqüenciamento do genoma da estirpe P4 também mostrou a presença de genes necessários para a utilização de diferentes açúcares, bem como para a quitina e celulose, como únicas fontes de carbono. Uma análise mais detalhada de seu genoma revelou seu potencial na produção de metabólitos secundários como terpenos, β-caroteno, antibióticos (frenolicina, nanaomicina e avilamicina A) e um potente antitumor, a azinomicina B. Os resultados sobre os aspectos fisiológicos e genéticos em condições de biorreator em diferentes condições mostraram que a estirpe P4 é capaz de crescer em diferentes níveis de pHs e temperaturas, além de apresentar alterações no genoma após longo período de cultivo. Além disso, foram realizados experimentos de PCR quantitativo com os principais genes envolvidos na degradação de hidrocarbonetos alifáticos, *alkS*, *alkB*, *rubA* e *alkTrans* e os resultados mostraram que estes são diferencialmente expressos quando a estirpe P4 é crescida em *n*-hexadecano como única fonte de carbono. Os resultados ora apresentados trazem novas informações sobre a genética e a fisiologia da estirpe *D. cinnamea* P4, sugerindo seu grande potencial para aplicações em processos de biorremediação e biotecnológicos.

Abstract

In a study on the effects of biostimulation on the bacterial diversity of oil-contaminated soil from Mata Atlântica, Brazil, we isolated *Dietzia cinnamea* strain P4, a strictly aerobic, non-spore forming and non-motile Gram-positive strain, able to grow in temperatures of up to 45°C, pH 10 and 10% NaCl. This strain was shown to be an excellent degrader of several *n*-alkanes and other linear and cyclic hydrocarbon derivatives. In the present study, the draft genome of strain P4 was determined using pyrosequencing, yielding 435 supercontigs, and the main features of this genome were described. The partially assembled genome contained 3,557,916 bp arranged in a single chromosome with an average G+C content of 70.9%. *D. cinnamea* P4 is the first species of the genus *Dietzia* to have its genome sequenced. It revealed the presence of complete pathways for all central metabolic routes. The complete sets of genes for the glyoxalate and reductive carboxylate cycles were also detected. The presence of genes coding for aerobic CO oxidation, formate/formaldehyde oxidation, reverse tricarboxylic acid cycle and 3-hydropropionate cycle for CO₂ fixation suggested its autotrophic growth. The strain P4 genome further revealed the presence of genes for utilization of diverse sugars as well as chitin and cellulose compounds as sole carbon sources. Secondary metabolism was suggested to occur due to the presence of genes for the biosynthesis of terpene compounds, frenolicin, nanaomycin and avilamycin A antibiotics. Furthermore, genes encoding azinomycin B synthesis, an important product with antitumor activity, were also observed. The complete Alk operon, responsible for *n*-alkane degradation, as well as clusters of genes coding for biphenyl ring dihydroxylation were found to be present in P4 genome. The cultivation of P4 in a bioreactor under different conditions revealed that P4 strain is able to grow at different pH levels and temperatures, and presented changes in its genome after long period of cultivation. Moreover, qPCR experiments were performed showing that the genes *alkS*, *alkB* and a putative transporter gene *alkTrans* were highly expressed when P4 strain grew in *n*-hexadecane as the only carbon source. This study brings new insights into the genetics and physiology of P4 strain, suggesting its relevance in biotechnology and bioremediation applications.

Índice

Resumo	viii
Abstract	ix
Índice	x
Lista de figuras	xiii
Lista de tabelas	xvi
Lista de abreviaturas	xvii
1-Introdução	1
1.1 O petróleo	2
1.2 A biorremediação	3
1.3 Breve histórico sobre a era genômica	<u>6</u>
1.4 A terceira geração de sequenciadores	<u>11</u>
1.5 A microbiologia na era genômica	<u>15</u>
1.6 A contribuição da técnica de PCR em tempo real no estudo da biorremediação	<u>17</u>
1.7 Degradação de alcanos	<u>18</u>
1.8 <u>O sistema AlkB</u>	<u>22</u>
<u>1.9 Os genes do sistema AlkB no ambiente</u>	<u>24</u>
1.10 O gênero <i>Dietzia</i>	<u>25</u>
2- Justificativa	26
3- Objetivos	27
4- Etapas desenvolvidas	<u>28</u>
5- Materiais e Métodos	<u>29</u>
<u>5.1 Estirpes bacterianas</u>	<u>29</u>
<u>5.2 Meios de Cultura - <i>Escherichia coli</i> e <i>D. cinnamea</i> P4</u>	<u>29</u>
5.3 Soluções	<u>30</u>
5.4 Construções das curvas de crescimento de <i>D. cinnamea</i> P4 em diferentes fontes de carbono, obtenção do DNA genômico e sequenciamento por meio da técnica de pirosequenciamento.....	<u>30</u>
5.5 Construção das sequências codificantes de proteínas (CDS) presentes no genoma da estirpe P4 utilizando programas de bioinformática e análise das similaridades entre as CDSs de <i>D. cinnamea</i> P4 e as sequências de aminoácidos de organismos relacionados depositados em bancos de dados	<u>32</u>

5.6 Avaliação do tempo de geração celular da estirpe P4 em condições de biorreator e prováveis gerações de populações distintas da estirpe P4 nestas mesmas condições, por meio da análise morfológica das colônias e mudança de pigmentação	34
5.7 Análise das alterações no genoma das populações evoluídas por meio de BOX-PCR	37
5.8 Obtenção do RNA total da estirpe P4 crescida em diferentes condições de crescimento em biorreator de 2 litros, síntese dos cDNAs referentes aos RNAs e análise dos níveis de expressão dos prováveis genes <i>alkS</i> , <i>alkB</i> e <i>alkTrans</i> de <i>D. cinnamea</i> P4 por meio de PCR quantitativo em tempo real	38
6- Resultados	43
6.1 – Características gerais do genoma de <i>D. cinnamea</i> P4	43
6.2 – Similaridades entre as CDS de <i>D. cinnamea</i> P4 com actinomicetos relacionados	46
6.3 – Metabolismo central	48
6.4 – Fosforilação oxidativa	52
6.5 – Aquisição de energia	53
6.6 – Fixação autotrófica de carbono	56
6.7 – Metabolismo de hidrocarbonetos	60
6.8 – Biodegradação de compostos nitrogenados	65
6.9 – Degradação de colesterol	70
6.10 – Oxigenases	72
6.11 – Assimilação de nitrogênio	76
6.12 – Assimilação de compostos inorgânicos	77
6.13 – Enzimas celulolíticas	79
6.14 – Metabolismo secundário	82
6.15 – Transporte transmembrânico	84
6.16 – Duplicação, reparo e recombinação do DNA	87
6.17 – Transcrição e tradução	88
6.18 – Respostas a estresses	88
6.19 – Biossíntese da parede celular	94
6.20 – Crescimento em diferentes meios de cultura	95
6.21 – Crescimento em C-16 em diferentes condições	96
6.22 – Crescimento em C-16 por longo período de cultivo	98
6.23 – Análise das alterações morfológicas e genômicas	100

6.24 – Análise dos níveis de expressão dos prováveis genes *alkS*, *alkB* e *alkTrans*

.....	102
<u>7. Discussão</u>	<u>105</u>
8. Conclusões	117
<u>9. Referências bibliográficas</u>	<u>118</u>
<u>10. Anexo</u>	<u>134</u>

Lista de figuras

Figura 1: Esquema representativo da formação do petróleo	2
Figura 2: Número de espécies bacterianas com genomas sequenciados e agrupados por classe, depositados no <i>National Center for Biotechnology Information</i> (NCBI).....	7
Figura 3: Esquema representando os principais passos da técnica de sequenciamento automático por capilaridade, utilizando a metodologia de Sanger	9
Figura 4: Esquema representativo das principais técnicas de sequenciamento de terceira geração	12
Figura 5: Via metabólica de Alk descrita em <i>Pseudomonas putida</i> (<i>oleovorans</i>) Gpo1	19
Figura 6: Controle sobre a expressão dos <i>operons alkBFGHJKL</i> e <i>alkST</i> em <i>P. putida</i> GPo1 por meio das proteínas Crc e Cyo (RUIZ-MANZANO, YUSTE & ROJO, 2005).....	21
Figura 7: Modelo proposto para a proteína transmembrana AlkB em <i>P. putida</i> GPo1 (ROJO, 2005).....	23
Figura 8: Perfil do DNA genômico de <i>Dietzia cinnamea</i> P4	31
Figura 9: Fluxograma representando as etapas das análises das sequência de <i>contigs</i> (DNA) e das CDSs (sequências de aminoácidos) de <i>D. cinnamea</i> P4 com os diversos bancos de sequências disponíveis	33
Figura 10: Sistema de fermentação de bancada composto por um birreator de 2 litros, 3 bombas peristálticas, sistemas de transferências de meio de cultura e fonte de carbono e sistema de monitoramento de dados em tempo real	35
Figura 11: Número de <i>contigs</i> obtidos com diferentes tamanhos após o sequenciamento do genoma da estirpe P4	43
Figura 12: Distribuição dos nucleotídeos do genoma da estirpe P4	44
Figura 13: Distribuição funcional dos prováveis genes presentes no genoma de <i>D. cinnamea</i> P4	46
Figura 14: Distribuição das CDSs de <i>D. cinnamea</i> P4 com similaridade com organismos relacionados	47
Figura 15: Esquema representativo da via central glicolítica com os genes presentes no genoma da estirpe P4	49
Figura 16: Esquema mostrando o ciclo do glioxilato (ciclo menor) sugerido pela anotação do genoma de <i>D. cinnamea</i> P4	50

Figura 17: Análise filogenética da provável subunidade L de CODH (DCP40505) de <i>D. cinnamea</i> P4 com as formas I e II da subunidade L de CODH de outras bactérias relacionadas	54
Figura 18: Esquema representativo do complexo Nar de <i>D. cinnamea</i> P4 empregado na obtenção de energia utilizando nitrato como aceptor final na cadeia de transporte de elétrons	56
Figura 19: O ciclo do ácido cítrico reverso sugerido no genoma da estirpe P4 de <i>D. cinnamea</i>	57
Figura 20: Ciclo do 3-hidroxiacetoato proposto pela análise do genoma da estirpe P4. Duas moléculas de CO ₂ são condensadas com uma molécula de Acetil-CoA, produzindo glicoxalato	58
Figura 21: Esquema representativo das prováveis vias envolvidas na assimilação dos compostos de C1 presente no genoma de <i>D. cinnamea</i> P4	59
Figura 22: Vias metabólicas propostas no genoma de <i>D. cinnamea</i> P4. (a) Via metabólica de compostos de hidrocarbonetos alifáticos	61
Figura 23: A via de degradação de compostos de bifenil proposta pela anotação dos prováveis genes presente no genoma de <i>D. cinnamea</i> P4	64
Figura 24: Esquema representativo da reação de hidroxilação de compostos de 4-hidroxiacetofenona (KAMERBEEK <i>et al.</i> , 2001)	65
Figura 25: Análise filogenética entre as sequências de aminoácidos DCP41901-03 de <i>D. cinnamea</i> P4 com sequências das enzimas nitril hidratase de organismos relacionados:	66
Figura 26: Esquema proposto para a reação de hidroxilação de nitriloacetato (adaptado de KIM <i>et al.</i> , 2006)	67
Figura 27: Fórmula da reação catalisada pela enzima nitroalcano dioxigenase. (a) Substrato da reação, nitroalcano, oxigênio molecular e água. (b) Produtos da reação, aldeído ou cetona, peróxido de hidrogênio, nitrito e íons	67
Figura 28: Alinhamento entre as sequências deduzidas de aminoácidos das CDSs DCP41915-19 de <i>D. cinnamea</i> P4 utilizando o programa CLUSTALW disponível no BioEdit apresenta a baixa similaridades entre as cinco sequências	69
Figura 29: Via de degradação de colesterol proposta para <i>D. cinnamea</i> P4 (adaptado de van der GEIZE <i>et al.</i> , 2007)	71
Figura 30: Análise filogenética entre as sequências deduzidas de aminoácidos DCP41295 e DCP41296 de <i>D. cinnamea</i> P4 e as sequências de enzimas colesterol oxidase de organismos	

relacionados, utilizando o programa PhyML disponível no sítio http://www.phylogeny.fr/version2_cgi/simple_phylogeny.cgi	72
Figura 31: Esquema representativo do sistema de assimilação de enxofre por meio dos sistemas Tau e Cys provavelmente presente no genoma de <i>D. cinnamea</i> P4	78
Figura 32: Esquema representativo sugerindo as via biossintéticas dos compostos de geranyl, undecaprenil, fitil e β -caroteno presente no genoma de <i>D. cinnamea</i> P4	83
Figura 33: Distribuição das classes de prováveis transportadores presentes no genoma de <i>D. cinnamea</i> P4, classificados através das ferramentas BLAST, BALSTALL, TMHMM e TransportDB (veja material e métodos)	85
Figura 34: Análise filogenética entre as sequências de aminoácidos DCP43195e DCP43196 de <i>D. cinnamea</i> P4 com sequências deduzidas de aminoácidos de SecA1 e SecA2 de organismos relacionados	86
Figura 35: Curva de crescimento celular de <i>D. cinnamea</i> P4 em meio líquido L.B	95
Figura 36: Curva de crescimento celular de <i>D. cinnamea</i> P4 em meio líquido B.H. contendo C-16	96
Figura 37: Curvas de crescimentos de <i>D. cinnamea</i> P4 em birreator de 2 litros em diferentes pHs a 35°C	97
Figura 38: Curvas de crescimentos de <i>D. cinnamea</i> P4 em birreator de 2 litros em diferentes pHs a 45°C	98
Figura 39: Acompanhamento da manutenção celular de <i>D. cinnamea</i> P4 utilizando C-16 como única fonte de carbono em birreator de 2 litros	99
Figura 40: Perfil da amplificação do genoma de oito isolados oriundos do experimento de manutenção celular por longo período de cultivo em biorreator de 2 litros	100
Figura 41: Perfil da amplificação do genoma de quatro concentrações diferentes dos DNAs de IS6, <i>mix</i> e ISW	101
Figura 42: Níveis de expressão do provável gene <i>alkS</i> de <i>D. cinnamea</i> P4 crescida em C-16, em relação ao nível de expressão durante crescimento em glicose	102
Figura 43: Níveis de expressão do provável gene <i>alkB</i> de <i>D. cinnamea</i> P4 crescida em C-16, em relação ao nível de expressão durante crescimento em glicose	103
Figura 44: Níveis de expressão do provável gene <i>alkTrans</i> de <i>D. cinnamea</i> P4 crescida em C-16, em relação ao nível de expressão durante crescimento em glicose	104

Lista de tabelas

Tabela 1: Comparação de características selecionadas entre o genoma de <i>D. cinnamea</i> P4 e outras bactérias relacionadas	45
Tabela 2: Divisão das CDS de <i>D. cinnamea</i> P4 que apresentaram similaridades com as proteínas que compõem o sistema de fosforilação oxidativa	52
Tabela 3: Resultado da análise de similaridade das sequências de deduzidas de aminoácidos DCP41915-19, comparadas às sequências depositadas no banco NR, utilizando a ferramenta BLASTP (http://www.ncbi.nlm.nih.gov/BLAST/)	68
Tabela 4: Prováveis oxigenases identificadas no genoma de <i>D. cinnamea</i> P4	73
Tabela 5: Resultado da análise de similaridade entre as sequências deduzidas de aminoácidos de <i>D. cinnamea</i> P4 e as sequências depositadas no banco de dados CAZy, disponível no sítio http://ww.cazy.org	80
Tabela 6: Principais processos de defesa a estresse oxidativo sugerido pela anotação do genoma de <i>D. cinnamea</i> P4	90
Tabela 7: Classificação das prováveis proteínas envolvidas na resistência a drogas presentes no genoma de <i>D. cinnamea</i> P4	93

Lista de abreviaturas

16S rDNA – fragmento oriundo da amplificação por PCR do gene codifica a subunidade menor do ribossomo bacteriano

16S rRNA – parte da sequência da subunidade menor do RNA ribossomal bacteriano

BSA – soro albumina bovina

DNA – ácido desoxirribonucléico

dNTP – deoxinucleotídeo trifosfato

EDTA – ácido etilenodiaminotetracético

EMBOSS – *The European Molecular Biology Open Software Suite* (EMBOSS, banco de dados europeu)

ex. – por exemplo

FAD – flavina adenina dinucleotídeo

GC – nucleotídeos Guanina e Citosina

kb – quilo base (1000 pb)

ml – mililitros

mRNA – RNA mensageiro

NCBI – do inglês *National Center for Biotechnology Information* (Centro Nacional de Informação Biotecnológica, banco de dados americano)

ng – nanogramas

pb – pares de bases (ou bp – em inglês: *base pairs*)

PCR – do inglês *Polymerase Chain Reaction* (Reação em cadeia da polimerase)

RNA – ácido ribonucléico

rRNA – RNA ribossômico

RT-PCR – *Reverse transcriptase –PCR* (PCR com enzima transcriptase reversa)

seg – segundos

UFRJ – Universidade Federal do Rio de Janeiro

UV – ultravioleta

1. Introdução

A poluição é considerada, juridicamente, como a entrada de qualquer fator ao ambiente que provoque alteração de suas qualidades naturais. A lei federal nº 6.938/81 define poluição como “toda alteração das propriedades físicas, químicas e biológicas que possa constituir prejuízo à saúde, à segurança e ao bem-estar das populações e, ainda, possa comprometer a biota e a utilização dos recursos para fins comerciais e recreativos”. O ponto de vista científico, a poluição é definida como a entrada de elementos, compostos ou energia no ambiente em níveis prejudiciais ao funcionamento do sistema vivo local. Neste contexto, poluição é o resultado indesejável das transformações das características naturais de um ambiente, atribuindo um caráter nocivo ao mesmo. Entre os diversos danos decorrentes da poluição ao meio ambiente, os solos estão entre os primeiros elementos a serem afetados. Contudo, enquanto as fontes e depósitos de água são objetos de regulamentação ambiental em alguns países desde o século 19, o solo recebeu atenção similar apenas nas últimas décadas do século 20. Isso se deve principalmente ao fato que a relação direta entre a poluição do solo e a saúde humana ser menos óbvia quando comparada com a água e o ar (SCULLION, 2006). Os solos prejudicados pela presença de poluentes falham como suporte à vegetação, devido aos seus efeitos fitotóxicos ou devido a alterações dos ciclos biogeoquímicos (SIDDIQUI & ADAMS, 2002). Além disso, a poluição dos solos pode alcançar os lençóis freáticos, comprometendo a qualidade da água para consumo.

Os principais responsáveis pela poluição do solo são: (i) indústrias farmacêuticas e químicas, que produzem uma ampla variedade de xenobióticos e polímeros sintéticos; (ii) indústrias envolvidas no branqueamento de polpa de papel, que são as principais geradoras de compostos orgânicos clorados presentes no ambiente; (iii) indústrias de minérios, as quais liberam metais pesados nos ciclos biogeoquímicos; (iv) indústrias agrícolas e a própria agricultura em si, que despejam quantidades enormes de fertilizantes, pesticidas e herbicidas no solo e (v) indústrias petrolíferas, que podem liberar combustíveis fósseis, como carvão e petróleo, em grandes quantidades acidentalmente no ecossistema. Além disso, a combustão destes aumenta significativa concentração de CO₂ e milhares de toneladas de compostos de enxofre e óxido de nitrogênio são lançados na atmosfera, onde sofrem reações químicas e se transformam em ácido nítrico e sulfúrico (RIEGER *et al.*, 2002).

1.1 - O petróleo

O petróleo é originário da decomposição parcial de matéria orgânica, vegetal ou animal, que se deposita no fundo dos mares e lagos em condições anaeróbias. A formação do petróleo ocorre de modo semelhante à formação da hulha, onde os restos vegetais cobertos de argila durante milhões de anos, sob o efeito da pressão e da atmosfera, perdem oxigênio e se enriquecendo de carbono e hidrogênio (Figura 1). O petróleo é uma mistura complexa de hidrocarbonetos e outros componentes como enxofre, nitrogênio e oxigênio, os quais representam menos de 3% (v/v), além de elementos traços como alguns constituintes organometálicos, sobretudo complexos com níquel e vanádio (1%). Estudos recentes em espectrometria de massa de alta resolução permitiram a detecção de mais de 17.000 compostos presentes no petróleo (MARSHALL & RODGERS, 2004). O petróleo pode ser classificado como leve, médio e pesado de acordo com a presença de hidrocarbonetos alifáticos, naftalenos e aromáticos (ATLAS, 1981). Os hidrocarbonetos podem ser divididos em três grandes grupos: os alifáticos, que incluem os compostos de cadeia aberta, os alicíclicos ou ciclanos, onde os seus átomos estão arranjados em uma estrutura em anel, e os hidrocarbonetos aromáticos, os quais contém um ou mais anéis benzeno (LEYTHAEUSER & RÜCKHEIM, 1989). Contudo, recuperados de diferentes reservatórios, o petróleo varia muito na composição e propriedades físico-químicas.

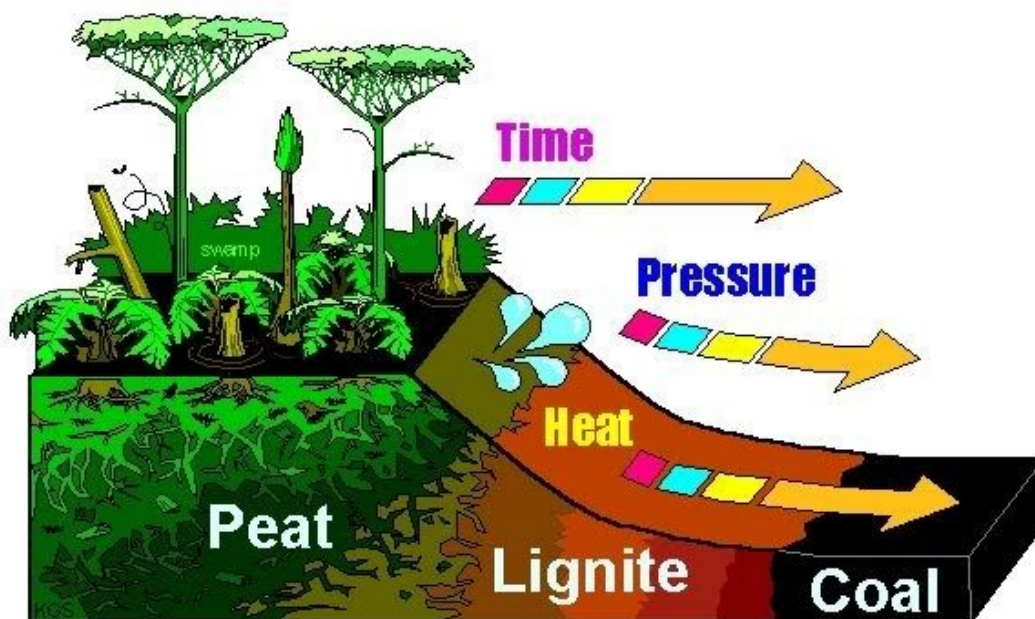


Figura 1: Esquema representativo da formação do petróleo

(http://media.photobucket.com/image/formation%20petroleum/chudetuchon/Ed_coal_formation.jpg).

Os impactos ambientais relacionados às atividades da indústria petrolífera são oriundos da extração, transporte, refino e uso de seus derivados, sendo apenas 10% de tudo atribuído a derrames acidentais (BALBA, AL-AWANDHI & AL-DAHER, 1998). Mesmo assim, aproximadamente 6×10^7 barris de petróleo foram derramados sobre 2×10^7 m³ de solo e cerca de 320 lagos de óleo foram criados no deserto do Kuwait durante a guerra do golfo (AL-SALEH & OBUEKWE, 2005). Atualmente, existem cerca de 12.000 sítios contaminados nos Estados Unidos e aproximadamente 400.000 no oeste da Europa. No total, cerca de US\$ 15 a 18 bilhões de dólares foram gastos na remediação de áreas contaminadas, apenas em 1998, e estes valores crescem a cada ano. A maioria destes sítios são solos e águas subterrâneas, contaminados por uma mistura de metais pesados (mercúrio, cádmio, chumbo e arsênio) e compostos orgânicos, os hidrocarbonetos de petróleo (WATANABE, 2001).

Devido às características dos hidrocarbonetos se aderirem fortemente a superfícies sólidas, como os solos, a remoção e/ou degradação destes compostos é um obstáculo às tecnologias empregadas atualmente. Tecnologias para a remediação de poluentes ambientais que contam com abordagens não biológicas são frequentemente de custos proibitivos e, geralmente, não geram os resultados esperados, podendo levar à formação de intermediários tóxicos com a possibilidade de sua perpetuação no local e entrada na cadeia alimentar ou contaminação de aquíferos (LOVLEY & LLOYD, 2000). Estudos de biodegradação de petróleo com o emprego de microrganismos têm se tornado um foco de pesquisa em diversas áreas da biologia, conhecida como biorremediação (ATLAS, 1981).

1.2 - A biorremediação

A biorremediação pode ser um processo espontâneo ou acelerado, onde mediante procedimentos biológicos, fundamentalmente microbiológicos, ocorre a degradação ou a transformação de contaminantes em formas menos nocivas ou não tóxicas (MORENO, GONZÁLEZ & BLANCO-SANTOS, 2004). Embora leveduras e plantas tenham capacidade de detoxificação, como mineralização, transformação ou imobilização de poluentes, as bactérias têm um papel crucial em ciclos biogeoquímicos de diversos compostos nocivos presentes no ambiente, apresentando um papel fundamental na manutenção sustentável da biosfera. Esta habilidade é atribuída principalmente pela capacidade de bactérias poderem usar aceptores de elétrons alternativos para o oxigênio, além de sua alta taxa de crescimento,

alta relação superfície-volume e plasticidade genômica (DÍAZ, 2004).

A partir de 1940, a microbiologia do petróleo teve um desenvolvimento significativo, principalmente nos Estados Unidos. Por meio de estudos foi possível comprovar o crescimento microbiano na presença de hidrocarbonetos como única fonte de carbono. Em amostras coletadas assepticamente de poços perfurados a uma profundidade de 300 m foi detectada uma ampla diversidade microbiana, incluindo bactérias anaeróbias como as redutoras de sulfato (BRS), archaeas metanogênicas e várias bactérias aeróbias e anaeróbias facultativas. Embora as técnicas de biorremediação tenham sido utilizadas por décadas, esta tecnologia ainda é considerada uma inovação, representando cerca de 10 a 15% dos métodos empregados no tratamento de solos e aquíferos contaminados nos Estados Unidos da América (ATLAS & PHILP, 2005). A “MicroBac International” (Round Rock, Texas, EUA), empresa especialista em remediação com emprego de microrganismos, afirma que o uso de produtos microbiológicos custa em torno de US\$ 25,00/m³, enquanto os custos de tratamentos padrões ou remoção de solos contaminados ficam entre US\$ 50,00 e US\$ 100,00/m³ (WATANABE, 2001).

O sucesso da aplicação de tecnologias de biorremediação em áreas contaminadas por petróleo requer o conhecimento dos parâmetros que afetam as características do poluente, isso porque a biodegradabilidade do petróleo por microrganismos é inerente à sua composição. Óleos pesados são geralmente mais difíceis de serem assimilados por microrganismos quando comparados com óleos leves (OKOH *et al.*, 2002). A temperatura também tem papel importante na biodegradação de hidrocarbonetos de petróleo, primeiramente pelo efeito direto na composição química do poluente. Em baixas temperaturas, ocorre um aumento na viscosidade do petróleo, enquanto a volatilidade dos hidrocarbonetos de baixa massa molecular é reduzida. Além disso, estratégias de biorremediação devem levar em consideração o conhecimento de quais microrganismos estão presentes no ambiente contaminado, sua capacidade metabólica e como os microrganismos irão responder às mudanças nas condições do ambiente como disponibilidade de nutrientes, estado físico-químico dos compostos hidrofóbicos, oxigenação, temperatura, salinidade e pressão (ATLAS, 1981).

Uma das principais razões para a persistência de alguns componentes tóxicos no ambiente é a sua biodisponibilidade limitada. Estudos de biodegradação de contaminantes no solo têm mostrado que sua liberação da matriz sólida para a fase aquosa é frequentemente um

passo limitante neste processo (VOLKERING, BREURE & RULKENS, 1998). A aplicação de biosurfactantes permite reduzir a concentração micelar crítica (CMC) e a tensão interfacial em soluções aquosas e mistura de hidrocarbonetos, permitindo que os hidrocarbonetos possam ser solubilizados (DESAI & BANAT, 1997). O efeito mais importante dos biosurfactantes na interação entre o solo e o poluente, principalmente compostos de alcanos, é a estimulação do transporte de massa do poluente do solo para a fase aquosa. Dois mecanismos são sugeridos pelos quais os biosurfactantes aumentam a biodisponibilidade dos compostos orgânicos hidrofóbicos. Primeiro, os biosurfactantes podem solubilizar componentes hidrofóbicos dentro de estruturas micelares, deste modo aumentando a aparente solubilidade aquosa do poluente, disponibilizando-o para uso pelos microrganismos. Segundo, os biosurfactantes podem causar alterações químicas na estrutura de superfície de células bacterianas, tornando-a mais hidrofóbica, assim aumentando o contato físico direto entre o componente orgânico e o microrganismo. Esta alteração química ocorre principalmente devido à perda de lipopolissacarídeos (LPS) para o meio (AL-TAHHAN *et al.*, 2000). De qualquer modo, os dois mecanismos possibilitam os hidrocarbonetos entrarem na célula para serem utilizados como fontes de carbono e, conseqüentemente, promover sua biodegradação.

As técnicas de biorremediação podem ser classificadas em dois grupos: *ex situ*, que se aplicam às tecnologias que envolvem a remoção do material contaminado e posterior tratamento, e *in situ* onde o tratamento do material contaminado é realizado no local de origem. Diversos métodos de biorremediação são empregados atualmente. A bioatenuação ou atenuação natural caracteriza-se por ser uma remediação passiva, onde o processo consiste em monitorar o declínio do poluente por meio da capacidade intrínseca microbiana presente no local, sem que haja uma interferência do homem. A bioestimulação envolve o aumento da atividade da comunidade microbiana degradadora de xenobióticos, por meio de adição de água e de nutrientes inorgânicos, usualmente nitrogênio, fósforo e potássio, correção do pH e da quantidade de oxigênio. A técnica de bioaumento consiste na aplicação direta de microrganismos previamente selecionados, com capacidade biodegradativa ou modificados geneticamente, no ambiente contaminado. Muitos destes microrganismos selecionados são comercializados por empresas de biorremediação (ATLAS & PHILP, 2005).

Muitos microbiologistas estudiosos dos processos de biorremediação fazem uso dos estudos de *treatability*, onde amostras de ambientes contaminados são armazenadas em laboratório e a taxa de degradação dos contaminantes por microrganismos presentes no inóculo é monitorada. Tais estudos apresentam uma estimativa do potencial metabólico da comunidade microbiana presente no sítio contaminado. Quando há necessidade de um

conhecimento mais detalhado, tentativas são direcionadas para o isolamento do organismo responsável pela biodegradação (ROGERS & McCLURE, 2003).

Um avanço significativo no campo da ecologia microbiana aconteceu com a utilização de genes frequentemente encontrados em todos os microrganismos, como o 16S rRNA em bactérias e o 18S rRNA em fungos. Com isso foi possível caracterizar filogeneticamente os microrganismos que compõem a comunidade microbiana local. Esta descoberta foi um avanço no campo da biorremediação, pois por meio de análises destas sequências de rRNA de microrganismos presentes em sítios contaminados foi possível determinar as espécies de microrganismos associados com o processo de biorremediação. Deste modo, vários microrganismos que nunca haviam sido relacionados com a degradação de hidrocarbonetos passaram a ser conhecidos e relacionados à biorremediação de áreas contaminadas (PACE *et al.*, 1986).

Uma limitação na caracterização de espécies bacterianas presentes em comunidades de microrganismos obtidas em locais de contaminação é o pequeno conhecimento a cerca de suas características fisiológicas (PACE, 1997). O avanço de métodos de sequenciamento de DNA e análise da expressão de genes aliados às técnicas de engenharia genética promoveu um aumento no entendimento do metabolismo dos microrganismos e sua aplicação na microbiologia ambiental. Técnicas, como mutagênese randômica, mutagênese sítio dirigido e *DNA shuffling*, têm sido amplamente empregadas para se obter variantes de enzimas existentes com capacidades catabólicas e biossensoras aperfeiçoadas. Além destas enzimas serem mais estáveis sob condições específicas, como altas temperatura e pressão e exposição a solventes, possibilitam a incorporação de novas rotas metabólicas que permitem ao microrganismo utilizar uma maior diversidade de substratos como fonte de carbono (SAMANTA, SINGH & JAIN, 2002).

1.3 - Breve histórico sobre a era genômica

A era genômica começou no final dos anos 80, catalisada pelo avanço na tecnologia de sequenciamento de DNA. Estudos iniciais foram constituídos principalmente de projetos em grande escala para o sequenciamento completo do genoma de certos organismos. Posteriores avanços nas técnicas de Biologia Molecular e o surgimento da bioinformática permitiram a aceleração na descrição de genes através do sequenciamento em massa e a possibilidade de análises computacionais comparativas entre genes de organismos relacionados, assegurando o grande impulso desta nova área de conhecimento dentro da Biologia (GIBAS & JAMBECK,

2001).

Bilhões de bases nucleotídicas sequenciadas encontram-se depositadas no GenBank, o maior e mais conhecido banco de seqüências da atualidade. Entre esta imensa massa de informação estão os genomas bacterianos, sendo os principais grupos as espécies Gram-negativas, com cerca de 1100 genomas de organismos do filo Proteobacteria, seguido da classe Bacilli com 457 e Actinobacteria com 217 espécies com genomas sequenciados (Figura 2). A maioria dos resultados de sequenciamento depositados no banco GenBank e em outros bancos de dados disponíveis foram obtidos por meio da técnica de Sanger, técnica que rendeu o segundo prêmio Nobel em 1980 a Frederick Sanger (SMITH *et al.*, 1986).

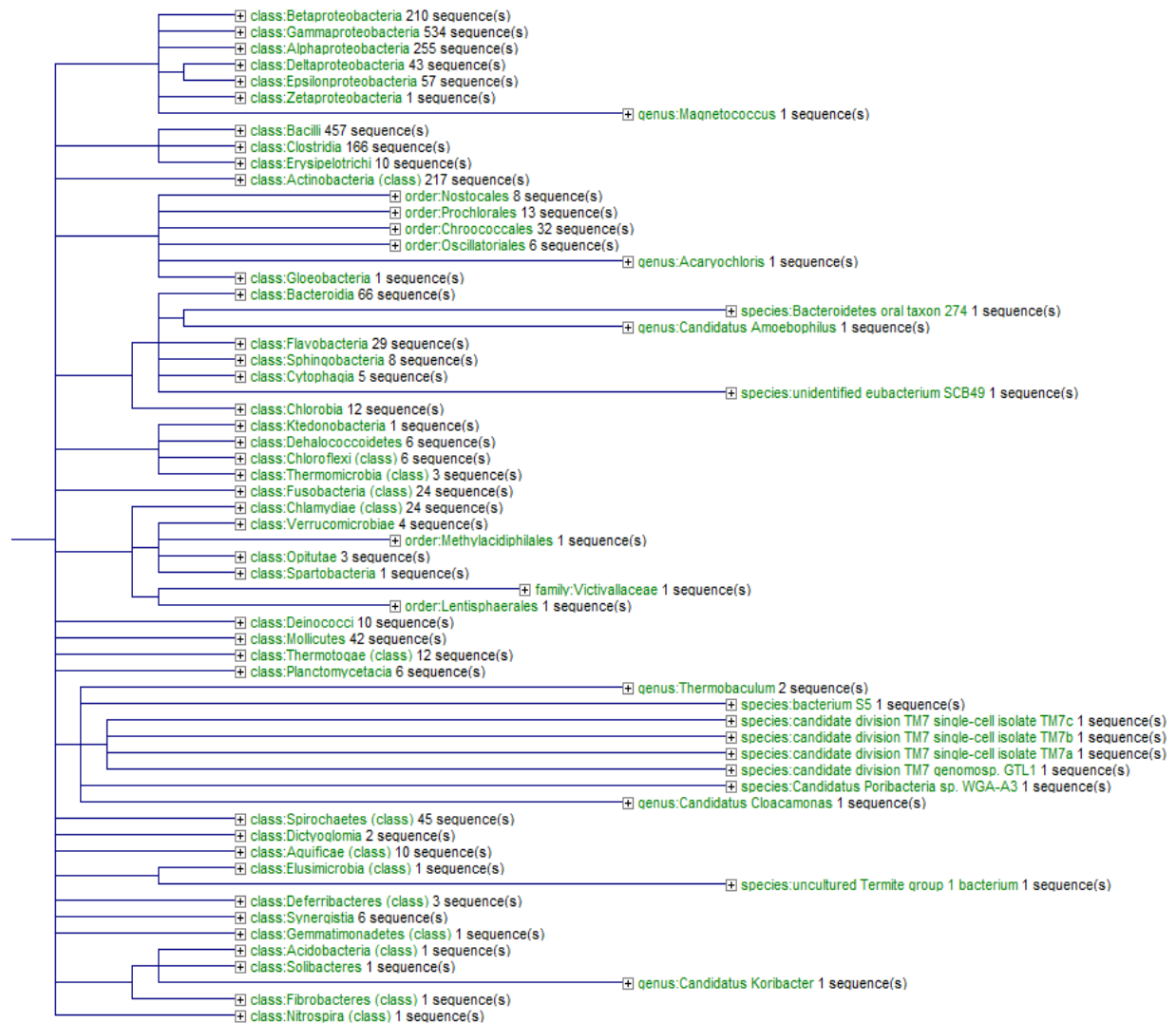


Figura 2: Números de espécies bacterianas com genomas sequenciados agrupados por classe, depositados no *National Center for Biotechnology Information* (NCBI).

Na técnica tradicional para o sequenciamento de DNA através do método de Sanger, quatro reações independentes são preparadas contendo, além dos nucleotídeos trifosfatados habituais, mais um dideoxynucleotídeo (ddNTP) sem o final 3'-OH. A incorporação de um ddNTP acarreta na terminação prematura da síntese da fita complementar de DNA. No final, em cada uma das quatro reações, haverá fragmentos de vários tamanhos diferentes. Posteriormente, estes fragmentos são analisados em eletroforese em gel de poliacrilamida, permitindo sua separação e conseqüentemente a determinação da seqüência.

O sequenciamento automático utiliza o método de Sanger combinado com o sistema de eletroforese capilar e varredura a laser. Neste novo sistema, os capilares são colocados em um suporte e submetidos à varredura do feixe de laser. Os fragmentos de DNA, previamente marcados com fluoróforos, emitem fluorescência que é filtrada e detectada por um fotomultiplicador. As fluorescências detectadas pelo fotomultiplicador são processadas por um computador e os dados armazenados na forma de eletroferogramas. Este sistema permite detectar até 10 mil pares de bases por hora (MATHIES, PECK & STRYER, 1990).

A técnica de sequenciamento *shotgun* utilizada durante o sequenciamento pioneiro do DNA humano pela equipe de Craig Venter da empresa Celera é um método que emprega enzimas para cortar o DNA em centenas ou milhares de pedaços aleatoriamente, os quais são clonados em vetores e em seguida utilizados no sequenciamento automático. Os aparelhos de sequenciamento automático decifram fragmentos de DNA com cerca de 500 a 850 bases. Apesar da técnica de Sanger apresentar um custo elevado e exigir uma infra-estrutura considerável, talvez o maior fator limitante desta técnica ainda seja a necessidade de montagem de uma biblioteca de clones contendo as seqüências de interesse (Figura 3).

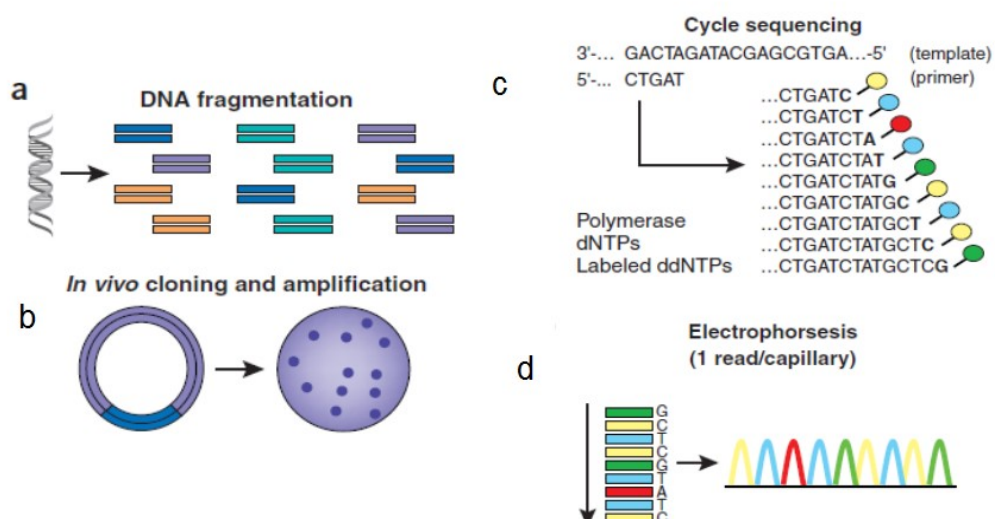


Figura 3: Esquema representando as principais etapas da técnica de sequenciamento automático por capilaridade, utilizando a metodologia de Sanger. (a) Fragmentação aleatória do DNA, (b) clonagem dos fragmentos de DNA em plasmídeos e transformação dos sistemas de ligação inserto:vetor em bactérias, (c) amplificação dos fragmentos clonados, empregando a técnica de terminação prematura do *amplicon* através de nucleotídeos marcados com fluoróforo, (d) detecção e análise dos sinais luminosos emitidos pelos nucleotídeos marcados com fluoróforo e armazenamento dos dados na forma de eletroferogramas.

Os resultados oriundos do sequenciador automático são apresentados na forma de gráficos de eletroferogramas e o processo de interpretação destes gráficos é chamado de *basecalling* (GIBAS & JAMBECK, 2001). Um programa muito utilizado para realizar o processo de *basecalling* é o PHRED, que utiliza a programação dinâmica para gerar uma saída de dados por meio do registro da base e a probabilidade de erro correspondente a essa base (EWING *et al.*, 1998). Estas probabilidades estão na forma de pontuações que representam a probabilidade logarítmica negativa da base, sendo que quanto maior a pontuação menor a probabilidade de erro. O índice varia entre 0 a 99, onde dez indica que a probabilidade desta base não corresponder ao registrado é de 10% (1/10), 20 indica que a probabilidade é de 1% de erro (1/100) e assim por diante (EWING *et al.*, 1998). Outra ferramenta largamente empregada em projetos de sequenciamento com a técnica de Sanger é o CAP3, um programa que realiza a montagem de uma sequência consenso chamada de *contig*, a partir de alinhamentos múltiplos entre um conjunto de sequências similares. O CAP3 é a terceira geração do programa de montagem de sequências CAP. O CAP3 apresenta a capacidade de remover as regiões 5' e 3' de baixa qualidade das sequências e utilizar os valores atribuídos às bases nucleotídicas pelo programa PHRED para computar as sobreposições dos consensos (HUANG & MADAN, 1999).

Vários conceitos são importantes na descrição das comparações de sequências de nucleotídeos. A **identidade de sequência** refere-se à presença da mesma base ou aminoácido na mesma posição em duas sequências alinhadas. O termo **similaridade** é frequentemente empregado nas comparações entre sequências de proteínas, quando dois aminoácidos são considerados similares com base em suas estruturas e características físico-químicas, sendo que geralmente a substituição de um pelo outro não afeta significativamente a estrutura ou a função da proteína. O conceito de **homologia** indica a relação evolutiva entre as sequências, quando apresentarem uma probabilidade de compartilharem o mesmo ancestral comum. Os genes ditos como homólogos podem ser classificados como **parálogos** e **ortólogos**. Os genes

parálogos são resultado de uma duplicação gênica a partir de um ancestral comum, podendo ter funções diferentes. Os genes ortólogos representam genes derivados de um ancestral comum que divergiram ao longo da evolução dos organismos, porém podem ainda apresentar funções similares (BAXEVANIS & OUELLETTE, 2001).

Uma vez dispondo-se da sequência de DNA, outro problema é a sua análise comparativa com seqüências de outros organismos relacionados ou não. A construção de algoritmos de busca em bancos de dados de seqüências é um dos grandes desafios na área da bioinformática. Estes algoritmos para busca precisam apresentar as seguintes características gerais: **1) velocidade**: devido ao tamanho dos bancos de dados atuais, os programas têm que ser ágeis para processar os dados rapidamente; **2) sensibilidade**: o programa precisa informar as similaridades significantivas; **3) estatísticas rigorosas**: o programa deve promover um modo para avaliar a significância dos resultados; e **4) facilidade de uso**: uma interface *Web* que possibilite que um único banco de dados público de seqüências de genoma ofereça todos os serviços por meio de uma interface uniforme e amigável aos usuários. Ferramentas de busca entre seqüências nucleotídicas ou proteicas por alinhamento como FASTA e BLAST (*Basic Local Alignment Search Tool*) (ALTSCHUL *et al.*, 1990) facilitaram e popularizaram as análises comparativas. O programa para consulta a bancos de seqüências FASTA, pertencente à família FAST, foi projetado inicialmente para pesquisar apenas seqüências de proteínas, sendo posteriormente incorporadas ferramentas de buscas de similaridades entre seqüências de DNA e comparações entre DNA e aminoácidos, quando necessário (PEARSON, 1988). O BLAST é uma família de ferramentas disponíveis na página <http://www.ncbi.nlm.nih.gov/BLAST/> para análise de similaridade entre seqüências. As buscas de similaridades entre uma seqüência a ser testada com as seqüências depositadas nos bancos de dados são feitas a partir da identificação de fragmentos entre as duas seqüências (ALTSCHUL *et al.*, 1990).

1.4 - A terceira geração de sequenciadores

Na última década, cinco novos sistemas de sequenciamento têm despontado no cenário do sequenciamento de DNA: a plataforma de pirosequenciamento 454 GS20/Roche, a plataforma AB SOLiD, os sistemas Solexa GA, Heliscope e recentemente a técnica desenvolvida pela Pacific Biosciences. Estas novas técnicas proporcionam uma imensa massa de informações em alta velocidade e com custos menores, quando comparadas com o sequenciamento automático de Sanger. Outra vantagem é a possibilidade de determinar as seqüências a partir de uma única fita de DNA, o que elimina a etapa de clonagem dos

fragmentos de DNA. Contudo, as limitações decorrentes do custo elevado do sequenciamento por base (mesmo sendo menor quando comparado com o método de Sanger), o alto índice de erros durante o sequenciamento e o pequeno tamanho das sequências (algumas técnicas não ultrapassam 40 bases) fazem do sequenciamento de Sanger um método ainda competitivo (SHENDURE & JI, 2008).

A plataforma 454 GS20 foi disponibilizada em 2005 pela Roche e foi substituída recentemente pela plataforma 454 GS FLX, sendo a mais utilizada atualmente (Figura 4a). O método empregado pela plataforma 454 GS FLX utiliza fragmentos de DNA, que podem ser de várias origens inclusive da fragmentação do DNA genômico, os quais são ligados a adaptadores e em seguida aderidos a um substrato esférico chamado de *beads*. Cada *bead* com a fita de DNA aderida é então mergulhada em uma emulsão de óleo onde os fragmentos de DNA são amplificados por meio de PCR com a utilização de uma DNA polimerase, produzindo várias cópias da mesma sequência de DNA. Após o término da PCR e a desnaturação das fitas de DNA sintetizadas é realizado o pirosequenciamento através da síntese da fita complementar. Para isso, os quatro desoxirribonucleotídeo trifosfatados (dNTP) diferentes são adicionados sequencialmente. Se o nucleotídeo correto for adicionado à reação este será utilizado na síntese da fita complementar, o que acarreta na liberação de pirofosfato (PPi) o qual é convertido à adenosina trifosfato (ATP) pela enzima ATP sulfúrilase. A enzima Luciferase utiliza esta molécula de ATP para produzir fóton de luz. Os fótons de luz produzidos proporcionalmente ao ATP disponível são detectados e analisados pelo sistema. Caso o nucleotídeo adicionado à reação não seja incorporado à fita crescente de DNA, não haverá a reação descrita anteriormente e conseqüentemente não ocorrerá nem a emissão de luz nem o registro da base.

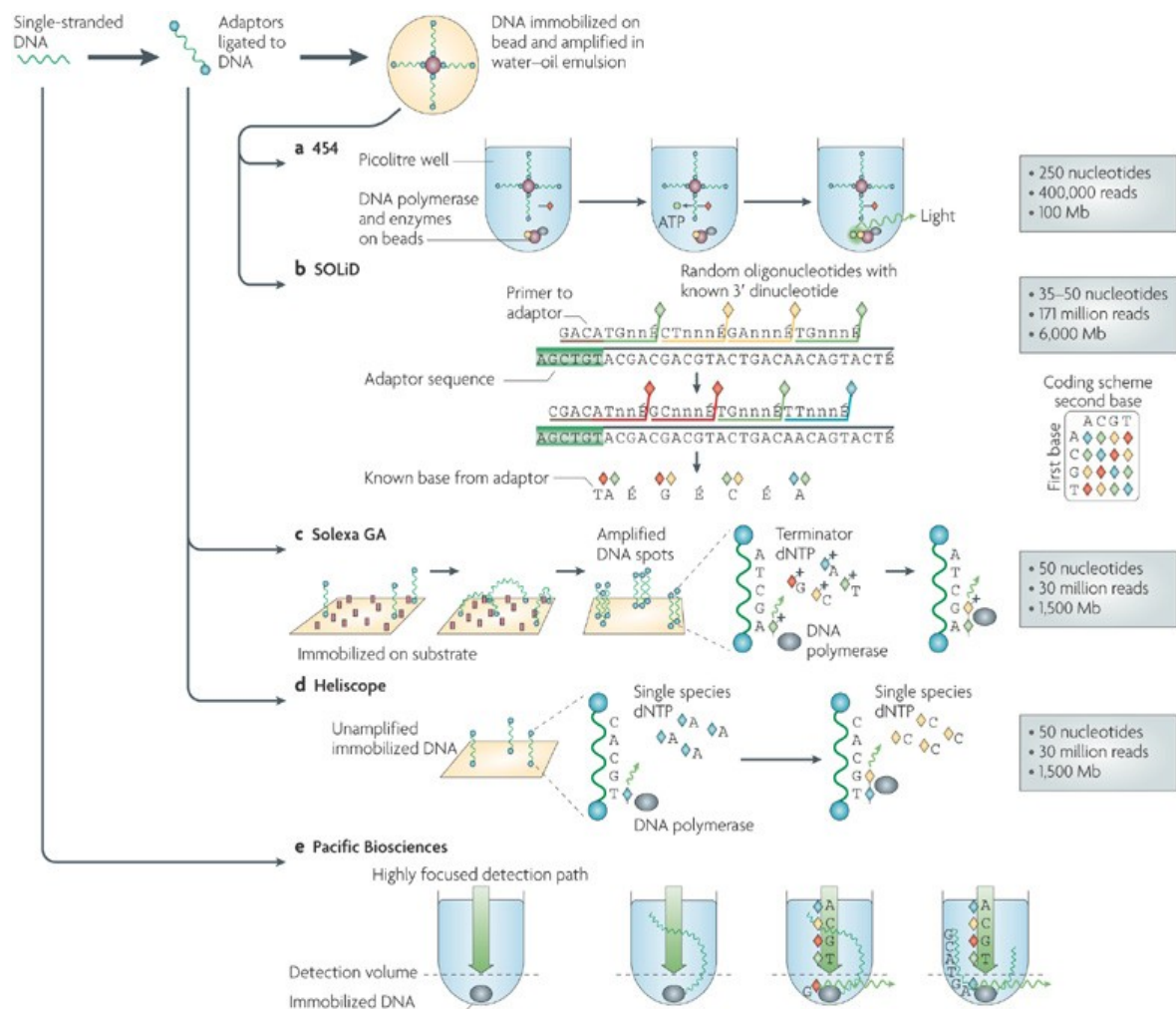


Figura 4: Esquema representativo das principais técnicas de sequenciamento de terceira geração. (a) 454 GS FLX Roche e (b) ABI SOLiD utilizam *beads* como substrato para o DNA molde, (c) Illumina Solexa emprega a ligação de adaptadores nas extremidades dos fragmentos de DNA, os quais são ligados por complementariedade a outros adaptadores previamente ligados a um substrato plano, (d) a técnica da HeliScope incorpora uma cauda poli(A) que hibridiza a uma cauda poli(T) previamente ligada a um substrato plano e (e) Pacific Biosciences, onde uma fita simples de DNA polimerase é fixada no fundo de um poço e a síntese da fita complementar permite o sequenciamento em tempo real (<http://www.nature.com/nrmicro/journal/v7/n4/full/nrmicro2088.html>).

A técnica disponibilizada pela plataforma ABI SOLiD (Figura 4b), descrita pela primeira vez por Shendure e colaboradores (2005), também emprega as *beads* como suporte para os fragmentos de DNA a serem amplificados. Nesta técnica, o DNA ligado às *beads* está

sujeito a hibridizações de iniciadores com sequências degeneradas, porém os dois últimos nucleotídeos da extremidade 3'-OH são conhecidos e contém o fluoróforo correspondente na última base. Outra diferença está no emprego da DNA ligase na síntese da fita complementar, ao contrário da DNA polimerase empregada na técnica 454 GS FLX Roche. Durante os ciclos de amplificação do DNA, os quais ocorrem em um substrato plano de vidro, a incorporação de um nucleotídeo na extremidade 3'-OH modificada do iniciador emite um sinal correspondente à base incorporada que é captado por quatro canais de imagens, uma para cada nucleotídeo. Normalmente são geradas *reads* com cerca de 36 nucleotídeos, embora já tenham sido detectadas *reads* com 70 nucleotídeos.

A plataforma Solexa foi inicialmente comercializada em 2006, sendo posteriormente adquirida pela empresa Illumina em 2007 (ANSORGE, 2009) (Figura 4c). O princípio do método Solexa reside no sequenciamento por síntese química, empregando nucleotídeos terminadores reversíveis para cada uma das quatro bases marcadas com um fluoróforo correspondente, além de uma DNA polimerase especial capaz de incorporá-los. Os fragmentos de DNA a serem sequenciados são previamente ligados a adaptadores em ambas as extremidades 5' e 3'. Posteriormente, os mesmos são ligados a um suporte plano densamente preenchido com adaptadores complementares, por meio de hibridização. Assim, a dupla fita de DNA liga as suas duas extremidades ao suporte através de dois adaptadores, assumindo uma forma de ponte. Após a desnaturação do fragmento de DNA em fita simples, os adaptadores atuarão como iniciadores para a síntese de uma nova fita complementar. No final da síntese, a nova dupla fita também assume a forma de ponte, devido à ligação de ambas as extremidades no suporte sólido. Estes processos de enriquecimento, por meio da síntese dos fragmentos de DNA, são necessários para se obter uma intensidade de sinais de luzes suficiente para ser detectada. Após várias etapas de ampliações, uma reação de sequenciamento é preparada contendo os quatro nucleotídeos modificados e a DNA polimerase especial. Durante a síntese de uma nova fita complementar a incorporação de um nucleotídeo modificado e marcado com o fluoróforo emitirá a luz correspondente à base, que será detectada e analisada por um sistema computacional.

O sequenciamento por meio da técnica HeliScope (BRASLAVSKY *et al.*, 2003) difere das outras plataformas descritas anteriormente por não necessitar da etapa de enriquecimento clonal através de PCR (Figura 4d). Um sistema de detecção altamente sensível é empregado diretamente sobre uma única molécula de DNA de interesse, durante a síntese da fita complementar. Outra diferença reside no preparo da fita de DNA a ser sequenciada. Primeiramente, o DNA é fragmentado aleatoriamente e em seguida uma cauda poli-A é

incorporada a estes fragmentos, sem a necessidade de qualquer método de PCR. Os fragmentos de DNA/poli-A são então capturados por hibridização e fixados em um suporte plano, previamente preparado com oligômeros poli-T. São realizados então ciclos de síntese da fita complementar empregando uma mistura contendo nucleotídeos marcados com fluoróforos correspondentes a cada base nucleotídica e um terminador, que é adicionado para impedir a continuação da síntese da fita complementar após a incorporação de um nucleotídeo marcado. Os sinais emitidos após a incorporação de um nucleotídeo são detectados e armazenados por um sistema computacional.

O método disponibilizado pela Pacific Biosciences é o mais revolucionário entre as cinco tecnologias descritas acima (Figura 4e). Nesta técnica, uma única molécula de DNA polimerase é imobilizada no fundo de um poço com altura e diâmetro da ordem de nanômetros, o que permite uma alta concentração de substratos e alta velocidade de síntese. Na reação, são adicionados nucleotídeos marcados com fluoróforos e uma fita simples de DNA que servirá de molde para a síntese da fita complementar. Durante a reação da síntese, os nucleotídeos são adicionados à fita complementar, o que acarreta na clivagem do fluoróforo e a emissão de luz correspondente à base nucleotídica. O sinal é captado continuamente durante a extensão da fita complementar, o que faz desta tecnologia uma das mais rápidas disponíveis no mercado.

Cada uma das cinco tecnologias de sequenciamento aqui descritas possui suas vantagens e limitações. Contudo uma limitação comum a todas as tecnologias residem ao pequeno tamanho das sequências geradas, chamadas de *reads*, normalmente entre 30 a 100 nucleotídeos. O tamanho limitado destas sequências acarreta um problema durante a montagem das *reads*. Cada vez mais há a necessidade de novos programas capazes de realizar a montagem de um número imenso de *reads* em sequências maiores, chamadas de *contigs*. Outro problema acontece durante a fase de detecção, onde o laser provoca efeitos deletérios sobre as enzimas e os fragmentos de DNA.

Com o advento das novas tecnologias de sequenciamento, uma grande variedade de programas foi desenvolvida para a análise dos dados provenientes das diferentes plataformas (454 Roche, SOLiD, Solexa, Heliscope e Pacific Biosciences). As funções destas ferramentas de bioinformática incluem alinhamento de sequências *reads*, *base-calling*, detecção de polimorfismo, anotação de genomas e associação *de novo* de *reads* não alinhadas. Isso porque os algoritmos dos programas de alinhamento BLAST disponíveis foram desenvolvidos para serem utilizados em sequências de maior extensão, sendo improvável sua aplicação em alinhamentos de sequências com tamanhos de 50 a 100 nucleotídeos gerados pelas

plataformas de sequenciamento de terceira geração. Atualmente, já existem programas que utilizam algoritmos bem estabelecidos, como o Smith-Waterman, para alinhamento de sequências de pequenos tamanhos, como por exemplo, os programas SOAP, SHRiMP e MAQ (LI *et al.*, 2008; ANSORGE, 2009).

1.5 - A microbiologia na era genômica

Com o DNA genômico da bactéria *Haemophilus influenzae* sequenciado e disponibilizado ao público (FLEISCHMANN *et al.*, 1995), acreditava-se que mais algumas dezenas de organismos modelos representativos dos domínios Bacteria e Archaea seriam suficientes para descrever todo o *pool* de genes presentes em procariotos. Após 15 anos e milhares de genomas totalmente sequenciados e disponibilizados já se reconhece que a diversidade microbiana foi imensamente subestimada.

Uma boa referência para descrever a variabilidade genômica é o conceito de pan-genoma. Medini e colaboradores (2008) sugerem dividir o conceito de pan-genoma em três elementos: os genes compartilhados por todas as estirpes da espécie, chamados de *core genome*, que estariam envolvidos na manutenção básica da célula; um conjunto de genes dispensáveis que seriam compartilhados entre apenas algumas espécies e um conjunto de genes específicos apenas para uma determinada estirpe, os quais provavelmente confeririam alguma vantagem competitiva na sobrevivência em um determinado ambiente. Desta forma, o conceito de pan-genoma reflete a pressão seletiva do ambiente sobre as espécies, produzindo novas recombinações de genes. Vários processos naturais de transferência de informações gênicas, como a transdução, a conjugação e a transformação, podem estar envolvidos na variabilidade intra-espécie (OCHMAN, LAWRENCE & GROISMAN, 2000).

A variabilidade genética é facilmente detectada mesmo dentro de uma única população, aumenta exponencialmente em uma única espécie bacteriana (pan-genoma) e aumenta ainda mais em um grupo de microrganismos de um mesmo nicho ambiental (metagenoma). Esta plasticidade genômica é visível em muitas espécies bacterianas. Um exemplo pode ser dado por um estudo comparativo dos genomas das estirpes O157:H7 e K12 de *Escherichia coli* que revelou que cerca de 1300 genes são específicos apenas para a primeira estirpe (PERNA *et al.*, 2001). Quando uma terceira estirpe, a CFT073, foi incluída no estudo, os resultados obtidos mostraram que apenas 39,2% dos genes são compartilhados pelas três estirpes (WELCH *et al.*, 2002).

Estes estudos mostram a dificuldade de caracterizar uma espécie a partir de um único genoma e levanta a questão do número de genomas que precisam ser sequenciados para

realmente alcançar o conhecimento sobre determinada espécie. Em outro estudo sobre a diversidade genômica de oito estirpes de *Solibacter usitatus*, a qual possui um genoma com tamanho médio de 10 Mb, mostrou que cerca de 30 genes não foram detectados nos genomas de outras estirpes de *S. usitatus* previamente sequenciados (TETTELIN *et al.*, 2005).

Na natureza, os microrganismos ocupam ambientes diversos em uma estrutura complexa de comunidade, com inter-relações entre os mesmos por meio de transferência horizontal de genes. Isto enfatiza a importância do ambiente sobre a evolução genômica de um indivíduo membro de uma comunidade. Assim, uma abordagem chamada de metagenômica, emprega métodos de sequenciamento e outras técnicas para estudar diretamente os microrganismos em seu ambiente natural (MEDINI *et al.*, 2008), visualizando a comunidade local como uma única unidade de estudo, sem a necessidade de isolamento do microrganismo. Este método tem identificado milhares de novas espécies de bactérias e milhões de novos genes, corroborando ainda mais a imensa diversidade de genes a serem explorados. O emprego das novas tecnologias no sequenciamento de microrganismos tem permitido a aquisição de giga bases de informações em um período relativamente curto de tempo. Estas novas estratégias são amplamente empregadas em metagenômica, pan-genoma, epigenética, descobertas de RNA não codantes e sítios de ligações de proteína no genoma.

Diversos projetos de sequenciamento de genoma bacterianos têm utilizado alguma das novas tecnologias descritas acima em conjunto com a técnica de Sanger. Isso se deve ao fato que as metodologias disponíveis pelas plataformas de terceira geração normalmente não conseguem preencher todas as lacunas de um genoma. Um exemplo pode ser dado pelo sequenciamento do genoma da bactéria de solo *Myxococcus xanthus* (VELICER *et al.*, 2006). O sequenciamento realizado pela plataforma 454 GS20 gerou um total de 2,5 milhões de *reads*, com um tamanho médio de 100 nucleotídeos. Após a montagem destas *reads* por meio de programas computacionais foram geradas 104 *contigs*, deixando 104 lacunas para serem preenchidas. Destas, 88 lacunas foram preenchidas utilizando o método de Sanger. Porém, uma parcela destes projetos emprega apenas as tecnologias de terceira geração, não dando importância ao fechamento das lacunas. Esses são então denominados de *draft genomes*. Para determinadas aplicações e objetivos, um sequenciamento do tipo *draft genome* é suficiente para trazer informações relevantes a respeito da genética do microrganismo.

1.6 - A contribuição da técnica de PCR em tempo real no estudo da biorremediação

A técnica de PCR em tempo real ou PCR quantitativo (qPCR) é uma ferramenta poderosa e versátil empregada na análise de genes de interesse em uma determinada condição

biológica. A metodologia de qPCR permite detectar e quantificar, a cada ciclo, o acúmulo de produtos de amplificações de sequências de DNA através de uma molécula repórter fluorescente, como a *Taqman* ou corantes que se ligam à molécula de DNA, como o *SYBR Green I* (GINZINGER, 2002). Atualmente, estas duas metodologias, *Taqman* e *SYBR Green I*, são as mais empregadas. O sistema *Taqman* utiliza a atividade exonucleotídica 5'-3' da *Taq* polimerase para digerir sondas modificadas que se anelam em regiões internas de um gene de interesse. Estas sondas possuem em sua porção terminal 5' uma molécula fluorescente, chamada de *reporter*, enquanto na porção terminal 3' há outra molécula com atividade inibitória sobre o fluoróforo *reporter*, denominada de *quencher*. Quando o gene de interesse é amplificado, a sonda é degradada e o *reporter* (contendo o fluoróforo e o *quencher*) é liberado. Uma vez distantes um do outro, a inibição do *quencher* sobre o *reporter* é eliminada e o fluoróforo emite a luz que será detectada e quantificada. O sistema de *SYBR Green I* utiliza um corante que se liga ao sulco menor da molécula de DNA. A cada ciclo há um aumento na quantidade dos produtos e, conseqüentemente, há um aumento proporcional no sinal emitido pelo corante, que passa a ser detectado e quantificado como níveis de expressão do gene.

A utilização das metodologias de qPCR no campo da biorremediação de petróleo é de grande importância principalmente na detecção e na quantificação de genes envolvidos na degradação de hidrocarbonetos e outros xenobióticos, a partir de amostras ambientais contaminadas ou não (SALMINEN, TUOMI & JØRGENSEN, 2008; WASMUND *et al.*, 2009; SCHULZ *et al.*, 2010). Outra aplicação relevante é na determinação da diversidade microbiana local e sua associação com processos de biorremediação (POWELL *et al.*, 2006; SCHULZ *et al.*, 2010). Trabalhos envolvendo os nível de expressões de mRNA de genes catabólicos de hidrocarbonetos foram descritos para o regulador principal dos genes metabólicos da via de degradação de *n*-alcanos em *Pseudomonas putida* GoP1 (DINAMARCA *et al.*, 2003). A quantificação dos níveis transcricionais de genes envolvidos na degradação de hidrocarbonetos aromáticos em condições de laboratório também é descrita para os compostos de benzeno, etilbenzeno, tolueno e xileno (BTEX), os quais são considerados altamente danosos ao meio ambiente (Da SILVA & ALVAREZ, 2004; KÜHNER *et al.*, 2005; GUNSCH *et al.*, 2006).

1.7 - Degradação de alcanos

Os alcanos estão entre os principais compostos presentes no petróleo, chegando a representar até 50% de sua massa, mas podendo ser também produzidos por algas verdes,

bactérias, plantas e animais (ROJO, 2009). Estes são caracterizados por serem lineares (*n*-alcanos), ramificados (*iso*-alcanos) ou cíclicos (*ciclo*-alcanos), com tamanhos variando entre 1 e 50 carbonos. Desde o início do século passado, quando foi isolada pela primeira vez uma bactéria degradadora de metano, microrganismos tais como bactérias (a maioria pertencente às subdivisões *Alphaproteobacteria*, *Betaproteobacteria*, *Gammaproteobacteria* e ao grupo dos Actinomicetos) e fungos filamentosos e leveduras têm sido isolados de ambientes contaminados de petróleo e descritos como degradadores de hidrocarbonetos alifáticos e aromáticos (COLEMAN, BUI & HOLMES, 2006). Embora muitos microrganismos utilizem hidrocarbonetos como fonte de carbono e energia, a maioria prefere outros tipos de substratos para crescimento, ao contrário de certas espécies que utilizam hidrocarbonetos como fonte preferencial de carbono. Estes microrganismos são chamados de hidrocarbonoclasticos e tem como principal representante a bactéria *Alcanivorax borkumensis* (SCHNEIKER *et al.*, 2006).

Os *n*-alcanos são moléculas apolares e apresentam baixa reatividade química, especialmente nas posições terminais (LABINGER & BERCAW, 2002). Isso porque os átomos de carbono e hidrogênio são quase igualmente eletronegativos, tendo todos os seus elétrons compartilhados, assim resultando em baixa polaridade e, conseqüentemente, baixa reatividade. Embora os *n*-alcanos sejam quimicamente inertes, na presença de oxigênio muitos microrganismos podem ativá-los oxidando os grupos metil terminal ou subterminal, transferindo um átomo de oxigênio a partir do O₂ produzindo álcool, enquanto o outro átomo de oxigênio é reduzido a H₂O. Esta reação normalmente é realizada por um conjunto enzimático, conhecido como “sistema AlkB”, composto por três proteínas: uma alcano hidroxilase transmembrana (AlkB), uma ou duas rubredoxinas (AlkF e AlkG) e uma rubredoxina redutase (AlkT) (van BEILEN *et al.*, 2002a). O produto desta primeira reação é novamente oxidado por uma álcool desidrogenase (AlkJ) e novamente por uma aldeído desidrogenase (AlkH). Em seguida, são produzidos ácidos graxos pela enzima Acil-coenzima A (AlkK). Assim, os ácidos graxos são direcionados para a via de β -oxidação, culminando na produção de energia pela célula (Figura 5) (van BEILEN, WUBBOLTS & WITHOLT, 1994; van BEILEN *et al.*, 2004). A oxidação subterminal difere das reações descritas acima. O produto gerado pela oxidação subterminal, um álcool secundário, é convertido a cetona e oxidado novamente por uma monoxigenase do tipo Bayer-Villiger, produzindo éster. A molécula de éster é então hidrolisada por uma esterase, resultando em uma molécula de álcool e outra de ácido graxo que podem ser utilizadas como fonte de carbono pela célula. Ambas as vias de degradação de alcanos podem coexistir em alguns microrganismos (ROJO, 2009).

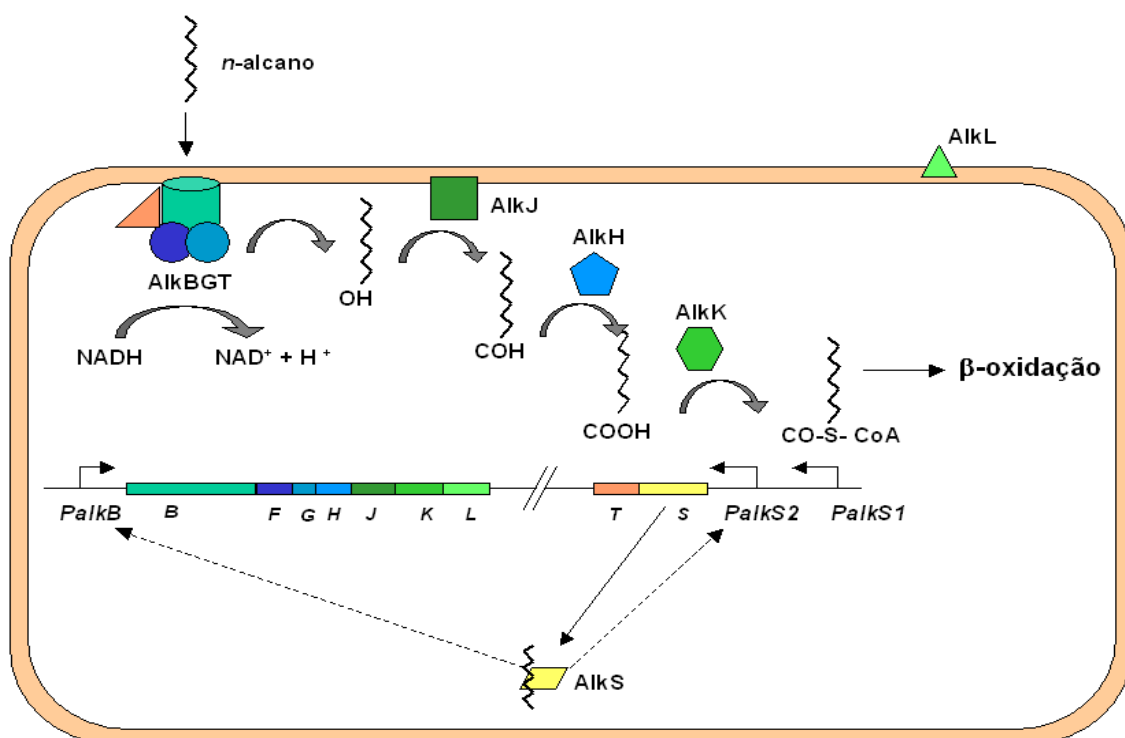


Figura 5: Via metabólica de Alk descrita em *Pseudomonas putida (oleovorans)* Gp01 (van BEILEN *et al.*, 2002b).

As enzimas envolvidas na assimilação de *n*-alcanos são codificadas por genes localizados tanto em cromossomos como em plasmídios bacterianos. Em *Pseudomonas putida (oleovorans)* GP01, onde a via metabólica Alk é melhor conhecida, os genes estão agrupados em dois *operons*: *alkBFGHJKL* e *alkST*, presentes no plasmídeo OCT. Como mostrado na Figura 5, na presença de *n*-alcanos a proteína reguladora AlkS induz a transcrição dos *operons* *alkBFGHJKL* e *alkST* a partir dos promotores *P_{alkB}* e *P_{alkS2}*, respectivamente. A indução do *operon* *alkST* irá resultar no acréscimo da transcrição de todos os genes da via Alk, por meio do aumento dos níveis da proteína AlkS (DOUGHTY *et al.*, 2006). Contudo, na ausência de *n*-alcanos ou na presença de outras fontes de carbono, como a glicose ou o succinato, ocorre uma diminuição na transcrição dos genes de toda a via Alk por meio de dois mecanismos: o *operon* *alkST* passa a ser expresso somente em níveis basais através do promotor *P_{alkS1}*, o qual é um promotor fraco, e há uma repressão nos níveis transcricionais dos genes da via Alk pela proteína Crc (*Catabolic repression control*) (Figura 6). A proteína Crc é um regulador global do metabolismo de *P. putida*, atuando na assimilação de hidrocarbonetos, de compostos aromáticos e na formação de biofilmes (DINAMARCA *et al.*, 2003; RUIZ-MANZANO, YUSTE & ROJO, 2005). Embora não se saiba corretamente o mecanismo de ação da Crc, dados na literatura sugerem que a Crc não atue por meio do

modelo clássico de ligação repressor-DNA (GORKE & STULKE, 2008). Dinamarca e colaboradores (2003), ao analisarem mutantes sem a repressão catabólica sobre o promotor do *operon* AlkS, encontraram mutantes para a enzima citocromo *o* ubiquinol oxidase (*cyoB*), importante componente da cadeia de transporte de elétrons em condições de aerobiose. Este fato indica que as condições de disponibilidade de oxigênio também determinam a escolha da fonte de carbono a ser utilizada pelo microrganismo.

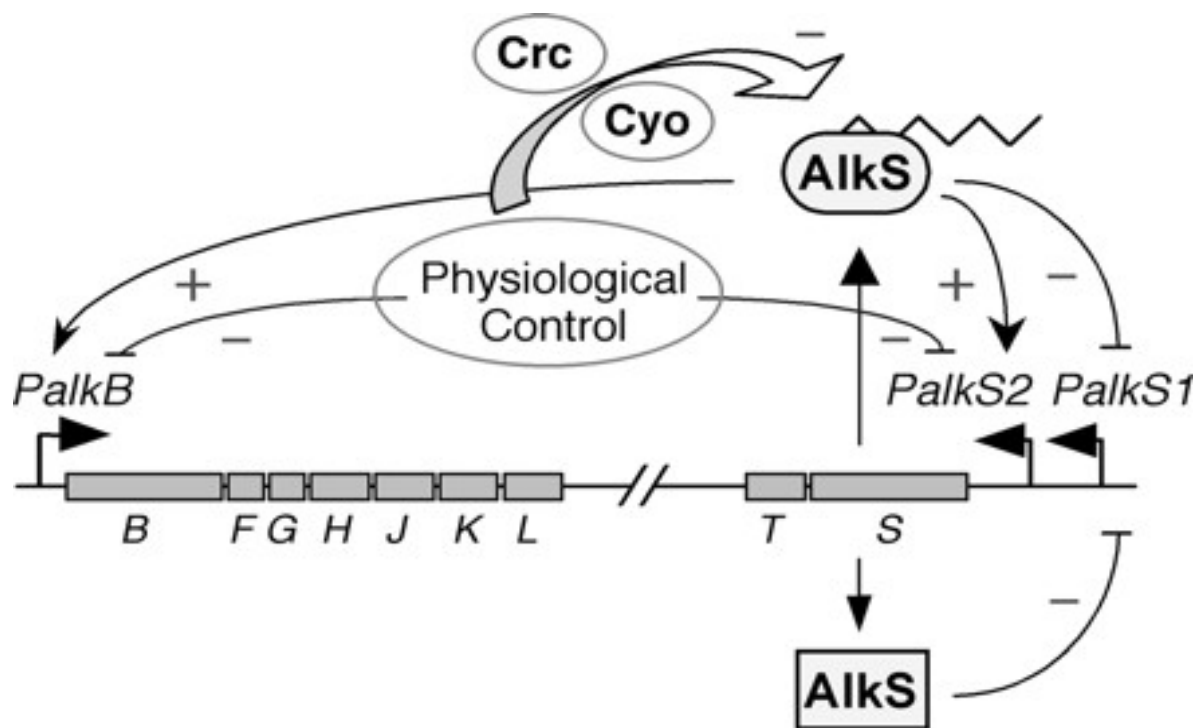


Figura 6: Controle sobre a expressão dos *operons* *alkBFGHJKL* e *alkST* em *Pseudomonas putida* GPo1 através das proteínas Crc e Cyo (RUIZ-MANZANO, YUSTE & ROJO, 2005).

A proteína transmembrana AlkL, também presente no *operon* Alk, não tem sua função conhecida, embora alguns dados na literatura sugiram que AlkL possa participar no transporte de *n*-alcanos pela bactéria. Entretanto, van Beilen e colaboradores (1992) mostraram que a deleção desta proteína não teve nenhum efeito no crescimento de *P. putida*. A impossibilidade de caracterizar uma proteína envolvida no transporte de moléculas de *n*-alcanos para dentro da célula levanta a questão de qual mecanismo é utilizado para a aquisição destas moléculas e sua utilização como fonte de carbono e energia. Os alcanos em geral são insolúveis na água, e esta característica de insolubilidade é maior conforme a sua massa molecular aumenta. O mecanismo pode diferir entre as diferentes espécies de bactérias e dependendo do peso

molecular da molécula e das características físico-químicas do ambiente onde se encontram as moléculas de alcanos (WENTZEL *et al.*, 2007). Moléculas de surfactantes foram descritas como facilitadores para a transferência de alcanos para o interior da célula, quando crescida em meio líquido, mas estes foram incapazes de realizar a transferência em meio sólido (NOORDMAN & JANSSEN, 2002). Desta forma, a caracterização de uma proteína envolvida no transporte de moléculas de alcanos para o interior da célula ainda não está totalmente elucidada.

Além da monoxigenase AlkB envolvida na degradação de compostos *n*-alcanos, as enzimas monoxigenases dependente de heme P450 também são descritas como capazes de realizar a primeira reação de hidroxilação destes compostos. As monoxigenases P450, também conhecidas como CYP, compõem uma grande classe de proteínas com cerca de 100 famílias diferentes de enzimas, sendo amplamente distribuídas desde entre os procariotos até nos eucariotos superiores como mamíferos e plantas. Contudo, as P450 são mais encontradas em fungos capazes de degradar hidrocarbonetos (ROJO, 2009). A primeira P450 relatada como responsável pela hidroxilação de compostos *n*-alcanos foi a CYP153A1 presente na bactéria *Acinetobacter* sp. EB104 (MAIER *et al.*, 2001). A partir deste trabalho, várias P450 envolvidas na hidroxilação de *n*-alcanos foram descritas em espécies dos gêneros *Mycobacterium*, *Rhodococcus* e do grupo das proteobactérias (SEKINE *et al.*, 2008). Ao contrário das monoxigenases AlkB, as P450 são proteínas solúveis, o que faz delas um cobiçado alvo de interesse da indústria de biotransformação (FUNHOFF *et al.*, 2006).

1.8 - O sistema AlkB

As oxigenases pertencem a uma ampla classe de enzimas relacionadas, que possuem a característica de catalisar a inserção seletiva de átomos de oxigênio em diversos compostos orgânicos. Estas enzimas possuem um papel importante na destoxificação em células hepáticas de mamíferos e na biossíntese de metabólitos secundários, moléculas de sinalização e hormônios. Porém, na degradação de hidrocarbonetos e na indústria de produtos químicos finos as oxigenases têm apresentado enorme importância na área de biocatálise, como por exemplo, na produção de compostos tais como pravastatina, índigo, hidroxiprolina e drogas humanas (van BEILEN *et al.*, 2003a). Em *Mycobacterium* sp. HXN-1500 os genes que codificam a monoxigenase P450, uma ferredoxina e ferredoxina redutase, foram capazes de converter limonene em álcool perilil, uma droga carcinogênica de interesse industrial (van BEILEN *et al.*, 2005).

O sistema AlkB é um complexo enzimático de três proteínas: uma alcano hidroxilase,

uma ou duas rubredoxinas e uma rubredoxina redutase. Comparações feitas entre as seqüências de alcano hidroxilases *alkB* de diversas bactérias Gram-negativa e Gram-positiva mostraram uma grande divergência entre elas, como observado especialmente nas estirpes dos gêneros *Pseudomonas* e *Rhodococcus* (SMITS *et al.*, 2002; WHYTE *et al.*, 2002). As distâncias das árvores filogenéticas baseadas em seqüências de aminoácidos apresentaram algumas vezes apenas 40% de identidade (van BEILEN *et al.*, 2003a). Apesar das seqüências do gene *alkB* apresentarem divergências significativas entre si, todas as monoxigenases AlkB possuem seis domínios α -hélice hidrofóbicos transmembrana, quatro sítios conservados ricos em resíduos de histidina essenciais para a atividade da enzima e provavelmente um centro dinuclear de ferro sem o anel heme (SHANKLIN & WHITTLE, 2003) (Figura 7).

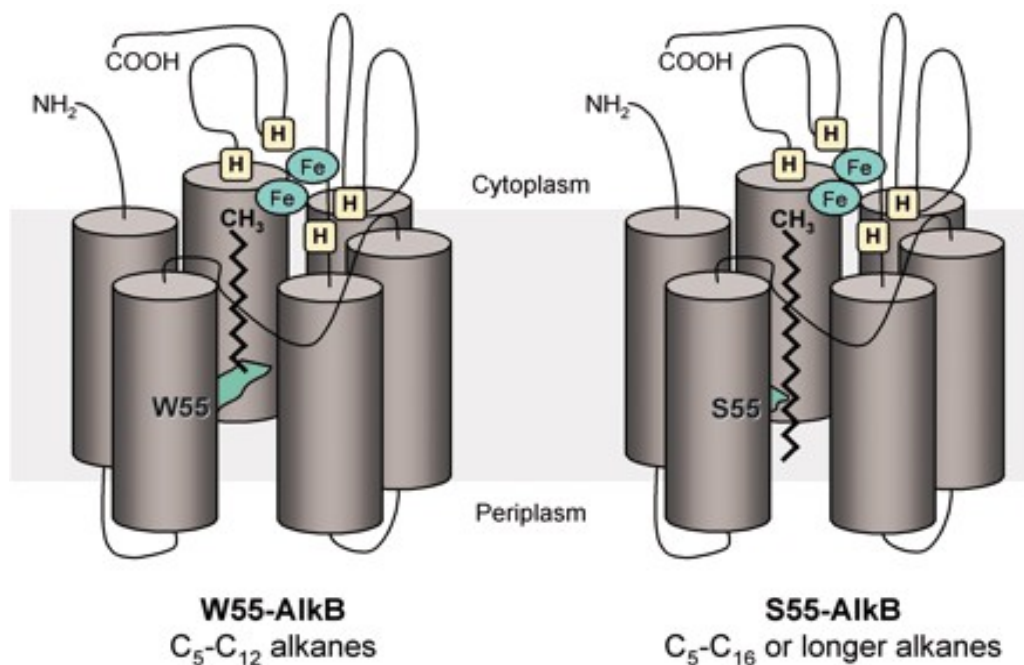


Figura 7: Modelo proposto para a proteína transmembrana AlkB em *P. putida* GPo1 (ROJO, 2005).

As proteínas rubredoxina (AlkG) e rubredoxina redutase (AlkT) são essenciais na atividade de hidroxilação de *n*-alcanos pelo sistema AlkB. Estudos bioquímicos em *P. putida* mostram que a rubredoxina redutase transfere elétrons de NAD(P)H via cofator FAD, em uma reação redox, para a rubredoxina, a qual então transfere para a proteína AlkB produzir álcool (van BEILEN *et al.*, 2002a). A enzima rubredoxina é uma proteína pequena com um sítio ativo ferro-enzofre. Estas enzimas podem ser divididas em dois grupos: as rubredoxinas tipo AlkG-1, que não participam na transferência de elétrons para a enzima AlkB e as

rubredoxinas tipo AlkG-2, responsáveis por esta reação redox. As enzimas tipo AlkG-1 estão presentes em organismos não-degradadores de hidrocarbonetos e provavelmente atuando em respostas a estresses oxidativos em bactérias anaeróbias. Em *P. putida* Gp01, a pré-proteína AlkG contém dois domínios rubredoxina ligados por uma pequena sequência peptídica, os quais são clivados posteriormente por ação de proteases. Na espécie *Rhodococcus erythropolis*, os genes que codificam as proteínas AlkG estão presentes em várias cópias logo após o gene *alkB*, seguido pelo gene *alkT*. Já em outras bactérias o gene *alkT* está localizado em outro locus, provavelmente porque esta proteína participa de outras vias metabólicas (WHYTE *et al.*, 2002). van Beilen e colaboradores (2002a) também relatam a distribuição dos genes acessórios *alkG* e *alkT* por todo o genoma bacteriano ao invés da clássica organização encontrada no operon OCT de *P. putida* Gp01.

1.9 - Os genes do sistema Alk no ambiente

Técnicas de PCR utilizando iniciadores degenerados para o gene *alkB* permitiram a identificação de várias estirpes bacterianas, Gram-negativo e Gram-positivo, capazes de assimilar *n*-alcanos como única fonte de carbono (SMITS *et al.*, 2002; van BEILEN *et al.*, 2002b). As espécies *P. aeruginosa*, *P. fluorescens* e *P. putida* contém várias estirpes bem caracterizadas em relação à presença dos genes da via Alk (van BEILEN & WITHOLT, 2004). Enquanto que em *P. putida* Gp01 apenas *n*-alcanos de tamanhos entre C₅ e C₁₂ são oxidados, muitas outras bactérias degradadoras de *n*-alcanos são capazes de oxidar cadeias maiores que C₁₀ (SMITS *et al.*, 2002). Por exemplo, a estirpe *P. aeruginosa* PAO1 apresenta dois genes para a enzima AlkB, *alkB1* e *alkB2*, e dois genes para a rubredoxina AlkG, *alkG1* e *alkG2*, os quais são expressos de modo diferente na presença de *n*-alcanos de tamanhos variados: AlkB1 oxida C₁₆-C₂₀ enquanto AlkB2 C₁₂-C₂₀ (MARIN, YUSTE & ROJO, 2003). van Beilen e colaboradores (2005) identificaram (através de técnicas de mutação e engenharia de proteínas) como a proteína AlkB seleciona o tamanho das cadeias de *n*-alcanos para serem hidroxiladas. A proteína selvagem AlkB ao ter o resíduo de aminoácido triptofano na posição 55 trocado pelo resíduo de serina, um aminoácido de menor volume, permite a entrada de cadeias de *n*-alcanos de tamanho maior (Figura 7). Já a estirpe *Acinetobacter* sp. M-1, onde os genes responsáveis pela assimilação de *n*-alcanos são uma alcano hidroxilase (*alkM*), uma rubredoxina (*rubA*) e uma rubredoxina redutase (*rubB*), localizados em diferentes posições no cromossomo, é caracterizada pela capacidade de degradar *n*-alcanos de diversos tamanhos, incluindo os *n*-alcanos de cadeia longa (C₁₃-C₄₄) (ROJO, 2009).

Estirpes pertencentes ao gênero *Rhodococcus* são capazes de degradar *n*-alcanos de

tamanhos variados, além de componentes aromáticos heterocíclicos e policíclicos. Esta versatilidade biodegradativa se deve à presença de genes que codificam vários tipos de mono e dioxigenases (van der GEIZE & DIJKHUIZEN, 2004). Whyte e colaboradores (1999) relataram que *Rhodococcus* sp. foi capaz de assimilar *n*-alcanos de tamanhos variados em diferentes condições de temperatura. Este mesmo grupo descreveu em diferentes estirpes de *Rhodococcus*, isoladas de ambientes distintos, pelo menos quatro genes homólogos a monoxigenases, *alkB1*, *alkB2*, *alkB3* e *alkB4* (WHYTE *et al.*, 1999). A presença de múltiplos genes homólogos a alceno hidroxilases em isolados de *Rhodococcus* evidencia a importância desta via na biodegradação destes compostos. Em *Nocardioides* sp. CF8 foram detectadas duas monoxigenases distintas capazes de oxidar *n*-alcanos: uma monoxigenase contendo cobre e uma monoxigenase transmembrana (HAMAMURA, YEAGER & ARP, 2001).

A detecção de genes envolvidos na degradação de xenobióticos em comunidades bacterianas em ambientes contaminados por petróleo tem confirmado a importância dos gêneros *Pseudomonas*, *Rhodococcus*, *Mycobacterium* e *Acinetobacter* como degradadores de *n*-alcanos (LUZ *et al.*, 2004; van BEILEN *et al.*, 2003a). Análises de solos contaminados da Antártida e do Brasil, quando comparados com solos não contaminados de localidades próximas, mostraram a presença de genes *alkB1* e *alkB2* de *Rhodococcus* sp. em ambas as localizações, mais a presença de genes *alkB* de *Pseudomonas* sp. nos solos da Antártida (LUZ *et al.*, 2004). Embora os gêneros Gram-positivos *Rhodococcus*, *Mycobacterium*, *Gordonia* e *Nocardioides* sejam reconhecidos como degradadores de hidrocarbonetos, muito pouco é conhecido sobre a fisiologia e sobre os genes responsáveis por esta degradação (BEARD & PAGE, 1998).

1.10 - O gênero *Dietzia*

A assimilação de hidrocarbonetos por bactérias do gênero *Dietzia* já foi relatada em diversos estudos de comunidades bacterianas em sítios contaminados por hidrocarbonetos, ao lado de outras Gram-positivo (CHAILLAN *et al.*, 2004; KLEINSTEUBER *et al.*, 2006). Bactérias pertencentes ao gênero *Dietzia* são amplamente distribuídas no ambiente, tendo sido isoladas de solos, rizosferas, lagos alcalinos, trato intestinal e pele de peixes marinhos (YASSIN, HUPFER & SCHAAL, 2006). Estirpes pertencentes às espécies *D. maris* (RAINEY *et al.*, 1995; BRITO *et al.*, 2006) e *D. psychrocaliphilia* (YUMOTO *et al.*, 2002) já foram isoladas e descritas como capazes de utilizar diversos *n*-alcanos como única fonte de carbono.

A espécie *D. cinnamea* é reconhecidamente patogênica ao ser humano, tendo sido

descrita a partir de um isolado obtido da medula óssea (YASSIN, HUPFER & SCHAAL, 2006). Entretanto, em um estudo sobre o efeito da bioestimulação na diversidade bacteriana em solos contaminados por petróleo em uma região da Mata Atlântica, Brasil, foi isolada uma estirpe denominada P4, que foi classificada como *Rhodococcus* (EVANS *et al.*, 2004) e posteriormente reclassificada como pertencente a *D. cinnamea* (von der WEID *et al.*, 2007). Esta estirpe apresenta uma forma arredondada a ovóide, não forma esporos e não apresenta mobilidade, cresce em temperaturas acima de 45°C, em pH entre 6 a 10, e cresce também na presença de até 10% NaCl. Suas células aparecem isoladas ou em cadeias e suas colônias apresentam coloração laranja e forma ligeiramente irregular. Em experimentos de biodegradação de compostos de petróleo, *D. cinnamea* P4 mostrou ser capaz de degradar uma faixa bastante ampla de *n*-alcanos (C₁₁-C₃₆), além de carbazol, quinolina, naftaleno, tolueno, gasolina e diesel, sem a presença de outra fonte de carbono (von der WEID *et al.*, 2007). A tolerância desta estirpe de *D. cinnamea* a faixas mais elevadas de temperatura, pH e salinidade, juntamente com sua capacidade em degradar diversos tipos de hidrocarbonetos, proporcionam a este microrganismo um enorme potencial de aplicação em biorremediação de solos contaminados por petróleo.

2. Justificativa

Tecnologias atuais para a remediação de poluentes ambientais que contam somente

com abordagens não biológicas são freqüentemente de custo elevado e podem resultar na formação de intermediários tóxicos que podem se perpetuar no local da contaminação e/ou entrar na cadeia alimentar ou contaminar aquíferos (LOVLEY & LLOYD, 2000). O uso da biorremediação pode ser mais barato que o uso de técnicas convencionais e pode eliminar completamente poluentes orgânicos ou converte-los em formas não tóxicas. Entretanto, um importante fator limitante na biorremediação de sítios contaminados por petróleo é a baixa taxa de degradação pelos microrganismos até então estudados (IWAMOTO & NASU, 2001). Embora dados na literatura relatem diversos microrganismos capazes de degradar compostos de petróleo, muito pouco é conhecido ainda sobre seus sistemas genéticos/bioquímicos responsáveis em assimilar os *n*-alcanos.

Conforme descrito por von der Weid e colaboradores (2007), a estirpe *D. cinnamea* P4 é eficiente na assimilação de cadeias de *n*-alcanos de tamanhos variados (C_{11} - C_{36}) e hidrocarbonetos aromáticos em condições de maior salinidade e temperaturas elevadas, condições essas mais próximas às encontradas em um ambiente contaminado. Com os avanços das novas tecnologias de sequenciamento, uma grande quantidade de informações sobre a genética de microrganismos tem sido gerada, permitindo a descrição de novos genes envolvidos em características importantes em diversos campos da ciência e da indústria. Aliado a isso, análises da expressão gênica em tempo real auxiliam no conhecimento sobre a genética e a fisiologia de microrganismos, bem como sua relação com o nicho onde habita. Os conhecimentos sobre a fisiologia e o genoma da estirpe *D. cinnamea* P4 fornecerão importantes subsídios para futuras aplicações no campo da biorremediação e biocatálise industrial.

3. Objetivos

Os objetivos do presente trabalho foram: (i) Sequenciar e anotar o genoma da estirpe

Dietzia cinnamea P4, (ii) analisar os níveis de expressão dos prováveis genes *alkS*, *alkB* e *alkTrans* através de PCR quantitativo em tempo real e (iii) avaliar o crescimento da estirpe P4 em condições de biorreator (controlada), determinando possíveis alterações no seu genoma através de métodos moleculares.

4. Etapas desenvolvidas

Para atingirmos os objetivos propostos, as seguintes etapas foram realizadas:

- Construção de curvas de crescimento de *D. cinnamea* P4 utilizando diferentes fontes de carbono; obtenção do seu DNA genômico e sequenciamento do genoma de P4 por meio da técnica de pirosequenciamento;
- Determinação das sequências codificantes de proteínas presentes no genoma da estirpe P4 (*coding sequences of proteins* - CDS), utilizando programas de bioinformática e análise das similaridades entre as CDS de *D. cinnamea* P4 e as sequências de aminoácidos de organismos relacionados depositados em bancos de dados (NR, CDD, TransportDB, THMMH, SignalP, KASS e CAZy);
- Avaliação do tempo de geração celular da estirpe P4 em condições de biorreator e determinação de possíveis alterações na população ao longo do tempo através de testes morfológicos (aparência das colônias e mudança de pigmentação) e testes moleculares (BOX-PCR);
- Obtenção do RNA total da estirpe P4 crescida em diferentes condições de crescimento em biorreator de 2 litros e posterior análise dos níveis de expressão dos prováveis genes *alkS*, *alkB* e *alkTrans* de *D. cinnamea* P4 por meio de PCR quantitativo em tempo real.

5. Materiais e Métodos

5.1 – Estirpes bacterianas

Foi utilizada a estirpe *D. cinnamea* P4 (DSM 44994, CIP 1092088), previamente isolada a partir de uma área de proteção ambiental na floresta Atlântica (Reserva Biológica do Poço das Antas), Rio de Janeiro, Brasil. Ela foi mantida a temperatura ambiente em meio LB-ágar inclinado acrescido de 1% de carbonato de cálcio, pH 7,0-7,2.

5.2 – Meios de cultura para o crescimento bacteriano

Meio Luria-Bertani (LB)

triptona	1 % (m/v)
extrato de levedura	0,5 % (m/v)
NaCl	0,5 % (m/v)

O pH foi ajustado para 7,0-7,2 com NaOH 2N. Para a confecção do meio LB sólido, este foi acrescido de ágar bacteriológico 1,5 % (p/v), recebendo a denominação de LB-ágar.

Meio Bushnell-Hass (BH) + *n*-Hexadecano/Benzeno

Sulfato de magnésio	0,02 % (p/v)
Cloreto de cálcio	0,002 % (p/v)
Fosfato de potássio	0,1 % (p/v)
Fosfato de amônio dibásico	0,1 % (p/v)
Nitrato de potássio	0,1 % (p/v)
Cloreto férrico	0,005% (p/v)
<i>n</i> -Hexadecano	1% (v/v)

Todos os meios foram autoclavados a 121°C por 20 minutos. Os compostos *n*-hexadecano foram adicionados após a autoclavação dos meios.

5.3 – Soluções

Tampão fosfato salino 10X (*Phosphate buffered saline 10X*)

NaCl	0,8 % (m/v)
Cloreto de potássio	0,002 % (m/v)
Fosfato disódio de hidrogênio	0,15 % (m/v)
Fosfato monopotássio	0,025 % (m/v)

O pH foi ajustado para 7,0 com NaOH 2N e autoclavado a 121°C por 20 minutos.

5.4 – Primeira etapa: Construção de curvas de crescimento da estirpe *D. cinnamea* P4 em diferentes fontes de carbono, obtenção do DNA genômico, sequenciamento por meio da técnica de pirosequenciamento.

5.4.1 – Construção de curvas de crescimento de *D. cinnamea* P4 em diferentes fontes de carbono

Visando determinar as fases de crescimento de *D. cinnamea* P4 em meio contendo diferentes fontes de carbono, um pré-inóculo de 30 ml de meio líquido LB foi preparado a partir de uma colônia isolada da estirpe P4 previamente crescida em LB-ágar. Em seguida, o mesmo foi incubado à temperatura de 28°C sob agitação de 180 rpm. Quando a densidade óptica da cultura atingiu 0.7 a 660 nm, 10 ml da cultura foram separados, centrifugados a 8.100 x g por 3 minutos e o precipitado ressuspense em 10 ml de tampão PBS 1X. As células ressuspensas em PBS foram centrifugadas nas mesmas condições e ressuspensas no mesmo tampão por 3 vezes. A seguir, essas células lavadas foram utilizadas como inóculo - 0,1% (v/v) - em 50 ml de meio líquido LB ou em meio líquido BH contendo *n*-hexadecano como única fonte de carbono (concentração final de 1% v/v) e de meio líquido BH sem a adição de qualquer fonte de carbono. Todos os meios inoculados foram incubados à temperatura de 28°C sob agitação de 180 rpm. Durante 10 dias, a cada 24 horas, uma amostra de 0,2 ml de cada cultura crescida foi retirada para medir sua densidade óptica em 660 nm. Paralelamente, nos mesmos intervalos de tempo, 100 µl de cada cultura foram diluídos, semeados em placas de LB-ágar e incubados a 32°C para a determinação da unidade formadora de colônias (UFC/ml). Todos os experimentos de contagem foram realizados em triplicata e os experimentos de crescimento foram repetidos pelo menos duas vezes.

5.4.2 – Sequenciamento do genoma da estirpe *D. cinnamea* P4

Após a determinação das diferentes fases de crescimento em cada um dos meios testados, a estirpe P4 foi crescida em 30 ml de meio líquido LB nas mesmas condições descritas no subitem 5.4.1, até atingir a fase exponencial. Em seguida, o DNA genômico de *D. cinnamea* P4 foi obtido seguindo o protocolo de SELDIN & DUBNAU (1985). A integridade e a quantidade do DNA obtido foram observadas por meio de eletroforese em gel de agarose 0,8% (SAMBROOK & RUSELL, 2001), conforme demonstrado na Figura 8.

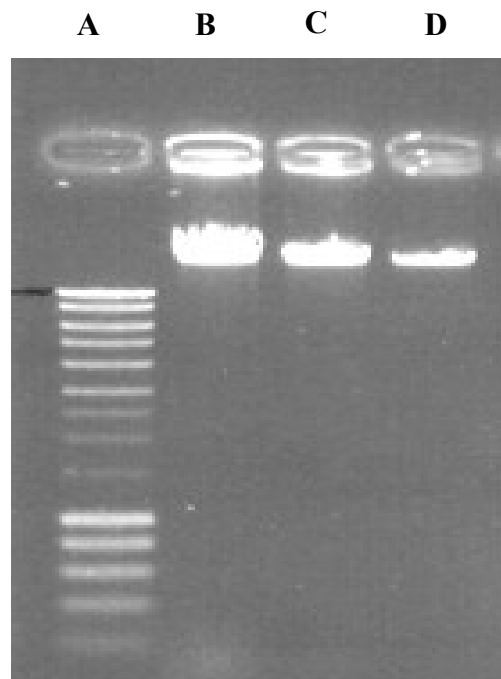


Figura 8: Perfil do DNA genômico de *Dietzia cinnamea* P4. (A) *Smart Ladder*, (B) 2 μ l, (C) 2 μ l da diluição de 1/10 da amostra B, (D) 1 μ l da diluição de 1/10 da amostra C.

Após verificar que o DNA obtido se encontrava íntegro, cerca de 2 μ g do DNA total foram acondicionados em tubo de polietileno de 2 ml e enviados para sequenciamento no Instituto de Microbiologia da Universidade de Copenhague, Dinamarca. A estratégia adotada para o sequenciamento do genoma de *D. cinnamea* P4 foi o pirosequenciamento disponibilizado pela plataforma 454 GS FLX da Roche.

5.5 – Segunda etapa: Construção das sequências codificantes de proteínas (CDSs) presentes no genoma da estirpe P4, utilizando programas de bioinformática e análise da similaridade entre as CDSs de *D. cinnamea* P4 e as sequências de aminoácidos de organismos relacionados depositados em bancos de dados.

5.5.1 – Análise de similiaridade utilizando as sequências de DNA

O sequenciamento do genoma de *D. cinnamea* P4 deu origem a 565.157 pequenas sequências, chamadas de *reads*. Associações de sobreposições de sequências agruparam as 565.157 *reads* em 435 grandes sequências, chamadas de *contigs*. As *contigs* foram utilizadas para buscar a presença de tRNAs, rRNAs e elementos de inserção (IS) utilizando os servidores tRNAScan-SE (<http://lowelab.ucsc.edu/tRNAscan-SE/>), rDNA Database e IS Finder (<http://www-is.biotoul.fr/>), respectivamente.

5.5.2 – Construção das sequências codificantes de proteínas (CDSs)

Utilizando o programa GLIMMER versão 3.02 disponibilizado pelo Centro de Bioinformática e Biologia Computacional da Universidade de Maryland (<http://www.cbcu.umd.edu/software/glimmer/>) (DELCHER *et al.*, 2007) e o programa *The European Molecular Biology Open Software Suite* (EMBOSS) disponibilizado no sítio <http://emboss.sourceforge.net/index.html> foram geradas 5482 *coding sequences of proteins* (CDS) a partir das 435 *contigs*. Após análise mais criteriosa foram selecionadas 5071 CDSs, as quais foram utilizadas para as análises de similaridade com sequências de outros microrganismos relacionados.

5.5.3 – Análise computacional das sequências

Para as análises de similaridade com sequências de genes funcionais de bactérias relacionadas foi empregado o programa BLASTALL disponibilizado pelo *National Center for Biotechnology Information* (NCBI) no sítio <ftp://ftp.ncbi.nih.gov/blastall> e o banco de dados CDD disponível no mesmo sítio. Paralelamente, todas as CDSs foram analisadas manualmente utilizando o programa BLASTP disponível no sítio <http://blast.ncbi.nlm.nih.gov/Blast.cgi> utilizando o banco NR como referência. Os domínios transmembrânicos presentes nas CDSs foram identificados através dos programas TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>) e TransportDB (<http://www.membranetransport.org/>), enquanto as vias metabólicas foram categorizadas utilizando o programa KASS e o banco de dados KEGG (<http://www.genome.jp/kegg/>). As

sequências peptídeo-sinais nas prováveis proteínas de *D. cinnamea* P4 foram localizadas empregando o programa SignalP versão 3.0 disponível no sítio <http://www.cbs.dtu.dk/services/SignalP/output.php>. As prováveis enzimas presentes entre as CDSs da estirpe P4 envolvidas na degradação, modificação e criação de ligações glicosídicas foram identificadas através de buscas de similaridade com as sequências depositadas no banco *Carbohydrate-Active Enzymes Database* (CAZy), disponível no sítio <http://www.cazy.org/> (Figura 9).

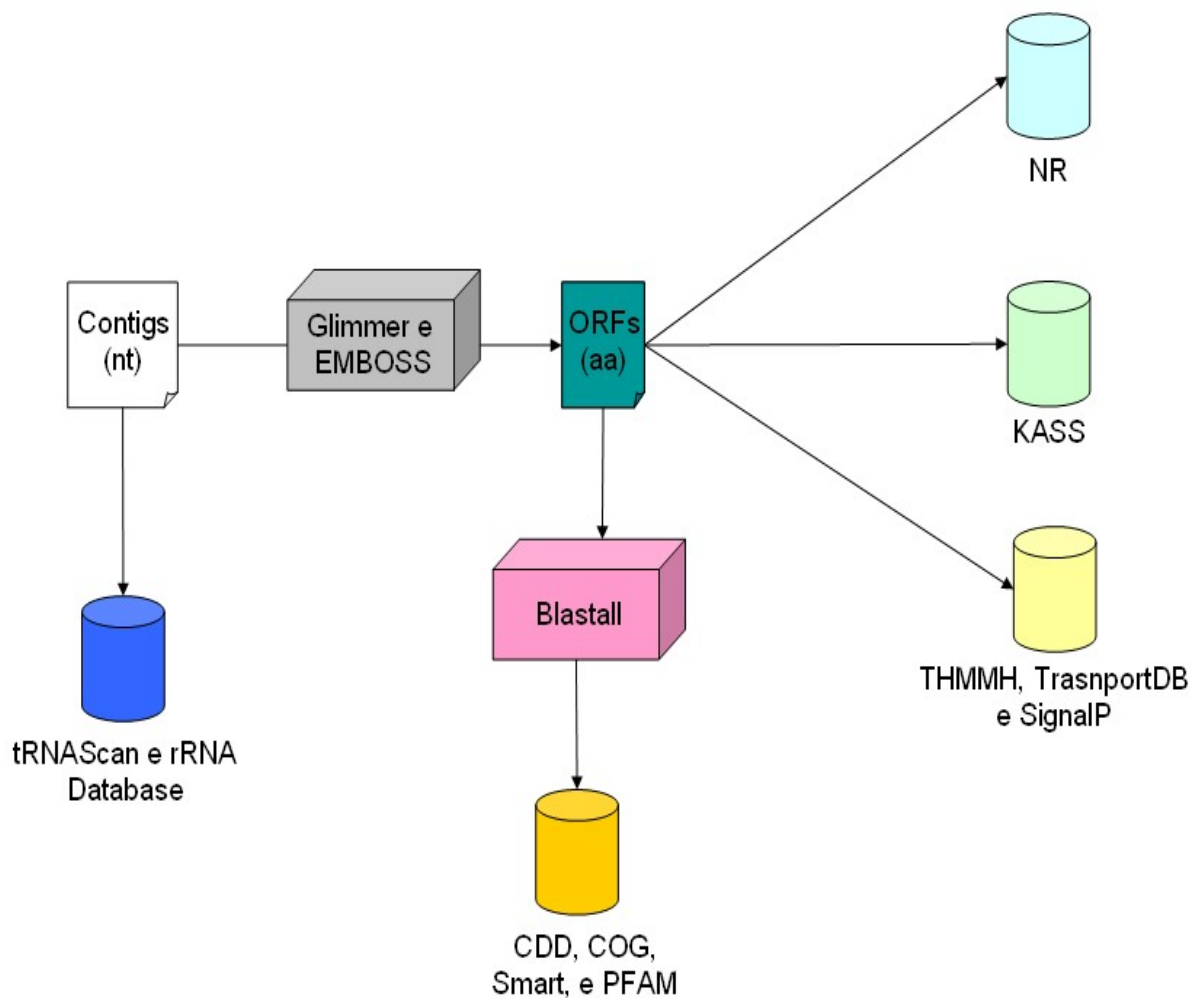


Figura 9: Fluxograma representando as etapas das análises das sequências de *contigs* (DNA) e das CDSs (sequências de aminoácidos) de *D. cinnamea* P4 com os diversos bancos de sequências disponíveis.

5.5.4 – Anotação dos resultados de similaridade

A significância de alinhamento entre as sequências relacionadas foi determinada pelo valor de E (e-value), que categoriza o índice de confiabilidade dos dados analisados em uma combinação de sequências (MCGINNIS & MADDEN, 2004). Desta forma, o critério adotado foi um valor de E de alta significância $E < 10^{-20}$, média significância $10^{-19} \leq E \leq 10^{-7}$, baixa significância $10^{-6} \leq E \leq 1$, e insignificante $E \geq 1.1$. Além do valor de E, também foi levada em consideração a porcentagem de cobertura, a qual indica a extensão da sobreposição da sequência analisada com as sequências similares de outros organismos depositadas nos bancos de dados. Esta condição adicional foi empregada porque uma sequência analisada poderia apresentar um valor de E de alta significância, porém apenas sobre uma pequena extensão de sobreposição. Assim, foram considerados como prováveis genes funcionais as CDSs com valor de E de alta ou média significância e com índice de cobertura de pelo menos 30%.

5.6 – Terceira etapa: Avaliação do tempo de geração celular da estirpe P4 em condições de biorreator e determinação de possíveis alterações na população ao longo do tempo através de testes morfológicos (aparência das colônias e mudança de pigmentação).

5.6.1 – Determinação do tempo de geração celular de *D. cinnamea* P4 em biorreator de 2 litros

Preparo do biorreator e dos meios de cultura

Visando a análise do crescimento e da manutenção das células de P4 em biorreator de dois litros, um pré-inóculo em meio líquido LB ou BH contendo *n*-hexadecano como única fonte de carbono (na concentração final de 1%) foi preparado em erlenmeyers de 125 ml, seguindo os mesmos passos descritos no subitem 5.3.1.

Em um biorreator de bancada com uma cuba de vidro refratário com capacidade útil de dois litros, foi preparado um litro de meio líquido LB ou BH. O biorreator estava equipado com três chicanas, duas turbinas de seis pás planas, dispensador de ar em forma de L perfurada, controle de pH e oxigênio dissolvido por meio de eletrodos e sensores de temperatura e espuma. O controle de temperatura foi realizado pelo sistema *ADI 1018 Thermo Cicular*, através de uma camisa externa. Além disso, o biorreator também estava equipado com três bombas peristálticas para a adição de base, ácido e meio de cultura, e com outra bomba peristáltica adicional para a retirada de meio de cultura. Os parâmetros de pH, oxigênio dissolvido, temperatura e nível de espuma foram controlados pelo sistema externo

ADI 1010 Bio controller do fabricante Applikon (Figura 10). Este sistema foi conectado a um computador, o qual permitiu a coleta de dados a cada hora por meio do software *BioXpert v2*, do mesmo fabricante.

As condições básicas de operação para o crescimento de *D. cinnamea* P4 foram: (1) frequência do agitador de 250 rpm, (2) temperatura de 35°C, (3) pH 7 ou 9 mantidos pela adição de NaOH 2N e HCl 2N, (4) concentração do oxigênio dissolvido de 10% de saturação mantida no nível mínimo através do fluxo de ar e da frequência do agitador e (5) controle da pressão interna realizado por meio de um manômetro de mercúrio com tubo em forma de “U”. A adição de *n*-hexadecano foi realizada após o biorreator ter sido autoclavado. Para isso foi utilizado um sistema de transferência estéril através de tubos de silicone e injeções.

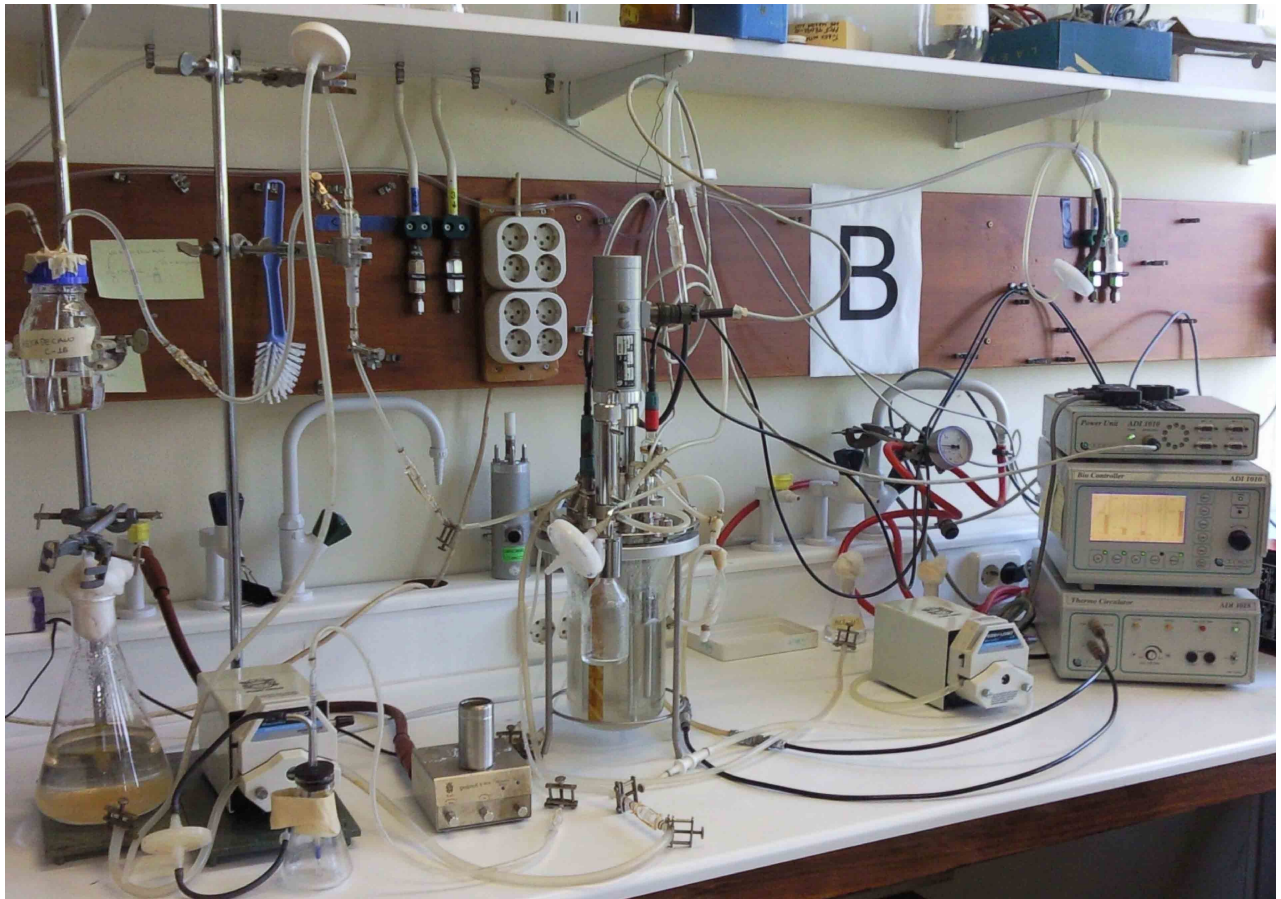


Figura 10: Sistema de fermentação de bancada composto por um biorreator de 2 litros, 3 bombas peristálticas, sistemas de transferência de meio de cultura e fonte de carbono e sistema de monitoramento de dados em tempo real.

Após a autoclavação do biorreator, este foi resfriado à temperatura ambiente, o pH ajustado para a condição desejada (pH 7 ou 9) e, quando o meio de cultura utilizado era o BH, foi adicionado *n*-hexadecano na concentração final de 1%. Em seguida, 10 ml do pré-inóculo descrito anteriormente foi adicionado através de um sistema de transferência estéril semelhante ao descrito para *n*-hexadecano.

Coleta de amostras e medidas de crescimento de D. cinnamea P4 em biorreator em diferentes condições de pH

Após um período de 12 horas de cultivo, foram coletadas amostras de 2 ml a cada 2 horas, por um período de 3 dias para determinar a taxa de crescimento celular. Cada amostra foi centrifugada e lavada 3 vezes, seguindo os mesmos passos descritos no subitem 5.3.1. Em seguida, 100 µl da cultura foram semeados em LB-ágar e as placas incubadas a 28°C para determinação da UFC/ml. Paralelamente, 1 ml da amostra foi separada e sua densidade óptica determinada em 660 nm. Após este período de 72 horas, a periodicidade de coleta de amostras foi a cada 24 horas. Todas as amostragens foram realizadas em triplicata e pelo menos 2 experimentos independentes de crescimento foram realizados para as condições de diferentes fontes de carbono e de pH (7 e 9).

5.6.2 –Análise das alterações no genoma de *D. cinnamea* P4 após longo período de cultivo por meio de BOX-PCR.

Para a detecção de possíveis alterações no genoma da estirpe P4 após longo período de cultivo na presença de *n*-hexadecano, foram escolhidas aleatoriamente cerca de 20 colônias crescidas em placa LB-ágar da amostragem referente ao 45º dia (correspondendo a aproximadamente 300 gerações). Estas colônias foram inoculadas separadamente em 10 ml de meio líquido LB e incubadas à temperatura de 28°C, sob agitação de 180 rpm. Após as culturas terem alcançado a fase exponencial, o DNA genômico de cada cultura foi obtido seguindo o protocolo de SELDIN & DUBNAU (1985). O DNA resultante foi analisado quanto a sua integridade e foi quantificado por meio de eletroforese em gel de agarose 0,8% (SAMBROOK & RUSELL, 2001). Além disso, também foi obtido o DNA genômico das amostras referentes ao 45º dia, retiradas diretamente do biorreator. Essa amostra recebeu a denominação de *mix*.

Três concentrações diferentes de DNA, 10 ng, 1 ng e 0,1 ng, obtidas de cada uma das 20 colônias escolhidas, da cultura *mix* e da estirpe antes de ser incubada em birreator foram utilizadas em uma reação de BOX-PCR (Reação em Cadeia da Polimerase). As condições para a reação de BOX- PCR foram as seguintes:

Componentes	Concentração final
H ₂ O livre de nucleases	-
Tampão 10X com MgCl ₂ 25 mM/μl	1X
dNTP Mix 10 μM/μl	0,5 μM/μl
DMSO 100%	10 %
Iniciador BOX GTG5 30 μM/μl	1,2 μM/μl
Enzima <i>Taq</i> Polimerase 1U/μl	0,2 U/μl
DNA genômico de <i>D. cinnamea</i> P4	1ng

O programa de amplificação utilizado foi:

- 1) Desnaturação: 95°C por 5 minutos.
- 2) Desnaturação: 95°C por 1 minuto.
- 3) Anelamento: 65°C por 1 minuto.
- 4) Extensão: 72°C por 5 minutos.
- 5) Repetição das etapas 2, 3 e 4 por 40 vezes.
- 6) Extensão final: 72°C por 8 minutos.

Os produtos da PCR foram submetidos à eletroforese em gel de agarose 2% por 24 horas, para a total separação de todos os fragmentos amplificados. Todos os experimentos de BOX-PCR foram realizados em triplicata e pelo menos dois experimentos de amplificação por PCR foram feitos em separado.

5.7 – Quarta etapa: Obtenção do RNA total da estirpe P4 crescida em diferentes condições de crescimento em biorreator de 2 litros, síntese dos cDNAs referentes aos RNAs e análise

dos níveis de expressão dos prováveis genes *alkS*, *alkB* e *alkTrans* de *D. cinnamea* P4 por meio de PCR quantitativo em tempo real.

5.7.1 – Amplificação dos genes *alkS*, *alkB* e *alkTrans*

A partir das sequências obtidas através do sequenciamento do genoma de *D. cinnamea* P4, foram desenhados iniciadores para os prováveis genes *alkS*, *alkB* e *AlkTrans* utilizando o programa *Clone Manager*. Esses iniciadores estão listados a seguir.

-AlkB-F (Eurogentec) - 21 nt - 5'-AGC GAG AAG TGG CTG TCC AA-3'

-AlkB-R (Eurogentec) - 21 nt - 5'-GGT GAT GTC GCC GTC GTA GT-3'

-AlkS-F (Eurogentec) - 18 nt - 5'-ATG CCG ATG CGG GCG ATG-3'

-AlkS-R (Eurogentec) - 18 nt - 5'-TGT TCG AGG CCA TGC ACG-3'

-AlkTrans-F (Eurogentec) - 20 nt - 5'-TCC TCA TCC TCT CCG TCT TC-3'

-AlkTrans-R (Eurogentec) - 20 nt - 5'-CGG TCA TCT GGT CGT TCA TC-3'

-GTG5 (Eurogentec) - 15 nt - 5'-GTG GTG GTG GTG GTG-3'

Para verificar a eficiência desses iniciadores, os mesmos foram empregados na amplificação por PCR dos respectivos genes, utilizando como molde o DNA genômico da

estirpe P4. As condições para as reações de PCR foram as seguintes:

Componentes	Concentração final
H ₂ O livre de nucleases	-
Tampão 10X	1X
MgCl ₂ 25 mM/μl	1,5 mM/μl
dNTP Mix 10 μM/μl	0,5 μM/μl
Iniciador 3' 10 μM/μl	1 μM/μl
Iniciador 5' 10 μM/μl	1 μM/μl
Enzima <i>Taq</i> Polimerase	0,2 U/μl
DNA genômico de <i>D. cinnamea</i> P4	1 ηg

O programa de amplificação utilizado foi:

- 1) Desnaturação: 95°C por 5 minutos.
- 2) Desnaturação: 95°C por 1 minuto.
- 3) Anelamento: 58°C por 1 minuto.
- 4) Extensão: 72°C por 3 minutos.
- 5) Repetição das etapas 2, 3 e 4 por 35 vezes.
- 6) Extensão final: 72°C por 7 minutos.

Os produtos da PCR foram analisados em eletroforese em gel de agarose 0,8% para a confirmação do tamanho dos produtos esperados.

5.7.2 – Obtenção do RNA total de *D. cinnamea* P4

Foram preparadas duas culturas da estirpe P4, uma em meio líquido LB e outra em BH contendo glicose ou *n*-hexadecano como única fonte de carbono, na concentração final de 1% (v/v). Para o inóculo nos meios (0,1% v/v), a P4 foi crescida conforme descrito no subitem 5.3.1. Na fase exponencial de crescimento da estirpe P4 em meio LB (correspondente à densidade óptica de 1.0 a 660 nm), 3 ml da cultura foram centrifugados a 8.100 g e o precipitado lavado 3 vezes, conforme descrito anteriormente no subitem 5.3.1. O mesmo procedimento foi empregado nas culturas contendo *n*-hexadecano, contudo as amostras foram coletadas após 1, 48 e 168 horas de crescimento, as quais correspondiam às fases lag, exponencial e estacionária, respectivamente.

A obtenção do RNA total da estirpe P4 nas diferentes condições de crescimento foi realizado utilizando o *Kit RNeasy Mini* do fabricante Quiagen. Entretanto, um passo adicional foi introduzido ao protocolo devido à dificuldade em promover a lise da estirpe P4. O

precipitado celular foi ressuspenso no tampão RLT fornecido pelo *Kit RNeasy Mini* e as células foram lisadas utilizando nitrogênio líquido. Os passos seguintes foram os recomendados pelo fabricante Quiagen. A integridade do RNA total obtido a partir de cada condição foi analisada por meio de eletroforese em gel de agarose 0,8%. A quantidade do RNA foi determinada por densidade óptica em espectrofotômetro NanoDrop® ND-1000 UV-Vis. Posteriormente, as concentrações de cada amostra de RNA total foram ajustadas para cerca de 100 ng/μl.

5.7.3 – Síntese dos cDNAs por meio de transcrição reversa

Cerca de 1 μg de RNA total de cada amostra descrita acima foi utilizado para a síntese de cDNA. Para isso, foi utilizado o *Kit cDNA Synthesis* do fabricante Bioline. Resumidamente, o protocolo do fabricante recomenda dois passos para a síntese de cDNA a partir de RNA total. No primeiro passo, os iniciadores randômicos são anelados à fita de RNA através da reação:

Componentes	Volume
0,5 μg de RNA total	5 μl
Iniciadores randômicos	1 μl
dNTP 10mM	1 μl
H ₂ O livre de nucleases	3 μl

Os componentes descritos acima foram incubados a 65°C por 10 minutos, em seguida incubados em gelo por 2 minutos e utilizados para a síntese do cDNA através da reação:

Componentes	volume
Reação de anelamento dos iniciadores randômicos	10 μl
Tampão RT 5X	4 μl

Inibidor de RNase	1 µl
Transcriptase reversa (200U/ µl)	0,25 µl
H ₂ O livre de nucleases	4,5 µl

A reação foi incubada a 42°C por 1 hora e em seguida a 70°C por 15 minutos. Após esse período, a mesma foi resfriada em gelo por dois minutos e acondicionada a -20°C. Posteriormente, 2µl de cada reação de cDNA foram separados para a análise de pureza e quantificação por densidade óptica em espectrofotômetro NanoDrop® ND-1000 UV-Vis. A seguir as concentrações de cDNA foram ajustadas para 0,5 ηg/µl cada.

5.7.4 – Quantificação dos níveis transcricionais dos prováveis genes *alkS*, *alkB* e *alkTrans* por meio de PCR quantitativo em tempo real (qPCR)

Amostras contendo 2µl de cDNA correspondentes às diferentes condições de crescimento descritas no subitem 5.7.2 foram utilizadas para comparar os níveis de expressão dos prováveis genes *alkS*, *alkB* e *alkTrans* por meio de qPCR empregando o *kit Master Mix SYBR Green* do fabricante Applied Biosystem e os respectivos iniciadores de cada provável gene. As condições da reação foram:

Componentes	Volume final
H ₂ O livre de nucleases	9,8 µl
Master Mix SYBR Green	12 µl
Inciador 3' 10 µM/µl	0,6 µl
Inciador 5' 10 µM/µl	0,6 µl
cDNA	2 µl

O programa de amplificação utilizado foi:

- 1) Ativação da Enzima: 95°C por 10 minutos.
- 2) Desnaturação: 95°C por 30 segundos.
- 3) Anelamento: 60°C por 1 minuto.
- 4) Extensão: 72°C por 1 minuto.
- 5) Repetição das etapas 2, 3 e 4 por 40 vezes.

5.7.4 – Quantificação dos níveis transcricionais dos prováveis genes *alkS*, *alkB* e *alkTrans* por meio de PCR quantitativo em tempo real (qPCR)

A detecção dos produtos foi feita pelo contínuo monitoramento do sinal fluorescente emitido pelo corante *SYBR Green*. Os valores quantitativos (CT) foram obtidos a partir do ciclo limiar onde o aumento no sinal associado ao crescimento exponencial dos produtos de PCR começa a ser detectado. Todas as amostras foram analisadas em triplicata e todo experimento realizado pelo menos duas vezes.

O nível de expressão gênica dos genes *alkS*, *alkB* e *alkTrans* foi determinado pela fórmula: $2^{-\Delta\Delta CT}$, onde ΔCT é a diferença entre a média do CT da amostra amplificada para o gene alvo e a média do CT da mesma amostra amplificada para o gene referência 16S e $\Delta\Delta CT$ corresponde à diferença entre o ΔCT da amostra de interesse e o ΔCT da amostra referência.

6. Resultados

I - Análise do genoma da estirpe *Dietzia cinnamea* P4

6.1 – Características gerais do genoma de *D. cinnamea* P4

O sequenciamento do genoma de *D. cinnamea* P4 foi realizado no Laboratório de Evolução e Microbiologia da Universidade de Copenhague, Dinamarca. A metodologia empregada foi o pirosequenciamento na plataforma 454 GS FLX da empresa Roche. Foram geradas 565.157 *reads*, o que corresponde a cerca de 30 leituras do genoma inteiro de *D.*

cinnamea P4. A associação subsequente realizada entre as *reads* resultou em 435 *contigs* com tamanho médio de 8.092 nt. Na figura 11 é apresentado o número de *contigs* e seus respectivos tamanhos. A maioria dos *contigs* apresentou tamanho médio entre 1.001 a 5.000 nt, seguido por três grupos com tamanhos entre 5.001 e 50.000 nt.

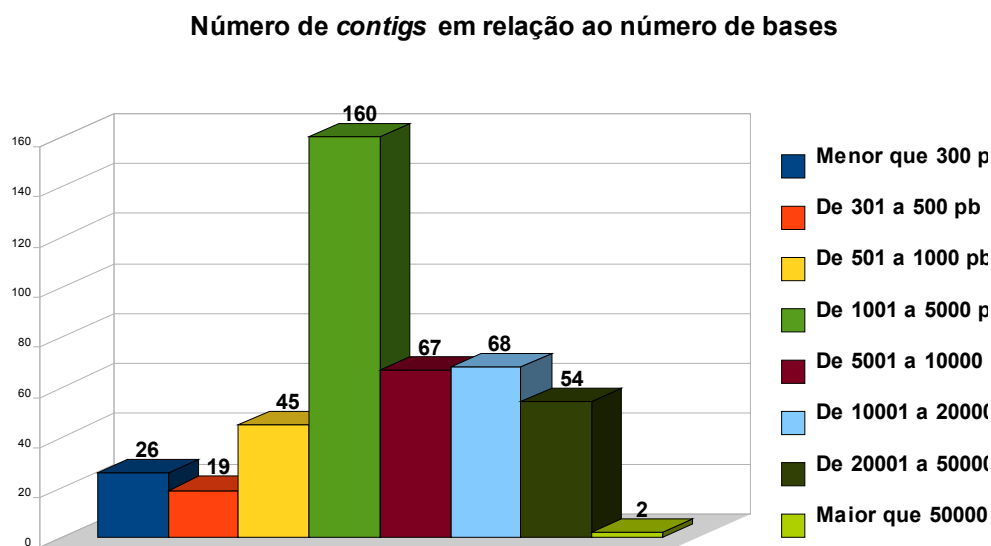


Figura 11: Número de *contigs* de diferentes tamanhos obtidos após o sequenciamento do genoma da estirpe P4.

A distribuição dos nucleotídeos no genoma da estirpe P4 é apresentada na figura 12. O seu conteúdo G+C% é de 70,9%, semelhante ao valor descrito para outros actinomicetos (Tabela 1). A estirpe P4 é a primeira espécie do gênero *Dietzia* a ter o seu genoma sequenciado. Uma das maiores dificuldades encontradas no sequenciamento foi devido ao seu alto conteúdo G+C. Durante a construção das sequências codificantes de proteínas (CDS), utilizando os programas GLIMMER e EMBOSS, também foram encontradas diversas dificuldades. Isso devido também à presença do alto conteúdo G+C no genoma de *D. cinnamea* P4, que acarreta na predição de um número elevado, mas falso de CDS, pela suposta presença de inúmeras sequências de terminação. Para minimizar a ocorrência de possíveis CDS falsas, foram atribuídas diversas “penalidades”, como por exemplo, restringir o tamanho das prováveis CDS. Apenas CDS construídas a partir de 600 nt foram consideradas, o que representa o dobro do que é comumente utilizado para outros genomas.

Distribuição das bases nucleotídicas

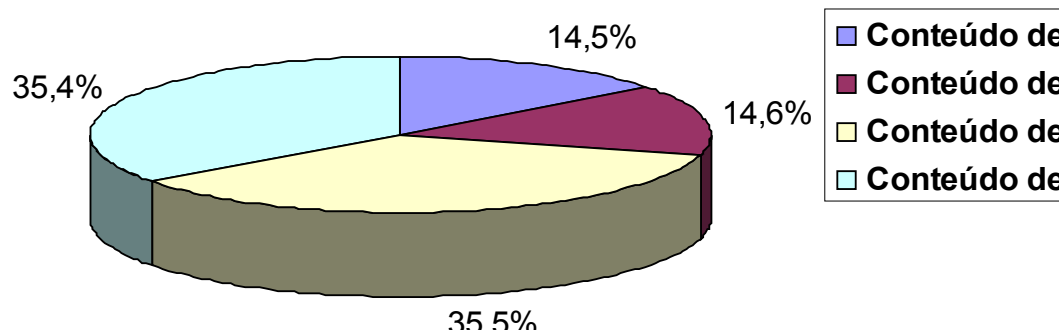


Figura 12: Percentagem de cada um dos nucleotídeos no genoma da estirpe P4.

Com a utilização dos programas GLIMMER 3.02 e EMBOSS, cerca de 5071 CDSs foram obtidas a partir das 435 *contigs*. A seguir, as CDSs foram utilizadas para buscas de similaridades com outras sequências de organismos relacionados depositados em bancos de dados, conforme descrito nos Materiais e Métodos.

O resumo das principais características sobre o genoma de P4 é apresentado na Tabela 1 e na figura 13. O número de proteínas envolvidas na regulação gênica foi menor do que o esperado, apenas 72 até o momento (subitem 6.17). As oxigenases, enzimas responsáveis na inserção de um ou dois átomos de oxigênio em compostos orgânicos estão bastante representadas no genoma de *D. cinnamea* P4 (veja subitem 6.7), aproximando em número das oxigenases relatadas em estirpes conhecidas como degradadoras de hidrocarbonetos de petróleo. As proteínas envolvidas no transporte de solutos estão representadas significativamente no genoma de *D. cinnamea* P4 (subitem 6.15), o que corrobora com a importância na aquisição de compostos presentes no meio extracelular e na extrusão de metabólitos, enzimas de ação extracelular e troca iônica.

Tabela 1: Comparação de características selecionadas entre o genoma de *D. cinnamea* P4 e outras bactérias relacionadas.

Organismo	Tamanho Mb ^a	%G+C ^a	Reguladores ^a	Oxigenases ^a	Proteínas Transportadoras ¹
<i>D. cinnamea</i> P4	3,5 ^b	70,9	72	64	312
<i>Rhodococcus</i> sp. RHA1	9,7	67,0	705	203	890
<i>Nocardia farcinica</i> IFM	6,3	70,8	448	96	838

10152					
<i>Burkholderia xenovorans</i>	9,7	62,6	745	134	1139
LB400					
<i>Corynebacterium</i>	3,3	65,6	148	27	402
<i>glutamicum</i> ATCC 13032					
<i>Acinetobacter</i> sp. ADP1	3,6	40,4	188	52	368
<i>Streptomyces coelicolor</i>	8,7	72,1	671	70	838
A3(2)					
<i>Polaromonas</i> sp. JS666	5,2	62,5	405	63	801

^a (McLEOD *et al.*, 2006)

^b Tamanho aproximado

Na figura 13 é apresentada em porcentagem a distribuição funcional dos prováveis genes presentes no genoma de *D. cinnamea* P4. A categoria de “Aquisição/conversão de energia” representou 7,7% de todas as outras categorias apresentadas. Esse foi um valor pouco acima do esperado quando comparado com outros genomas sequenciados de organismos relacionados. Normalmente esses organismos apresentam cerca de 4% a 7% de seus genes envolvidos neste tipo de processo metabólico. O actinomiceto de solo *Kocuria rhizophila* emprega 6,8% de seu genoma para aquisição/conversão de energia (TAKARADA *et al.*, 2008), enquanto a bactéria metanogênica *Methanococcus maripaludis* 4,7% (HENDRICKSON *et al.*, 2004), *Micrococcus luteus* (Fleming) 6,7%, *Arthrobacter aurescens* TC1 e *Arthrobacter* sp. FB24, 6,4% e 7,1%, respectivamente (MONGODIN *et al.*, 2006). O valor de 7,7% apresentado pela estirpe P4 em genes envolvidos em processos de aquisição/conversão de energia reflete o seu metabolismo versátil. Outra categoria de genes que foi representada acima do esperado (3,4% dos genes) foi a de “Respostas a estresses”. Essa observação poderia sugerir uma adaptação genética da *D. cinnamea* P4 para sobreviver em um ambiente complexo como o solo.

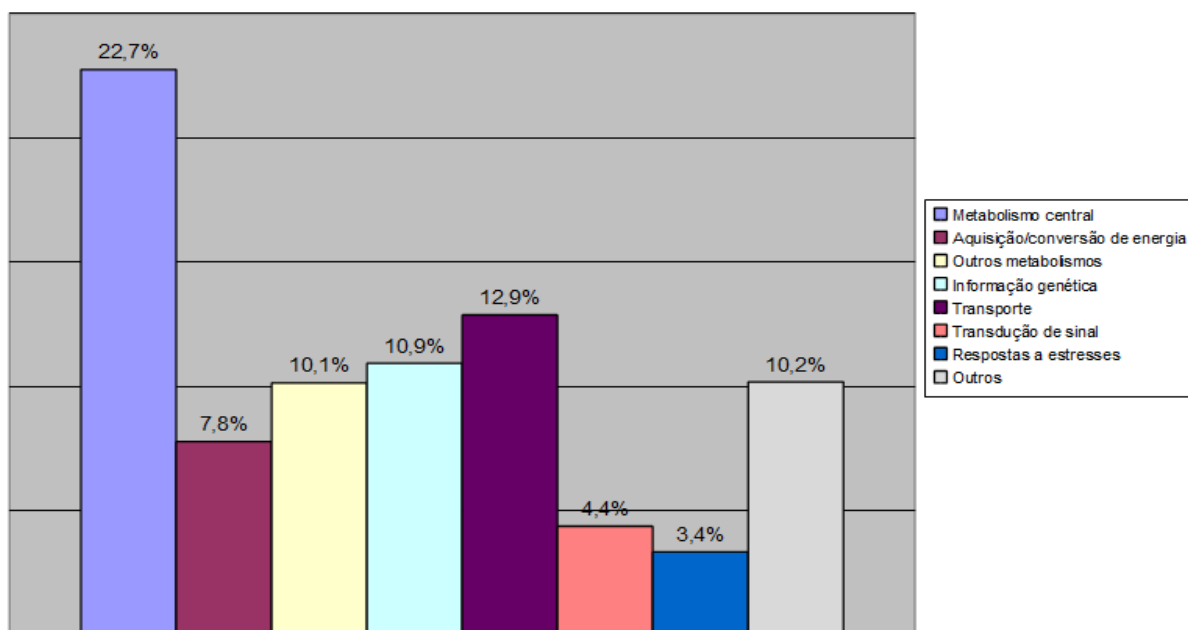


Figura 13: Distribuição funcional dos prováveis genes presentes no genoma de *D. cinnamea* P4.

6.2 – Similaridades entre as CDS de *D. cinnamea* P4 com actinomicetos relacionados

Como apresentado na figura 14, dentre as 5071 CDSs do genoma de *D. cinnamea* P4, cerca de 41% apresentaram similaridades com genes anotados no genoma de espécies do gênero *Rhodococcus*, 15% com espécies do gênero *Mycobacterium* e 13% com espécies do gênero *Corynebacterium* spp.; enquanto 14% foram similares com as estirpes *Gordonia bronchialis* DSM 43247, *Norcadia farcinica* IFM10152 e *Tsukamurella paurometabola* DSM 20162.

Principais organismos com similaridade com genes da estirpe P4

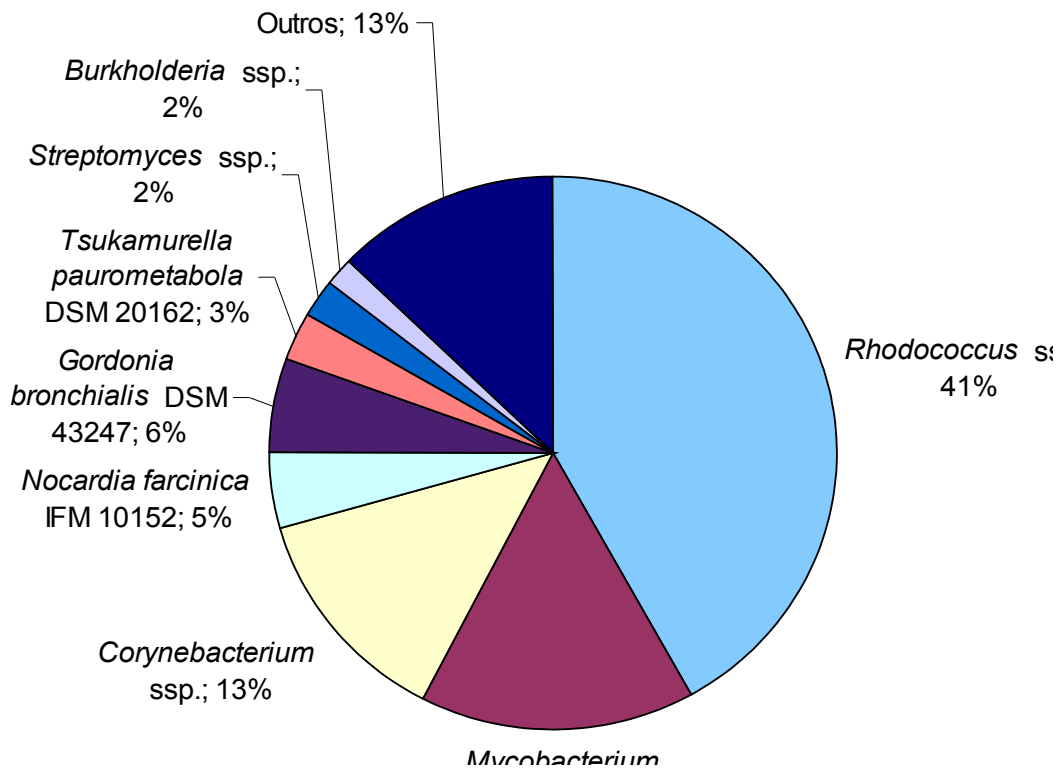


Figura 14: Distribuição das CDSs de *D. cinnamea* P4 com similaridade com organismos relacionados.

Conforme o esperado para a grande maioria de bactérias Gram-positivas com alto conteúdo G+C, não foram detectados genes hipotéticos envolvidos na síntese de polihidroxiturato, um composto normalmente utilizado como armazenamento de energia por células Gram-negativas, nem as prováveis enzimas capazes de hidrolisar tais compostos. O sequenciamento de *D. cinnamea* P4 também não revelou a presença de nenhum gene provável envolvido em mobilidade e síntese de componentes de flagelo, confirmando os resultados relatados por von der Weid e colaboradores (2007) e semelhante às observações encontradas no genoma do actinomiceto de solo *Arthrobacter aurescens* TC1 (MONGODIN *et al.*, 2006).

6.3 – Metabolismo central

Todos os genes hipotéticos necessários para a oxidação da glicose a piruvato (DCP40001-12) e para a subsequente conversão de piruvato a acetil Coenzima A (acetil-CoA) foram identificados no genoma da estirpe P4 (DCP40033, DCP40036). A acetil-CoA pode ser direcionada para o Ciclo de Ácido Cítrico (TCA), o qual é representado no genoma de *D. cinnamea* P4 através dos genes putativos citrato sintase (DCP40189), aconitase (DCP40199, DCP40200), isocitrato desidrogenase (DCP40188, DCP40195), oxaglutarato desidrogenase componentes E1 e E2 (DCP40196, DCP40194), succinil CoA sintase subunidade alfa e beta (DCP40186, DCP40185), succinato desidrogenase (DCP40191-92), fumarato hidratase (DCP40198) e malato desidrogenase (DCP40190). É sugerida a conversão de piruvato a L-lactato pela presença do gene putativo L-lactato desidrogenase (DCP40020) e a produção de etanol a partir de acetil-CoA através da presença dos genes hipotéticos álcool desidrogenase (DCP40031-32, DCP40037-38), aldeído desidrogenase (DCP40030, DCP40040), acetil-CoA sintase (DCP40034) (Figura 15). A capacidade de utilizar o glicerol como um intermediário da via glicolítica é indicada pela presença dos prováveis genes glicerol quinase (DCP40110) e glicerol 3-fosfato desidrogenase (DCP40111), sugerindo o uso de compostos de glicerol como fonte de carbono pela célula. A análise do genoma da estirpe P4 sugere a presença da via metabólica da gliconeogênese. Nesta via, uma molécula de oxalacetato é convertida a fosfoenolpiruvato através da provável enzima fosfoenolpiruvato carboxilase (DCP40013). Em seguida o composto fosfoenolpiruvato é direcionado através da via glicolítica, porém no sentido contrário. Todas as enzimas necessárias para conversão da fosfoenolpiruvato a glicose-6-fosfato operam em ambas as direções, com exceção da 6-fosfofrutoquinase, a qual é substituída pela reação catalisada pela enzima hipotética frutose-1,6-bifosfatase, fpb (DCP40013).

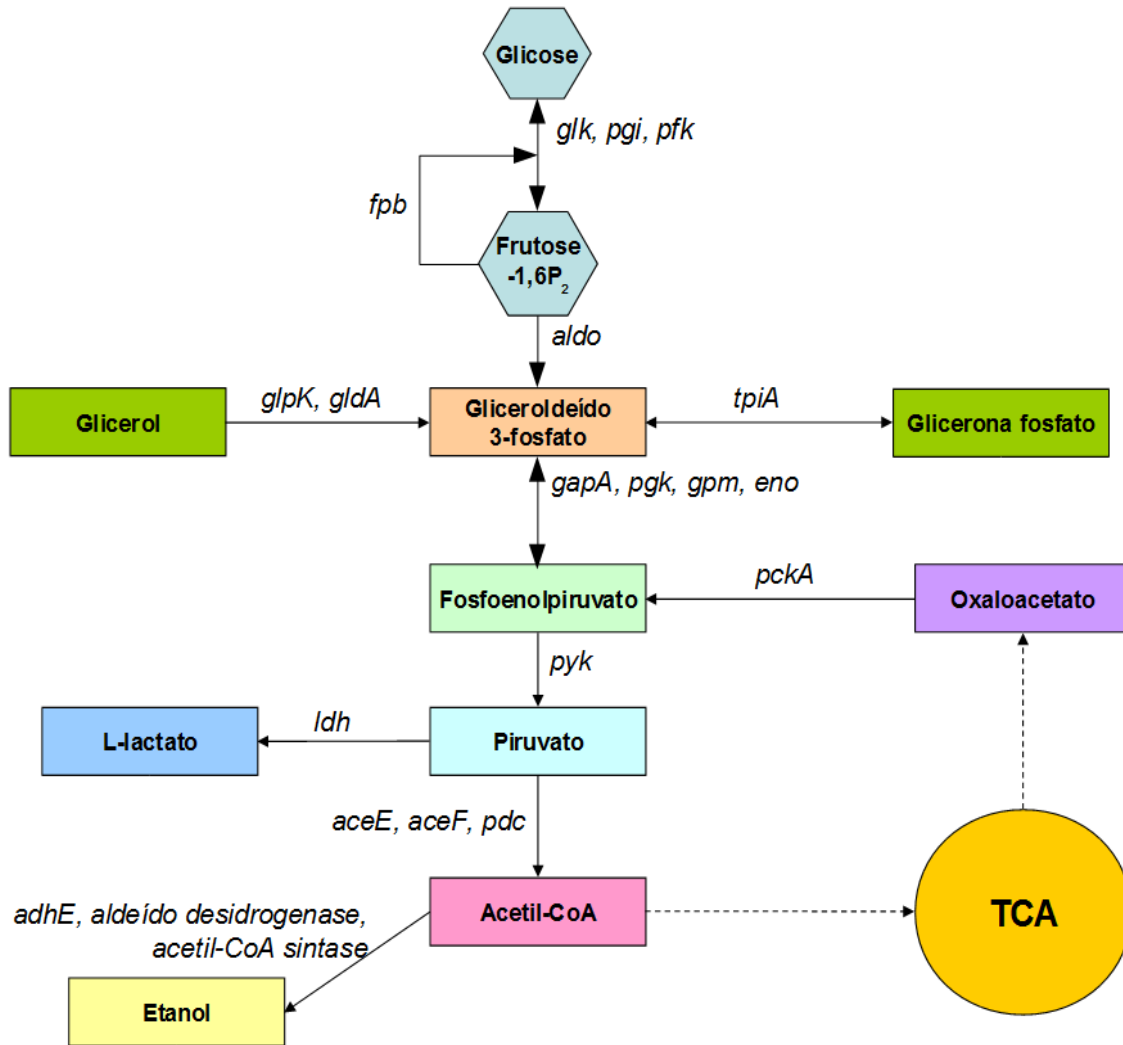


Figura 15: Esquema representativo da via central glicolítica e os genes presentes no genoma da estirpe P4. O composto inicial é a glicose que é convertida a piruvato e acetil-CoA. O piruvato pode ser convertido a lactato e a molécula de acetil-CoA pode ser direcionada para ser totalmente oxidada a CO_2 na via TCA ou para a produção de etanol. Abreviaturas correspondentes aos nomes dos genes: *glk*, glicoquinase; *pgi*, glicose 6-fosfato isomerase; *pfk*, 6-fosfofrutoquinase; *fpb*, frutose-1,6-bifosfatase; *aldo*, frutose-bifosfato aldolase; *tpiA*, triose fosfato isomerase; *gapA*, gliceroldeído 3-fosfato desidrogenase; *pgk*, fosfoglicerato quinase; *gpm*, fosfoglicerato mutase; *eno*, enolase; *pyk*, piruvato quinase; *aceE*, piruvato desidrogenase componente E1; *aceF*, piruvato desidrogenase componente E2; *pdc*, piruvato decarboxilase; *pckA*, fosfoenolpiruvato carboxiquinase, *adhE*, álcool desidrogenase; *ldh*, lactato desidrogenase; *glpK*, genes glicerol-3-fosfato:glicerol quinase; *gldA*, glicerol 3-fosfato desidrogenase (LEHNINGER *et al.*, 2007).

A detecção dos genes hipotéticos isocitrato liase e malato sintase (DCP4010-81) no genoma da estirpe P4 sugere a presença do ciclo do glioxilato, o qual consiste de um desvio do TCA, onde a enzima isocitrato liase (ICL) catalisa a clivagem do isocitrato em succinato e glioxilato. A reação seguinte consiste na condensação do glioxilato com acetil-CoA para a formação de malato pela ação da enzima malato sintase (LEHNINGER *et al.*, 2007).

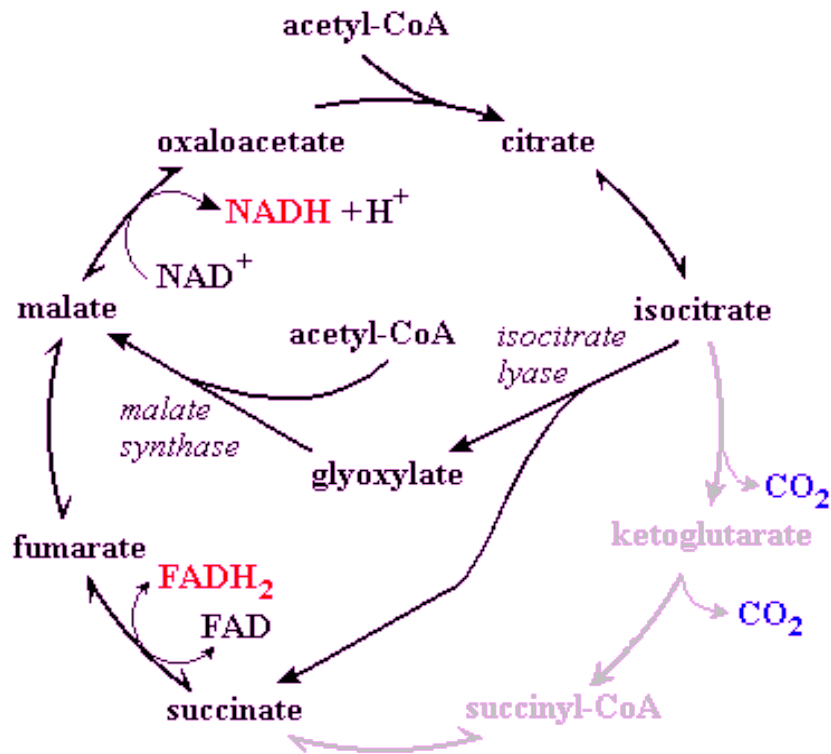


Figura 16: Esquema mostrando o ciclo do glioxilato (ciclo menor) sugerido pela anotação do genoma de *D. cinnamea* P4.

(<http://www.chemistry.uoguelph.ca/educmat/Chm452/lecture9.htm>).

A maioria dos genes prováveis que compõem as vias necessárias para a síntese de todos os 20 aminoácidos naturais foram identificados no genoma da estirpe P4, além de 49 genes tRNAs para todos os aminoácidos. Também foram anotados os genes hipotéticos para as vias biossintéticas dos aminoácidos D-alanina e selenocisteína. Os grupamentos amino necessários na síntese dos aminoácidos normalmente têm como fonte a amônia presente no ambiente, a qual pode ser obtida pela amônia permease (DCP40633) e pela via de assimilação da amônia, a qual emprega as enzimas glutamato desidrogenase, glutamato sintase, glutamato amônia liase e glutamina sintase (DCP40621-32). As vias de biossíntese dos nucleotídeos

purinas e pirimidinas são bastante complexas e longas. A síntese de purinas e pirimidinas envolve a construção átomo por átomo e são utilizadas diferentes fontes de carbono e nitrogênio. A anotação do genoma permitiu identificar muito dos genes hipotéticos envolvidos na síntese dos precursores dos cinco nucleotídeos.

Com a anotação do genoma foi sugerido ainda que a estirpe P4 tem a capacidade de sintetizar diversos cofatores e vitaminas envolvidas em funções de energia, respostas a estresses e síntese de metabólitos secundários (subitem 6.14). Foram identificados genes provavelmente envolvidos na biossíntese de piridoxina (DCP48030-35), tiamina (DCP48010-17), riboflavina (DCP48020-24), pantotenato e coenzima A (DCP48040-51), biotina (DCP48055-61), folato (DCP48065-70), porfirina, menaquinona, molibdopterina, lipoato, tiorredoxina, heme, quinolinato, cobalamina e do cofator NAD(P).

A síntese de ácidos graxos em *D. cinnamea* P4 é sugerida pela presença das enzimas β -cetoacil sintase I e III, β -cetoacil redutase, ácido graxo sintase, malonil CoA acil transferase, acetil CoA carboxilase subunidade alfa e beta, acetil CoA carboxilase subunidade biotina carboxilase (DCP40150-70). A síntese de ácidos graxos envolve a adição de dois átomos de carbono em cada etapa, com a adição de glicerol ao produto final da síntese de ácidos graxos, o qual pode ser suprido pela aquisição do meio externo através da provável glicerol permease (DCP3003, DCP43145) e por meio da via de síntese de glicerol. Esta última via provavelmente encontra-se presente em P4 devido a detecção dos prováveis genes responsáveis pela síntese da glicerol quinase, da glicerol 3-fosfato desidrogenase e da amido sintase (DCP40110-12).

A utilização pela estirpe P4 de ácidos graxos como fonte de carbono e energia é sugerida pela presença dos genes que compõem a via da β -oxidação. Na via da β -oxidação, após o ácido graxo ser ativado pela ligação de uma molécula de acetil-CoA pela enzima acil-Coa sintetase (DCP40157), uma série de reações cíclicas clivam dois carbonos do ácido graxo através das prováveis acil-CoA oxidase, acil-CoA desidrogenase, enoil-CoA hidratase, 3-hidroxi-acil-CoA desidrogeanse, acetil-CoA aciltransferase, acetil-CoA acetiltransferase e butiril-CoA desidrogenase, os quais serão empregados na produção de energia ou na construção de moléculas que necessitam de carbono como constituinte.

6.4 – Fosforilação oxidativa

A anotação do genoma de *D. cinnamea* P4 permitiu identificar os prováveis genes constituintes da cadeia de transporte de elétrons e do complexo ATP-sintase (Tabela 2). A transferência de elétrons através dos complexos inicia-se pelo complexo I ou NADH desidrogenase, o qual foi detectado em um provável *operon* contendo 12 genes (DCP40580-91) dos 14 descritos como constituintes do complexo I. Os genes hipotéticos não detectados durante a anotação foram *nuoAB*. O complexo II ou succinato desidrogenase é constituído de quatro subunidades, das quais duas foram identificadas: duas cópias da subunidade A (DCP40595, DCP40597) e duas cópias da subunidade B (DCP40596, DCP40559). Também foram identificados todos os prováveis genes do complexo III (DCP40557-57) ou citocromo bc1 e os prováveis genes do complexo IV ou citocromo c oxidase (DCP40520, DCP40540, DCP40560, DCP40598, DCP40599). O potencial eletroquímico gerado através do transporte de elétrons pela cadeia de complexos é empregado na síntese de ATP ao permitir o retorno dos H⁺ através da ATP-sintase. A presença da ATP-sintase é sugerida através das subunidades hipotéticas alfa (DCP40536), beta (DCP40538), delta (DCP40535), gama (DCP40537) e ATPase subunidade a (DCP40534).

Tabela 2: Sequências codificantes de proteínas de *D. cinnamea* P4 semelhantes a proteínas que compõem o sistema de fosforilação oxidativa.

Complexo	CDS	Função
Complexo I ou NADH desidrogenase	DCP40580-91	Oxidação de NADH e redução da quinona
Complexo II ou succinato desidrogenase	DCP40595, DCP40597, DCP40596, DCP40559	Oxidação de succinato a fumarato e redução da ubiquinona
Complexo III ou citocromo c redutase	DCP40557-57	Oxidação da ubiquinona e redução da citocromo c
Complexo IV ou citocromo c oxidase	DCP40520, DCP40540, DCP40560, DCP40598, DCP40599	Redução de O ₂ a H ₂ O
ATP-sintase	DCP40534-38	Síntese de ATP

6.5 – Aquisição de energia

6.5.1 Oxidação de monóxido de carbono

A anotação do genoma de *D. cinnamaea* P4 permitiu a identificação de genes hipotéticos para as enzimas *coxSLM* (DCP40504-06), as quais compõem o complexo denominado CODH, responsável pela oxidação de monóxido de carbono (CO) a dióxido de carbono (CO₂). Este complexo é composto por uma subunidade pequena CoxS (proteína ferro-enxofre), uma subunidade média CoxM (proteína FADH) e uma subunidade catalítica grande CoxL (proteína com centro molibdênio) (KING & WEBER, 2007). Apesar da disposição dos genes *coxSLM* variar entre as espécies oxidadoras de CO, a distribuição tradicional dos genes hipotéticos *coxSLM* é encontrada no genoma de *D. cinnamaea* P4 (SANTIAGO *et al.*, 1999; FUHRMANN *et al.*, 2003; PELZMANN *et al.*, 2009).

A literatura científica descreve dois tipos diferentes de CODHs: a forma I que oxida CO a CO₂ e a forma II com sua função ainda não totalmente compreendida (KING & WEBER, 2007). Desta forma, a sequência da provável subunidade maior L do complexo CODH presente na estirpe P4 foi comparada filogeneticamente com outras bactérias relacionadas, previamente descritas como oxidadoras de CO. Como apresentado na figura 17, a CDS CDP40505, proposta como a enzima CoxL presente no genoma de *D. cinnamaea* P4, mostrou maior proximidade filogenética às enzimas da forma I juntamente com a de outros actinomicetos relacionados como *Rhodococcus* e *Mycobacterium*.

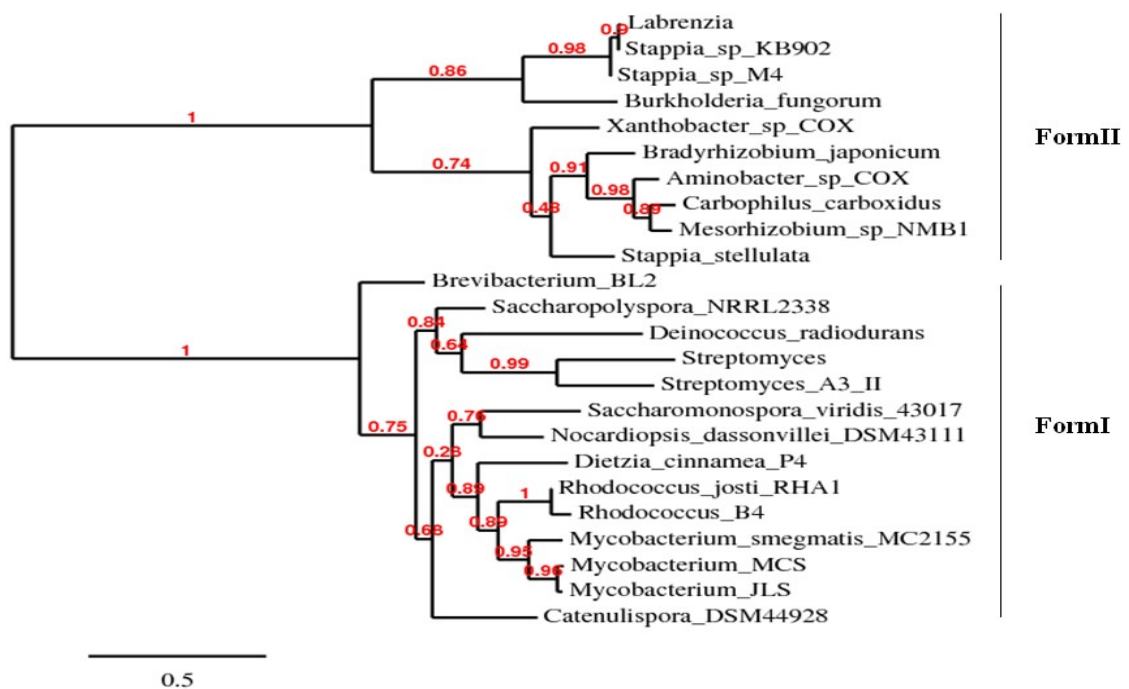


Figura 17: Análise filogenética da provável subunidade L de CODH (DCP40505) de *D. cinnamea* P4 com as formas I e II da subunidade L de CODH de outras bactérias relacionadas.

Também foram identificados três prováveis cópias do gene acessório *coxE* (DCP4500-02) e o provável fator de maturação proteica *xdhC/coxF* (DCP405003). A presença de *coxE* e *xdhC/coxF*, juntamente com o complexo CODH, corrobora a proposição de SANTIAGO *et al.* (1999), onde acredita-se que é essencial a presença dos genes acessórios *cox* com o centro CuMoO_2 , presente no complexo CODH, para a completa oxidação de CO.

6.5.2 Redução dissimilatória de nitrato

A anotação do genoma da estirpe P4 sugeriu a possível utilização de compostos de nitrogênio como fonte de energia, através da presença dos prováveis genes codificantes do complexo Nar ligado à membrana (DCP40600-13). Este complexo é envolvido em um processo respiratório que emprega a redução de nitrato (NHO_3^-) a nitrito (NO_2^-) através da via chamada de redução dissimilatória de nitrato (MORENO-VIVIÁN *et al.*, 1999). As enzimas do sistema Nar são compostas por uma subunidade catalítica NarG (DCP40604), uma subunidade solúvel NarH (DCP40605) e uma subunidade ligada à membrana NarI (DCP40607). As subunidades NarGH, também chamadas de subunidades α e β , são ancoradas na face citoplasmática da membrana através da subunidade NarI ou subunidade γ . Próximo aos genes *narGHI* foi identificado também o gene hipotético *narJ* (DCP40606), ao qual é sugerido pela literatura científica uma participação na associação e na estabilidade do complexo Nar (DUBOURDIEU & DEMOSS, 1992). Em outra localização no genoma de *D. cinnamea* P4 foi identificado o provável gene *napA* (DCP40636), o qual codifica a enzima dissimilatória nitrato redutase com função similar ao sistema Nar e localizada no periplasma. Entretanto, não foram identificados outros genes hipotéticos participantes deste complexo enzimático. Também foram localizadas três cópias do provável gene *moaA* (DCP40600-01, DCP40608), duas cópias de *moeA* (DCP40609, DCP40611) e uma única cópia do provável gene *moaC* (DCP40610). Os produtos da tradução destes genes são proteínas cofatoras contendo o centro de molibdênio que possuem papel indispensável na atividade catalítica da redutase NarG.

A regulação da expressão do *operon* Nar em *D. cinnamea* P4 provavelmente emprega o sistema de dois-componentes NarL e NarP (DARWIN & STEWART, 1995), o qual foi localizado no genoma através da presença do hipotético gene *narL* (DCP40614). Próximo a *narL* foi identificado um segundo gene provável que codifica uma proteína transdutora de

sinal histidina-quinase (DCP40635). Alguns relatos na literatura indicam que a proteína NarL é ativada através de sua fosforilação pela proteína histidina-quinase NarX. Entretanto, o papel exato da DCP40635 sobre a provável DCP40614 não foi possível ainda ser determinado apesar de estarem localizadas próximas uma da outra no genoma da estirpe P4, sugerindo um possível papel na regulação sobre o *operon* Nar. Devido à localização citoplasmática do complexo Nar, o nitrato tem que ser transportado para o interior da célula e o excesso de nitrito excretado. Dois genes prováveis codificando o transportador NarK foram localizados em diferentes pontos do genoma (DCP40602, DCP40612). Este transportador possui um papel de antiporte, o qual capta nitrato do meio externo em contrapartida da eliminação de nitrito (Figura 18).

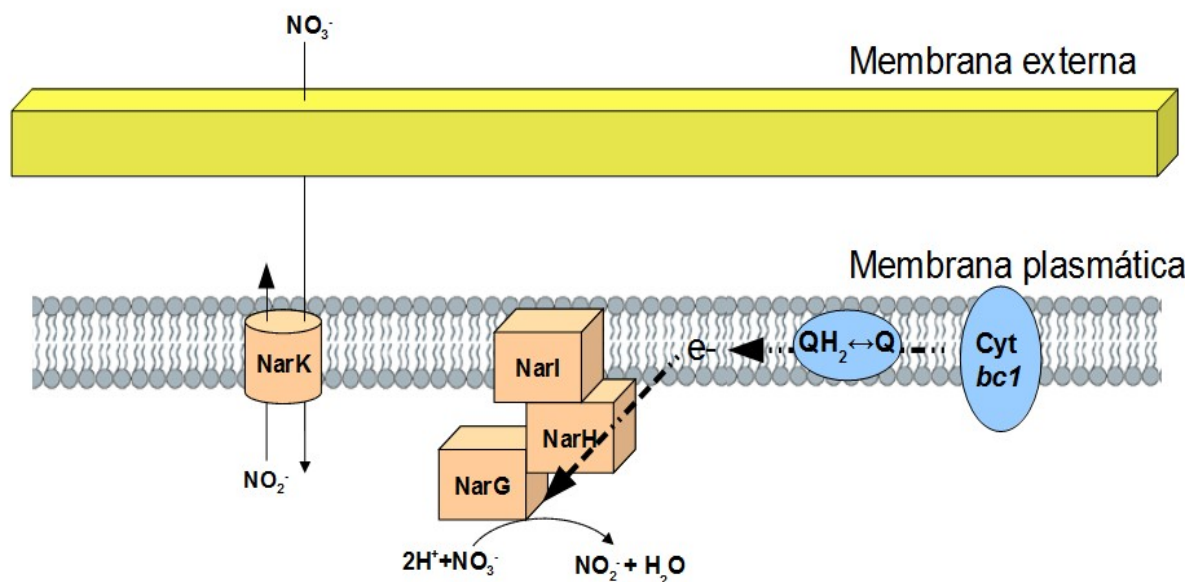


Figura 18: Esquema representativo do complexo Nar de *D. cinnamea* P4 empregado na obtenção de energia utilizando nitrato como aceptor final na cadeia de transporte de elétrons (MORENO-VIVIÁN *et al.*, 1999).

6.6 – Fixação autotrófica de carbono

6.6.1 Oxidação de monóxido de carbono (CO)

A energia conservada na oxidação aeróbia de CO pode ser direcionada para a fixação de CO₂ e, conseqüentemente para a sua conversão em biomassa para a célula. A análise do genoma da estirpe P4 sugere a existência de dois mecanismos possíveis para a fixação de CO₂: o ciclo do ácido cítrico reverso (rTCA) e a via do 3-hidroxiopropionato.

Como apresentado da figura 19, muitas das enzimas presentes no ciclo do ácido cítrico operam de maneira reversa, com a exceção da reação catalisada pela enzima citrato sintase

que é substituída pela reação da citrato liase no ciclo do rTCA (DCP41210, DCP41213). As outras enzimas que atuam no ciclo do rTCA são: (i) o complexo multi enzimático ferredoxina oxidoreductase (DCP41211), o qual participa em duas reações: carboxilação do succinil-CoA a α -cetoglutarato e a carboxilação de acetil-CoA a piruvato, (ii) a enzima fosfoenol-piruvato sintase (DCP41212) e (iii) a fosfoenol-piruvato carboxilase (DCP41214), responsável pela conversão de piruvato a gliceraldeído 3-fosfato, um componente da via glicolítica. A presença dos prováveis genes que codificam estas enzimas é um indicativo da existência deste mecanismo de fixação de CO_2 no genoma de *D. cinnamea* P4.

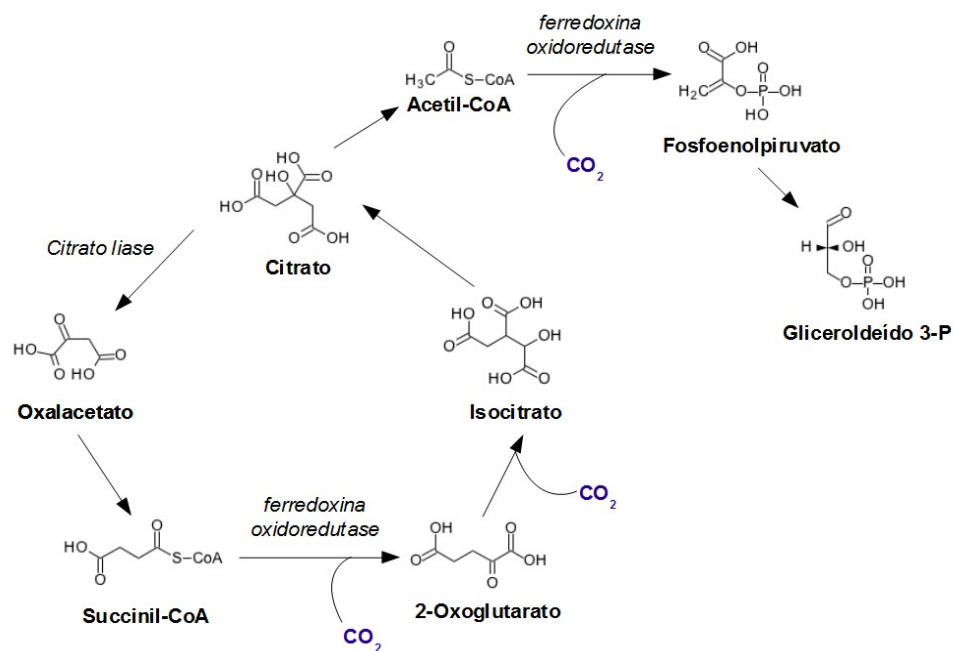


Figura 19: O ciclo do ácido cítrico reverso sugerido no genoma da estirpe P4 de *D. cinnamea* (ROTHSCHILD, 2008).

6.6.2 O ciclo do 3-hidroxi-propionato

Em *D. cinnamea* P4 a anotação genômica identificou os prováveis genes codificantes das duas subunidades alfa e beta da enzima acetil/propionil-CoA carboxilase dependente de biotina (DCP40516-17), a qual realiza duas reações de carboxilação entre acetil-CoA e CO_2 e propionil-CoA e CO_2 resultando em metilmalonil-CoA. Também foi identificada a enzima hipotética malonil-CoA redutase (DCP40515), responsável pelo rearranjo do composto metilmalonil-CoA, que é regenerado em acetil-CoA ou gera glioxilato. Esse último composto é direcionado para o ciclo de glioxilato descrito acima para ser utilizado como fonte de

energia e biomassa pela célula (Figura 20).

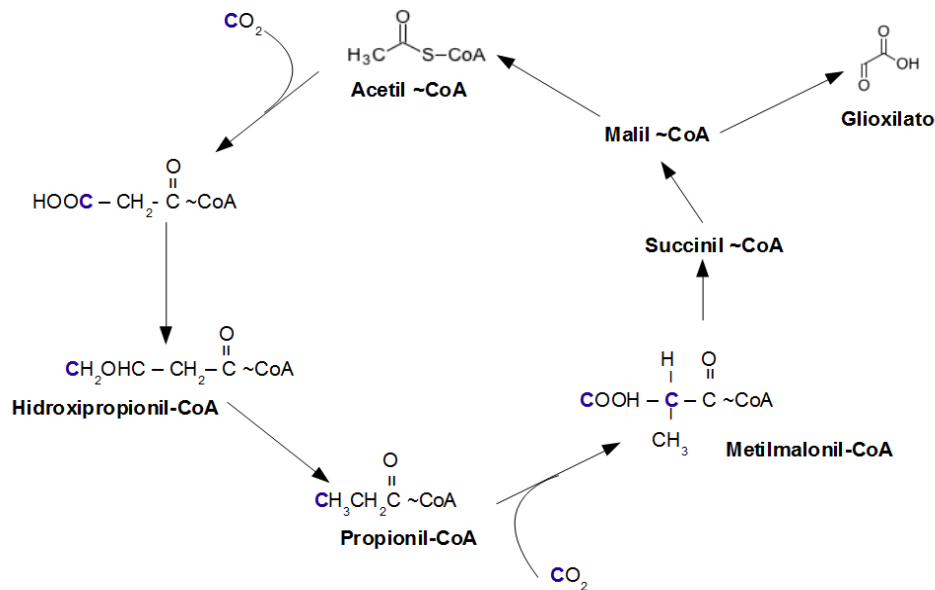


Figura 20: Ciclo do 3-hidroxiacetil-CoA proposto pela análise do genoma da estirpe P4. Duas moléculas de CO_2 são condensadas com uma molécula de Acetil-CoA, produzindo glioxalato (HERTER *et al.*, 2001).

6.6.3 Assimilação de composto de um carbono (C1)

Outros vários e prováveis genes envolvidos na assimilação de compostos C1 (Figura 21) foram detectados durante a anotação do genoma de P4. Embora *D. cinnamea* P4 não apresente os prováveis genes necessários para a oxidação de metano, muitos dos genes hipotéticos para a oxidação de formato, transferência de formaldeído e finalmente a assimilação de C1 pela célula estão presentes no seu genoma. A anotação do genoma da estirpe P4 identificou duas possíveis vias para a oxidação do formato. Na primeira via sugerida, os compostos de formato disponíveis no meio são transferidos para o interior da célula através de um transportador do tipo ABC (DCP43169), sendo subsequentemente oxidado a CO_2 pelo complexo formato desidrogenase, o qual é constituído por três subunidades FdhABC mais a proteína acessória FdhD. A análise do genoma de *D. cinnamea* P4 permitiu identificar um grupamento gênico constituído pelos prováveis genes *fdhABD* (DCP40575-77), mas não foi possível localizar o genes *fdhC*. O segundo mecanismo para a oxidação de formato emprega como substrato inicial o composto formaldeído. Em uma primeira reação o formaldeído é ligado ao composto tetrahydrofolato (H_4F) formando

metileno-H₄F, pela ação da formil-H₄F ligase. A via de síntese de H₄F envolve a ação de seis enzimas, das quais três foram detectadas no genoma da estirpe P4: as putativas dihidropteroato sintase, *folP* (DCP40700), dihidrofolato sintase, *folC* (DCP40701) e dihidrofolato redutase, *dys* (DCP40702). O composto metileno-H₄F é convertido a formato pela ação da enzima bifuncional 5,10-metileno-tetrahydrofolato desidrogenase/Metenil tetrahydrofolato ciclohidrolase, *folD* (DCP40566) e, em uma segunda reação subsequente, o formato é oxidado a CO₂ pelo mesmo complexo formato desidrogenase descrito anteriormente.

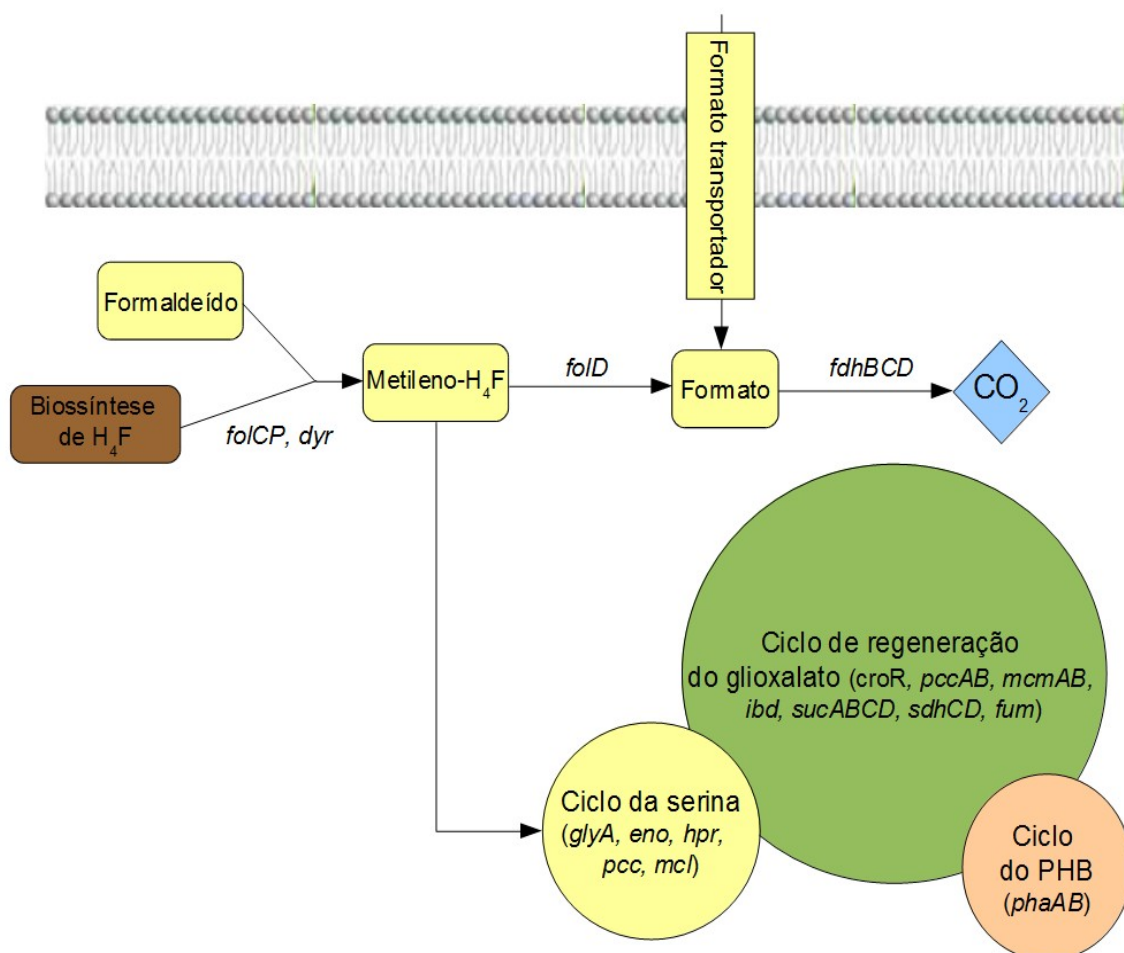


Figura 21: Esquema representativo das prováveis vias envolvidas na assimilação de composto de um carbono (C1) presente no genoma de *D. cinnamea* P4 (CHISTOSERDOVA *et al.*, 2004).

Compostos de formaldeído também podem ser assimilados como constituintes celulares através da participação de três vias metabólicas: o ciclo da serina, o ciclo de

regeneração do glioxilato e o ciclo poli- β -hidroxisobutilil-CoA (PHB) (CHISTOSERDOVA *et al.*, 2004). Em *D. cinnamea* P4, a anotação do seu genoma indica a provável assimilação de formaldeído a partir do composto metileno- H_4F , que entra no ciclo da serina através da enzima serina glioxilato aminotransferase. Entretanto, essa última enzima não foi detectada no genoma da estirpe P4. Outras prováveis enzimas identificadas como participantes deste ciclo são: serina hidroximetiltransferase, *glyA* (DCP40705); enolase, *eno* (DCP4003); hidroxipiruvato redutase, *hpr* (DCP40704); fosfoenolpiruvato carboxilase, *ppc* (DCP40706); malil-CoA liase, *mcl* (DCP40703). O produto gerado pelo ciclo da serina é direcionado para o ciclo de regeneração do glioxalato, o qual envolve um elaborado conjunto de reações procedendo a partir de derivados de ácidos carboxílicos com C3, C4 e C5 ligados a CoA. Incluem ainda no ciclo de regeneração do glioxalato, reações de carboxilações e descarboxilações e a participação de algumas enzimas do ciclo do TCA (KOROTKOVA & LIDSTROM, 2004). Como em outros casos, vários processos e a função de seus respectivos genes continuam ainda sem entendimento. Entre as enzimas que participam do ciclo de regeneração do glioxilato foram identificados as prováveis enzimas crotonase, *croR* (DCP40715); propionil-CoA carboxilase, *pccAB* (DCP40716); metilmalonil-CoA mutase, *mcmAB* (DCP40717-18); isobutilil-CoA desidrogenase, *ibd* (DCP40719); succinil-CoA sintetase, *sucABCD* (DCP40196, DCP40194, DCP40185-86); succinato desidrogenase, *sdhCD* (DCP40192-93) e fumarato, *fum* (DCP40198). O ciclo PHB é intimamente ligado ao ciclo de regeneração do glioxilato, sendo a primeira reação que ocorre em PHB (realizada pelas enzimas β -cetotiolase e acetoacetil-CoA redutase) também a primeira no ciclo de regeneração do glioxilato (CHISTOSERDOVA *et al.*, 2004). A anotação do genoma de *D. cinnamea* P4 possibilitou identificar as prováveis enzimas β -cetotiolase, *phaA* (DCP40710) e acetoacetil-CoA redutase, *phaB* (DCP40711) que fazem parte do ciclo metabólico PHB.

6.7 – Metabolismo de hidrocarbonetos

6.7.1 Hidrocarbonetos alifáticos

A anotação do genoma estudado confirmou a presença de prováveis genes envolvidos no metabolismo de alcanos e alguns hidrocarbonetos aromáticos (Figura 22a). Foi identificado um *operon* no genoma da estirpe P4 contendo os genes hipotéticos alcano 1-monoxigenase-rubredoxina A (DCP41333), um transportador de lipídeos (DCP41332) e um regulador transcricional da família TetR (DCP41331) (Figura 22c). O gene alcano 1-monoxigenase-rubredoxina A da estirpe P4 apresentou similaridade de 96% com *Dietzia* sp. E1. A configuração proteína fusionada, alcano 1-monoxigenase-rubredoxina A não é

comumente descrita em enzimas pertencentes a via Alk. Como a localização das duas enzimas é diferente, (transmembrânica e citoplasmática, respectivamente), uma provável clivagem proteolítica deve ocorrer durante a maturação e/ou transporte pós-traducional de ambas enzimas. A alceno 1-monoxigenase, que é responsável pela incorporação do átomo de oxigênio na molécula de alceno, é uma proteína transmembrânica com seis trechos de sua sequência atravessando a membrana plasmática. Dados da análise da CDS DCP41333 no banco de dados THMMH confirmam a mesma configuração desta com a apresentada pela proteína de *Dietzia* sp. E1. A proteína citoplasmática rubredoxina A (RubA) é responsável pela transferência de elétrons de NADH para a reação. As comparações de similaridades feitas entre o trecho da sequência da CDS que codifica a enzima RubA e sequências de RubA depositadas no banco de dados NR de organismos relacionados confirmam a porção RubA da CDS DCP41333 de *D. cinnamea* P4 como provável enzima RubA. A regulação da via Alk é realizada pelo regulador AlkS, o qual se liga no promotor próximo ao operon Alk, induzindo a transcrição dos genes que compõem a via metabólica (van BEILEN & WITHOLT, 2004). A análise filogenética entre a CDS DCP41331, que apresentou similaridade com regulador da família TetR, e sequências de aminoácidos de outros reguladores AlkS de organismos relacionados sugere um provável papel de DCP41331 na regulação da via metabólica responsável em degradar compostos de alcanos.

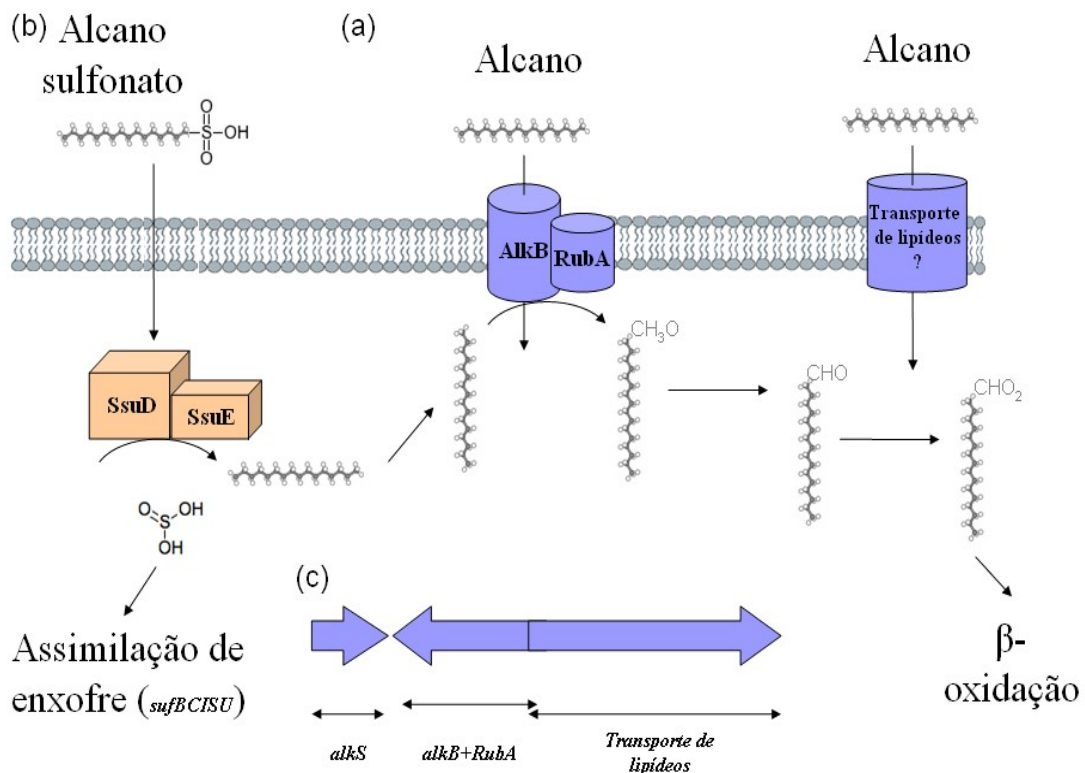


Figura 22: Vias metabólicas propostas para a estirpe *D. cinnamea* P4. (a) Via metabólica de compostos de hidrocarbonetos alifáticos. O complexo AlkB e RubA são responsáveis pela primeira reação de hidroxilação de compostos de *n*-alcanos, que após ser convertido a um ácido, e.g. ácido hexadecanóico, pela via Alk é direcionado para a via central de β -oxidação. (b) Via de assimilação de compostos de alcanosulfonatos. Após a clivagem de alcanosulfonatos em *n*-alcano e sulfito, os dois são assimilados pela via Alk e a pela via de assimilação de enxofre, respectivamente. (c) Esquema representativo da disposição dos genes envolvidos na assimilação de *n*-alcanos (van BEILEN *et al.*, 2002b).

Também foram localizados genes putativos envolvidos na assimilação de compostos de alcanosulfonatos como fontes de enxofre e carbono/energia pela célula. Os alcanosulfonatos são compostos de *n*-alcanos contendo uma molécula de enxofre em sua posição terminal, sendo amplamente distribuídos na natureza, pois são derivados de material vegetal em decomposição. Foram localizados também no genoma da estirpe P4 os prováveis genes constituintes do *operon ssuED* (DCP40672, DCP40672, DCP40660), cuja função proposta pela literatura científica é o papel na assimilação de sulfonatos alifáticos. O gene *ssuD* codifica uma monoxigenase que catalisa uma reação oxigenolítica, onde um átomo de oxigênio é incorporado na posição terminal do alcanosulfonato. Ao mesmo tempo, é realizado uma clivagem C-S do alcanofulfonato com a participação de SsuE na transferência de elétrons de FADH. Os produtos gerados são R-CHO, H⁺ e sulfito (SO₃²⁻) (Figura 22b). O sulfito pode ser ainda direcionado para a via de assimilação de enxofre (ver subitem 6.12.1). Estudos sugerem que a molécula de alcano liberado após a reação de clivagem catalisada pelas enzimas SsuD e SsuE é direcionado para a via Alk (REICHENBECHER & MURRELL, 1999).

6.7.2 Hidrocarbonetos aromáticos

A anotação do genoma de *D. cinnamea* P4 possibilitou a observação da presença da via catabólica de degradação do hidrocarboneto aromático bifenil. Os hidrocarbonetos aromáticos podem conter um anel aromático, denominados de monoaromáticos, cujos exemplos são os compostos de tolueno, benzeno, xileno e etilbenzeno. Podem também conter dois ou mais anéis aromáticos, denominados de poliaromáticos ou policíclicos aromáticos, tendo como exemplos os compostos de naftaleno, antraceno e pireno. Embora os genes que codificam a enzima responsável pela primeira reação de inserção de dois átomos de oxigênio, a multi subunidade bifenil dioxigenase (*bphA1A2A3A4*), não tenham sido localizados, foram

identificados os genes dihidrodeol desidrogenase, *bphB* (DCP41306), 2,3- dihidroxibifenil 1,2-dioxigenase, *bphC* (DCP41311, DCP41327) e 2-hidroxi-6-oxo-6-fenilhexa-2,4-dienoato hidrolase, *bphD* (DCP41326), que fazem parte da via superior bifenil (Figura 23a). Os produtos gerados pela via superior bifenil são 2-hidroxi-penta-2,4-dienoato e benzoato. O primeiro produto é direcionado para a degradação de benzoato via ligação de CoA, onde é convertido a piruvato e posteriormente a succinato pelas enzimas codificadas pelos genes *bphE* (DCP41308), *bphF* (DCP41310), *bphG* (DCP41309), *catF* (DCP41322) e *catIJ* (DCP41320-21) (Figura 23b). O segundo produto, o benzoato, é direcionado para a degradação de benzoato via hidroxilação ou via central β -ceto adipato, composta pelos genes benzeno dioxigenase, *benAB* (DCP41315-16); 2-hidro 1,2-dihidroxibenzoato desidrogenase, *benD* (DCP41302); catecol 1,2-dioxigenase, *catA* (DCP41318); muconato cicloisomerase, *catB* (DCP41319); β -ceto adipato enol-lactona hidrolase, *catD* (DCP41324-25) e β -ceto adipil coenzima A tiolase, *catF* (DCP41322) (Figura 23c). Não foi possível identificar o gene muconolactona isomerase, *catC*. Outro possível metabólito resultante da via superior bifenil é o 4-hidroxi-benzoato, que é convertido a protocatecuato pela enzima 4,5-protocatecuato dioxigenase (DCP41312) e, em seguida, direcionado para a via central β -ceto adipato (Figura 23d). Bactérias pertencentes à classe Betaproteobactéria e ao grupo Actinobacteria são as principais representantes dos microrganismos capazes de utilizarem compostos de bifenil como única fonte de carbono em condições de aerobiose (WILLIAMS, LOBOS & CHEETHAM, 1997).

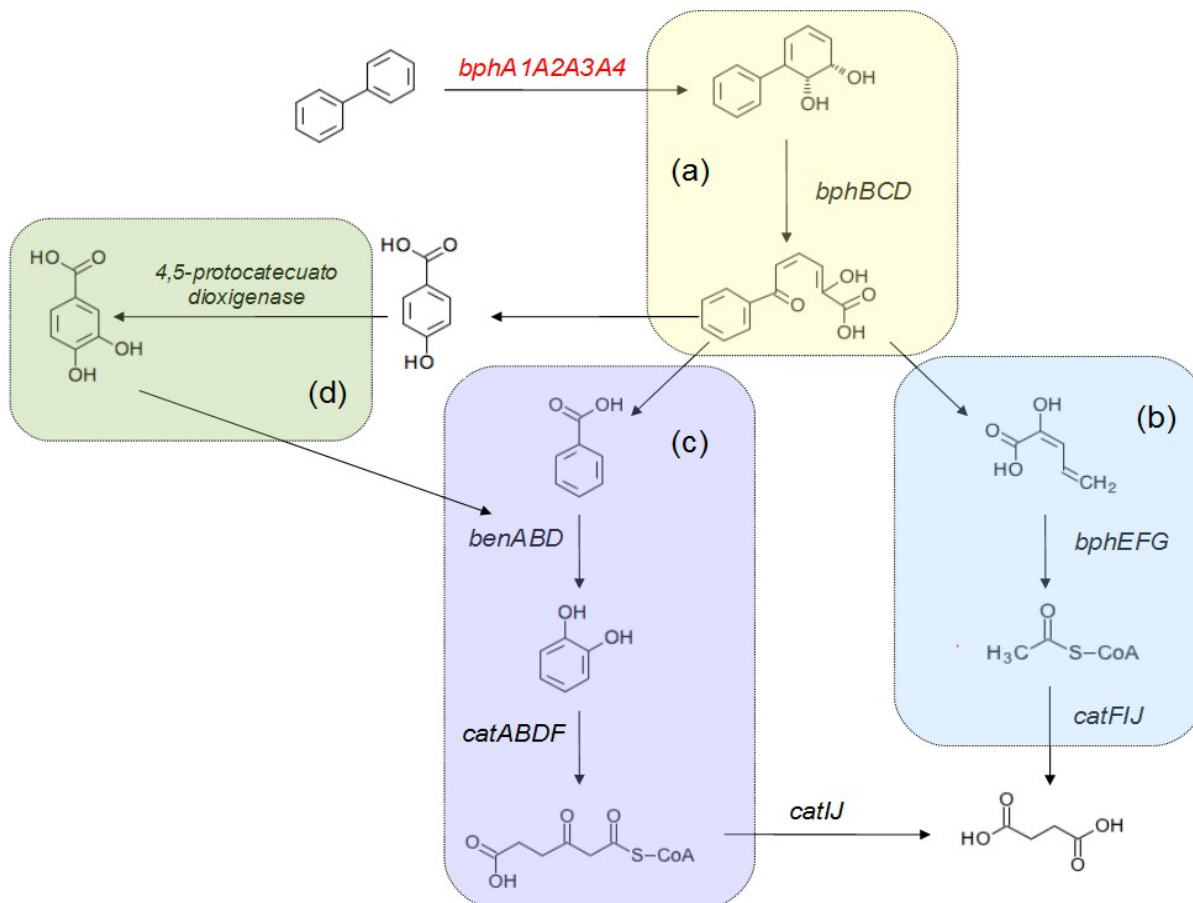


Figura 23: A via de degradação de compostos de bifenil proposta pela anotação dos prováveis genes presentes no genoma de *D. cinnamea* P4. (a) Via superior bifênil, (b) via de degradação de benzeno via ligação de CoA, (c) via central β -cetoadipato e (d) via de degradação de protocatecuato (van HAMME, SINGH & WARD, 2003).

A anotação do genoma da estirpe P4 revelou também prováveis genes necessários para a utilização de compostos de fenol, como única fonte de carbono. Os compostos de fenol, como os de bifenil são danosos ao meio ambiente devido a sua persistência no meio e sua toxicidade. A reação catalisada pela enzima fenol monoxigenase (DCP41328) converte fenol em catecol, o qual pode ser direcionado para a via central β -cetoadipato descrita no parágrafo anterior. Também foi identificada uma provável 4-hidroxiacetofenona monoxigenase (DCP41329). Esta enzima participa na reação de oxigenação de compostos de 4-hidroxiacetofenona resultando em 4-hidroxiifenil acetato (KAMERBEEK *et al.*, 2001) (Figura 24).

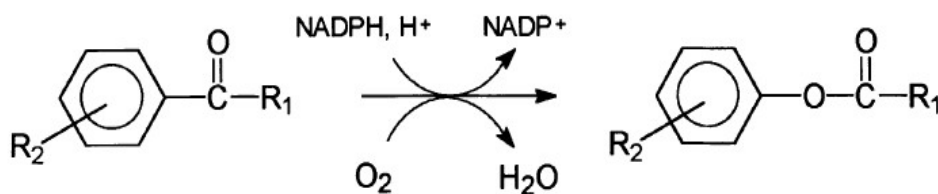


Figura 24: Esquema representativo da reação de hidroxilação de compostos de 4-hidroxiacetofenona (KAMERBEEK *et al.*, 2001)

6.7.3 Hidrocarbonetos cíclicos

Um gene hipotético que codifica a enzima ciclohexanona monoxigenase (DCP41330) foi identificado no genoma da estirpe P4. A via de degradação de hidrocarbonetos cíclicos é composta por seis enzimas responsáveis pela conversão de ciclohexanona a ácido adípico (CHENG *et al.*, 2000). Em *D. cinnamea* P4, a primeira reação é a de hidroxilação, a qual é catalisada pela enzima ciclohexanona monoxigenase. As reações subsequentes são catalisadas por álcool desidrogenases, hidrolases e uma provável 6-oxohexanóico ácido desidrogenase. Exceto a 6-oxohexanóico ácido desidrogenase, todas as enzimas hipotéticas foram identificadas no genoma da estirpe P4.

6.8 – Biodegradação de compostos nitrogenados

6.8.1 Nitrila

Possíveis genes que estejam envolvidos na degradação de compostos de nitrila foram localizados no genoma da estirpe P4. Os genes denominados de nitrilase/cianida hidratase estão presentes em três cópias no genoma (DCP41901-03), e uma amidase com cinco cópias presentes no genoma (DCP41904-09). As reações catalisadas por estas enzimas consistem em uma hidratação, $R-CN + H_2O \rightarrow R-CONH_2$, seguida de uma reação de clivagem, $R-CONH_2 + H_2O \rightarrow R-CONH + NH_3$, produzindo um ácido e amônia.

As enzimas capazes de metabolizar compostos de nitrilas, nitrilases e nitrilas hidratases fazem parte de um grupo diverso, composto por 13 subgrupos encontrados em três famílias de plantas, em um pequeno número de fungos e em vários gêneros de bactérias como, por exemplo, *Acinetobacter*, *Corynebacterium*, *Arthrobacter*, *Pseudomonas*, *Klebsiella*, *Nocardia* e *Rhodococcus* (PACE & BRENNER, 2001). Já as nitrilas hidratases podem ser classificadas em três subgrupos baseados em seu cofator, que pode conter Fe³⁺,

Co³⁺ ou um dos três metais: cobalto, cobre ou zinco (MARTÍNKOVÁ *et al.*, 2009). Análises filogenéticas realizadas com as CDSs DCP41902 e DCP41903 mostraram a similaridade entre estas sequências e outras sequências de nitrila hidratase dos grupos que contém Fe³⁺, Co³⁺ como cofator. Já a CDS DCP41901 mostrou maior similaridade com outro grupo membro da subfamília 3; entretanto, não foi possível realizar sua classificação em relação ao seu cofator (Figura 25).

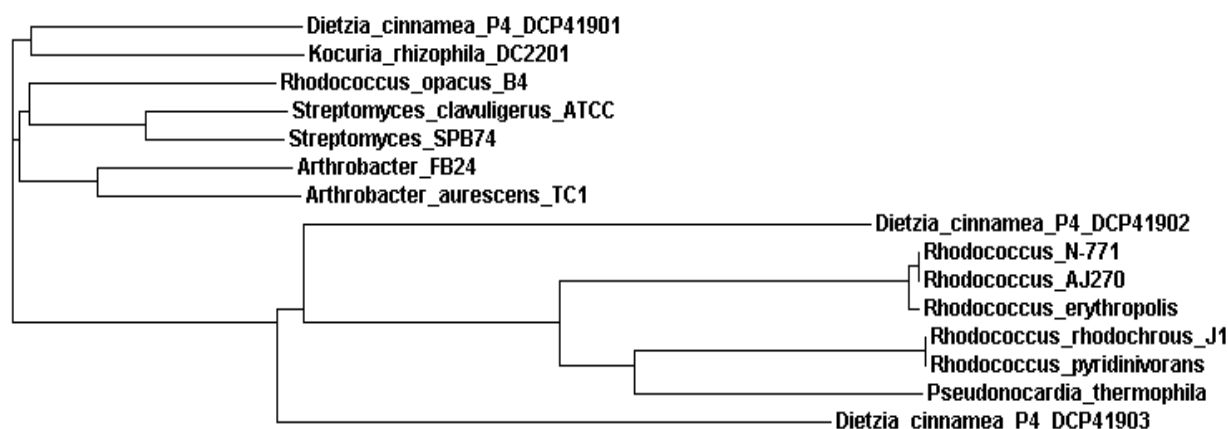


Figura 25: Análise filogenética entre as sequências de aminoácidos DCP41901-03 de *D. cinnamea* P4 e sequências da enzima nitrila hidratase de organismos relacionados.

6.8.2 Nitriloacetato

Também foram identificados genes hipotéticos envolvidos na degradação de compostos de nitriloacetato (NTA). A reação proposta para a degradação de nitriloacetatos em *D. cinnamea* P4 seria realizada pelos genes hipotéticos que formam o complexo enzimático composto por uma nitriloacetato monoxigenase (DCP41911) e uma nicotinamida adenina dinucleotídeo (NADH):FMN oxidoreductase (DCP41910), responsável em fornecer FMNH₂ para a oxidação de NTA utilizando NADH para reduzir FMN a FMNH₂ (KIM *et al.*, 2006), resultando em glioxilato e iminodiacetato (IDA) como produtos finais (Figura 26). O glioxilato é direcionado para o ciclo de glioxilato para ser utilizado como fonte de carbono, enquanto o iminodiacetato é convertido em glicina e glioxilato novamente pela ação da iminodiacetato desidrogenase. Essa última enzima não foi identificada no genoma da estirpe P4. A glicina é então convertida à serina pela enzima serina hidroximetil transferase (DCP41912), entrando em seguida na via do metabolismo da serina até a conversão de 3-glicerato bifosfato, pela ação da glicerato quinase (DCP41913).

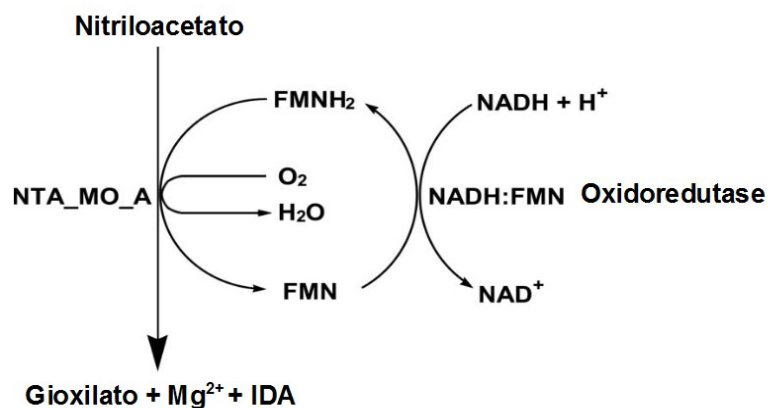


Figura 26: Esquema proposto para a reação de hidroxilação de nitriloacetato (adaptado de KIM *et al.*, 2006).

6.8.3 Nitroalcano

Com a anotação do genoma de *D. cinnamea* P4 foi observada a presença de cinco CDSs prováveis codificantes da enzima 2-nitropropano dioxigenase (DCP41915-19), pertencente à classe das nitroalcanos oxigenases. A flavoproteína nitropropano dioxigenase (NAO) realiza a reação de inserção de dois átomos de oxigênio no composto de nitroalcano produzindo cetonas ou aldeídos mais nitrito e peróxido de hidrogênio (Figura 27). A tabela 3 apresenta os resultados da busca de similaridades entre as CDSs DCP41915-19 e as sequências de organismos relacionados. Embora todas as CDSs apresentem alto valor de similaridade com o gene codificante para a 2-nitropropano dioxigenase de actinomicetos relacionados, o alinhamento entre as sequências DCP41915-19 não sugere uma relação filogenética entre as cinco diferentes CDS de *D. cinnamea* P4 que apresentaram similaridades com a enzima nitropropano dioxigenase (Figura 28).



Figura 27: Reação catalisada pela enzima nitroalcano dioxigenase. (do lado esquerdo) Substrato da reação, nitroalcano, oxigênio molecular e água. (do lado direito) Produtos da reação, aldeído ou cetona, peróxido de hidrogênio, nitrito e íons.

Tabela 3: Resultado da análise de similaridade das sequências deduzidas de aminoácidos DCP41915-19, comparadas às sequências depositadas no banco NR, utilizando a ferramenta BLASTP (<http://www.ncbi.nlm.nih.gov/BLAST/>).

CDS	Gene provável	E-value	Identidade / Cobertura ¹	Organismo
DCP41915	2-nitropropano dioxigenase	7,00E-150	78% / 100%	<i>Rhodococcus equi</i> ATCC 33707
DCP41916	2-nitropropano dioxigenase NPD	6,00E-165	80% / 100%	<i>Rhodococcus erythopolis</i> SK121
DCP41917	2-nitropropano dioxigenase NPD	1,00E-063	45% / 100%	<i>Geobacillus thermoglucosidasius</i> C56-YS93
DCP41918	2-nitropropano dioxigenase	7,00E-150	78% / 99%	<i>Rhodococcus equi</i> ATCC 33707
DCP41919	2-nitropropano dioxigenase	4,00E-125	68% / 98%	<i>Alcanivorax</i> sp. DG881

¹Extensão da sobreposição da sequência analisada com as sequências similares de outros organismos depositadas nos bancos de dados.

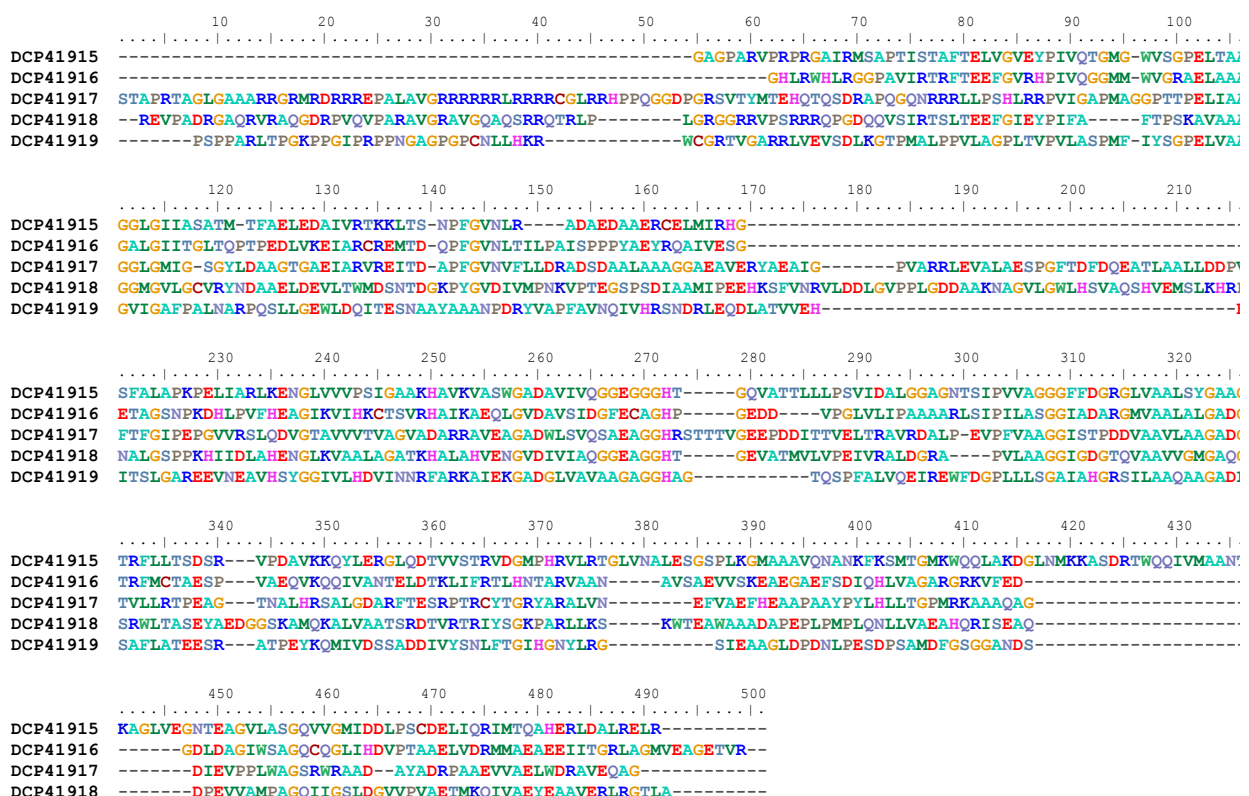


Figura 28: Alinhamento entre as sequências deduzidas de aminoácidos das CDSs DCP41915-19 de *D. cinnamea* P4, utilizando o programa CLUSTALW disponível no BioEdit. Os resultados mostram uma baixa similaridade entre as cinco sequências.

6.9 – Degradação de colesterol

Foi identificado no genoma de *D. cinnamea* P4 um provável *operon* contendo os genes hipotéticos 3-cetoesteróide 9 α -hidroxilase componente KshA, 3-cetoesteróide 9 α -hidroxilase componente KshB e 3-cetoesteróide 1 Δ -desidrogenase componente KstD. Além disso, em outro local do genoma, foi detectado uma segunda cópia do gene putativo 3-cetoesteróide 9 α -hidroxilase componente KshB (DCP41291-93, DCP41290). Os resultados da busca de similaridade entre as CDSs DCP41290-93 e as sequências de organismos relacionados depositados no banco NR foram bastante expressivos, com todos os valores de similaridade e cobertura acima de 73%. KshAB e KstD são enzimas chave na via catabólica de colesterol, onde catalisam a inserção de um átomo de oxigênio no anel B da molécula de 3-cetoesteróide e a abertura do anel A. O sequenciamento do genoma da estirpe P4 também revelou a presença de duas cópias do gene hipotético extradiol dioxigenase HsaC (DCP41298-99), também envolvidas na degradação de compostos de colesterol. A provável extradiol dioxigenase HsaC atua abrindo o anel B, com a concomitante aromatização do anel A. Em seguida, o produto 3,4-dihidroxi-9,10-seco-nandroste-1,3,5(10)-trieno-9,17-diona (3,4-DHSA) sofre uma metaclivagem pela ação da extradiol dioxigenase HsaC, resultando em ácido 4,5-9,10-diseco-3-hidroxi-5,9,17-trioxoandrosta-1(10),2-dieno-4-óico (4,9-DSHA) (Figura 29) (LARKIN, KULAKOV & ALLEN, 2005; McLEOD *et al.*, 2006). Uma enzima HsaD, também envolvida na via catabólica de colesterol, é descrita na literatura com a identificação de BhpD (van der GEIZE *et al.*, 2007), Esta foi anotada no genoma de *D. cinnamea* como participante da via de degradação de bifênil (subitem 6.7.2). As prováveis enzimas HsaEFG não foram identificadas no genoma de *D. cinnamea* P4.

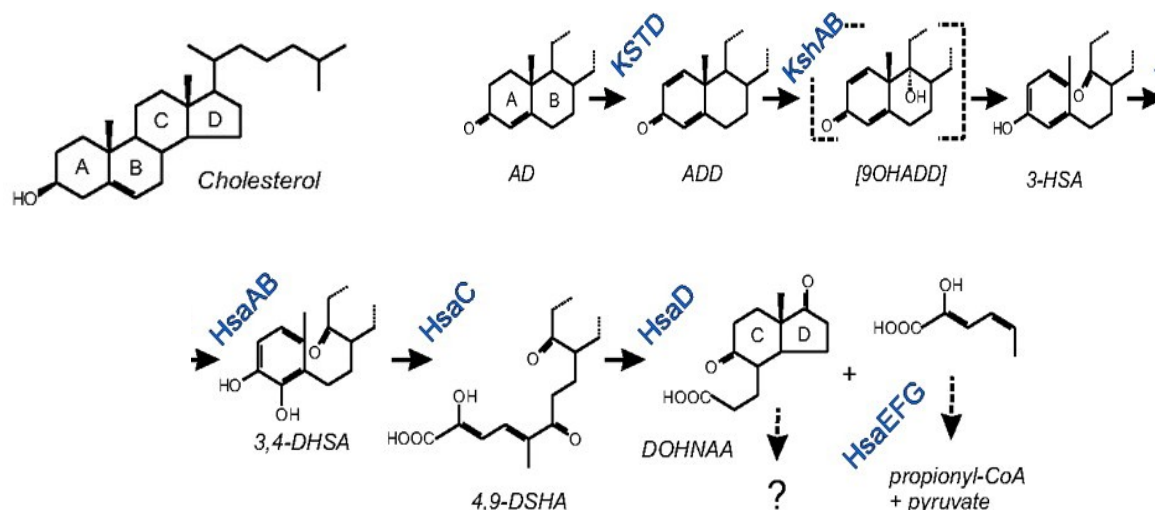


Figura 29: Via de degradação do colesterol proposta para *D. cinnamea* P4 (adaptado de van der GEIZE *et al.*, 2007).

Dois genes putativos que codificam a enzima colesterol oxidase foram detectados em diferentes locais do genoma de *D. cinnamea* P4 (DCP41295-96). A enzima colesterol oxidase é uma enzima di-nucleotídeo adenina flavina dependente de FAD que catalisa a oxidação de colesterol, resultando nos produtos peróxido de hidrogênio e o composto *colest-4-en-3-one* (CEO), um importante intermediário em muitas transformações de esteroides (DOUKYU, 2009). A literatura científica descreve duas classes diferentes de colesterol oxidase, com duas subclasses cada uma: a classe I contendo o cofator FAD ligado de maneira não covalente à enzima e a classe II com o cofator ligado de maneira covalente (SAMPSON & VRIELINK, 2003). A análise filogenética com as duas sequências de aminoácidos deduzidas DCP41295-96 e outras sequências de aminoácidos de organismos relacionados revelaram que as mesmas são pertencentes à classe I de colesterol oxidases, mas pertencentes a subclasses distintas (Figura 30).

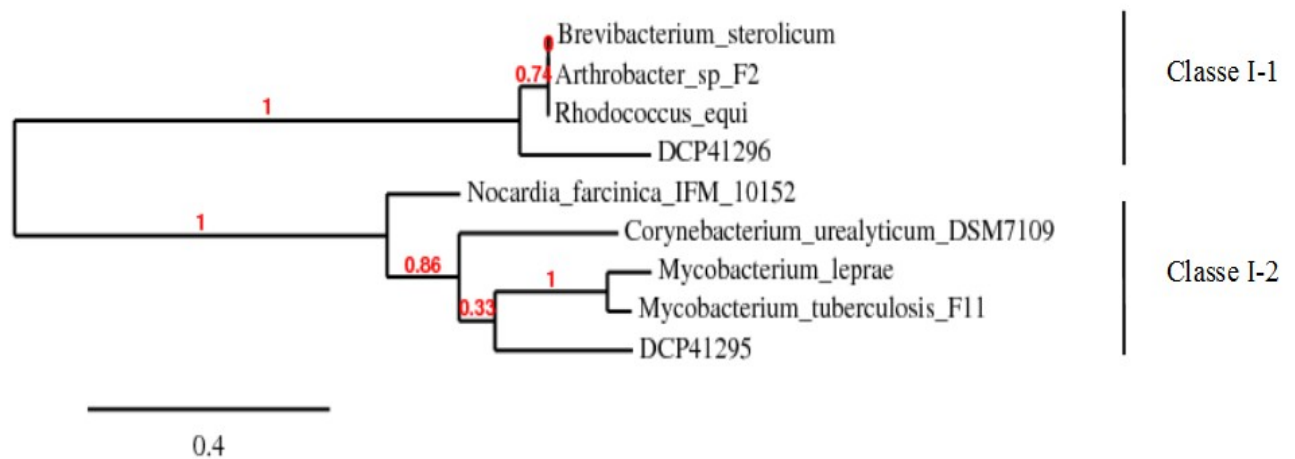


Figura 30: Análise filogenética entre as sequências deduzidas dos aminoácidos DCP41295 e DCP41296 de *D. cinnamea* P4 e as sequências de enzimas colesterol oxidase de organismos relacionados, utilizando o programa PhyML disponível no sítio http://www.phylogeny.fr/version2_cgi/simple_phylogeny.cgi.

6.10 – Oxigenases

O sequenciamento do genoma de *D. cinnamea* P4 permitiu a identificação de 45 e 19 genes prováveis que codificam monoxigenases e dioxigenases, respectivamente (Tabela 4, Anexo II). Sessenta e sete oxigenases hipotéticas presentes no genoma da estirpe P4 é um número relativamente alto quando se avalia a relação do número de oxigenases por Mb do genoma celular (Tabela 1). A classe monoxigenase é dividida em sete subclasses baseadas no cofator necessário para a catálise: (1) monoxigenases dependentes de heme, também referida com citocromo P450, (2) monoxigenases dependentes de flavina, (3) monoxigenases dependentes de cobre, (4) monoxigenases não-heme dependentes de ferro, (5) monoxigenases dependentes de pterina, (6) monoxigenases dependentes de outros cofatores e (7) monoxigenases independentes de cofator. Entre as 45 monoxigenases hipotéticas, nove pertencem à subclasse monoxigenases dependentes de heme, 36 à subclasse monoxigenases dependentes de flavina e apenas uma monoxigenases não-heme dependente de ferro (Tabela 4).

Tabela 4: Possíveis oxigenases identificadas no genoma de *D. cinnamea* P4.

CDS	Melhor índice de similaridade	Processo envolvido
DCP43000	Monoxigenase da família Luciferase	-
DCP43001	Monoxigenase da família Luciferase	-
DCP43002	Piridoxamina 5'-fosfato oxidase	Biossíntese da vitamina B6
DCP41328	Diclorofenol 6-monoxigenase	Metabolismo de hidrocarbonetos
DCP41329	4-hidroxi acetofenona monoxigenase	Metabolismo de hidrocarbonetos aromáticos
DCP43005	Monoxigenase da família Luciferase	-
DCP41333	Alcano 1-monoxigenase / Rubredoxina	Metabolismo de hidrocarbonetos alifáticos
DCP43007	L-Lactato 2- monoxigenase (LMO)	Oxidação de L-lactato a piruvato
DCP40660	Alcanosulfonate monoxigenase	Metabolismo de hidrocarbonetos alifáticos
DCP41330	Ciclohexanona monoxigenase	Metabolismo de hidrocarbonetos cíclicos/esteróides
DCP43010	Alcano monoxigenase	Metabolismo de hidrocarbonetos alifáticos
DCP43011	CypX, Citocromo P450	Monoxigenase sem função conhecida
DCP41291	3-cetoesteroide 9 α -hidroxilase componente KshA	Metabolismo de colesterol
DCP45313	NcnH, Nafto ciclinona hidroxilase	Biossíntese do antibiótico naftoquinona
DCP45314	Monoxigenase	Metabolismo de hidrocarbonetos alifáticos
DCP45315	Monoxigenase	Metabolismo de hidrocarbonetos alifáticos
DCP45316	Monoxigenase da família Luciferase	-
DCP45317	Monoxigenase da família Luciferase	-
DCP45318	P450, Linalool 8-monoxigenase	Metabolismo de álcool terpeno
DCP45319	Monoxigenase	Monoxigenase sem função conhecida
DCP45320	Monoxigenase 3	Monoxigenase sem função conhecida
DCP45322	Monoxigenase da família Luciferase	-
DCP45323	Monoxigenase da família Luciferase	-
DCP45325	Monoxigenase da família Luciferase	-
DCP45326	Citocromo P450 ácido graxo α -hidroxilase	Hidroxilação de ácidos graxos de cadeia longa

Monoxigenases

Monoxigenases	DCP45327	Citocromo P450 CYP123	Monoxigenase sem função conhecida
	DCP45328	Citocromo P450 CYP51	Biossíntese de esterol
	DCP45329	Monoxigenase	Metabolismo de hidrocarbonetos alifáticos
	DCP41295	Colesterol oxidase	Metabolismo de colesterol
	DCP45331	Monoxigenase da família Luciferase	-
	DCP45332	P450, Linalool 8-monoxigenase	Metabolismo de álcool terpeno
	DCP45333	Monoxigenase da família Luciferase	-
	DCP45334	IucD, Lisina/ornitina N- monoxigenase	
	DCP45335	Monoxigenase da família Luciferase	-
	DCP45336	P450, Linalool 8-monoxigenase	Metabolismo de álcool terpeno
	DCP45337	Monoxigenase da família Luciferase	-
	DCP45338	CypX, Citocromo P450 CYP136	
	DCP45339	Monoxigenase da família Luciferase	-
	DCP45340	L-Lactato 2- monoxigenase (LMO)	Oxidação de L-lactato a piruvato
	DCP41911	Nitrilotriacetato monoxigenase	Metabolismo de compostos nitrogenados
	DCP45342	Monoxigenase da família Luciferase	-
	DCP45343	P450, Linalool 8- monoxigenase	Metabolismo de álcool terpeno
	DCP45344	Monoxigenase da família Luciferase	-
	DCP41296	Colesterol oxidase	Metabolismo de colesterol
	DCP45309	Fenilacetato-CoA oxigenase	
Dioxigenases	DCP41915	2-nitropropano dioxigenase	Metabolismo de compostos nitrogenados
	DCP40654	TauD, taurino dioxigenase	Metabolismo de compostos sulforosos
	DCP41916	2-Nitropropano dioxigenase	Metabolismo de compostos nitrogenados
	DCP41298	Extradiol dioxigenase	Metabolismo de colesterol
	DCP45050	Óxido nítrico dioxigenase	Metabolismo de compostos nitrogenados
	DCP41311	BphC, 2,3-dihidroxi bifenil 1,2- dioxigenase	Metabolismo de hidrocarbonetos aromáticos
	DCP41312	Protocatecuato 4,5-dioxigenase	Metabolismo de hidrocarbonetos aromáticos
	DCP41315	BenA, Benzoato 1,2-dioxigenase	Metabolismo de hidrocarbonetos aromáticos
	DCP41302	BenDO, Benzoato dioxigenase (BenDO)	Metabolismo de hidrocarbonetos aromáticos
	DCP41318	CatA, 1,2-CCD, clorocatecol 1,2- dioxigenase	Metabolismo de hidrocarbonetos aromáticos

DCP41316	BenB, Benzoato 1,2-dioxigenase	Metabolismo de hidrocarbonetos aromáticos
DCP41299	Extradiol dioxigenase, enzima classe III, subunidade B	Metabolismo de colesterol
DCP41917	2-Nitropropano dioxigenase	Metabolismo de compostos nitrogenados
DCP41918	2-Nitropropano dioxigenase	Metabolismo de compostos nitrogenados
DCP45365	CobG, precorrina-3B sintase	Biossíntese da vitamina B12
DCP41327	BphC, 2,3-dihidroxi bifenil 1,2-dioxigenase	Metabolismo de hidrocarbonetos aromáticos
DCP41919	2-Nitropropano dioxigenase	Metabolismo de compostos nitrogenados
DCP45368	PcbC, Isopenicilina N sintase	Biossíntese de antibiótico tipo β -lactâmico

Entre os prováveis genes que codificam as oxigenases, 15 apresentaram similaridade com monoxigenases da família das luciferases. Não foi possível estabelecer uma relação entre os 15 genes possíveis que codificam monoxigenases da família das luciferases e um papel provável para sua ecofisiologia. Das 17 prováveis oxigenases, envolvidas na degradação de hidrocarbonetos, nove apresentaram similaridade com monoxigenases e oito com dioxigenases. Entre as monoxigenases, não foi possível determinar o tipo de substrato específico utilizado pelas três CDS DCP41335-37. Também foram detectados dois genes putativos L-lactato monoxigenase (DCP45307, DCP45340), responsáveis pela conversão de L-lactato a piruvato. Outros genes putativos observados foram o gene que codifica a isopenicilina N sintase (DCP45368), que está associado à síntese de antibiótico tipo β -lactâmico, e os genes hipotéticos piridoxamina 5'-fosfato monooxigenase (DCP45302) e precorrina-3B sintase dioxigenase (DCP45365), envolvidos na biossíntese das vitaminas B6 e B12, respectivamente.

Foram detectados nove prováveis oxigenases citocromos P450 no genoma de *D. cinnamomea* P4. Destas sequências codificantes de proteínas, quatro apresentaram similaridade com P450 linalool 8-monoxigenase, que é responsável pela utilização de compostos de linalool como fonte de carbono (DEPREZ *et al.*, 2002). Linalools são compostos de álcool terpenos sintetizados por plantas e encontrados especificamente nas flores. A utilização de produtos de origem vegetal, como fonte de carbono pela estirpe P4, sugere a presença de diferentes mecanismos para a aquisição de diferentes fontes de carbono disponíveis no solo. Também foi detectada uma provável P450 ácido graxo α -hidroxilase (DCP45326), envolvida na α -hidroxilação incomum de cadeias longas de ácidos graxos, como por exemplo, do ácido mirístico (LEE *et al.*, 2003), além de um putativo gene naftociclinona hidroxilase, *ncnH*

(DCP41442), envolvido na síntese do antibiótico naftoquinona (BRÜNKE *et al.*, 2001) e a hipotética monoxigenase esterol 14 α -demitilase, pertencente à família citocromo P450 CYP51 (DCP51297). A sequência da CDS DCP51297 compartilhou 72% de identidade com a sequência P450 CYP51 de *Rhodococcus jostii* RHA1 e *Rhodococcus opacus* B4. A reação catalisada pela enzima esterol 14 α -demitilase envolve a remoção do grupo 14 α -metil através de oxidações sucessivas, resultando na liberação de ácido fórmico e 8,14-dieno, esse último podendo ser utilizado na biossíntese de diversos compostos lipofílicos, e.g. esteróides, ácidos graxos e moléculas de sinalização (LEPESHEVA & WATERMAN, 2007).

6.11 – Assimilação de nitrogênio

A presença de genes putativos envolvidos na assimilação de compostos de nitrogênio por meio da conversão a amônia foi observada após o sequenciamento do genoma de *D. cinnamea* P4. A assimilação de nitrato inicia-se pelo provável mecanismo descrito no subitem 6.5.2 onde, por meio de um processo respiratório anaeróbico, o nitrato é reduzido a nitrito através do complexo Nar. O segundo passo é a redução de nitrito a amônia por um complexo composto por uma enzima nitrito redutase dependente de NAD(P)H (NirB) e uma enzima que tipicamente atua no processo de assimilação de nitrato em várias bactérias (NirD). Contudo, na estirpe P4, um mecanismo atípico parece ocorrer na redução de nitrito a amônia. Inicialmente, foi detectada a presença de três cópias do provável gene *nirB* (DCP41215, DCP41218, DCP41220) no genoma de P4. Entretanto, ao invés da presença de um gene *nirD* hipotético, foi anotado duas cópias do gene putativo *narB* que codifica a enzima ferredoxina-nitrate redutase (DCP41216-17), o qual também possui um papel semelhante à proteína NirD na conversão de nitrato a amônia. Foram identificados também dois possíveis genes envolvidos no transporte de nitrato e nitrito (DCP41213, DCP41219). A CDS DCP41213 apresentou 76% de homologia com um transportador da classe MFS (*major facilitador transport*) envolvido na extrusão de nitrito denominado de NirK, enquanto a CDS DCP41219 apresentou 85% de homologia com o transportador de formato/nitrito NirC.

A identificação dos genes hipotéticos para as proteínas glutamina sintase (DCP40626), as subunidades alfa e beta da glutamato sintase (DCP40626-27), glutamato desidrogenase (DCP40629-30), além de permeases para os compostos de glutamato (DCP40632) e amônia (DCP40633) indica uma possível utilização destes compostos nitrogenados como fonte primária de nitrogênio. Também é sugerida a capacidade da P4 em obter compostos de uréia do meio extracelular e a sua subsequente utilização como fonte de nitrogênio. Foram identificados os prováveis genes que codificam as subunidades alfa, beta e gama do complexo

urease (DCP41237-38), dois prováveis genes acessórios do complexo urease *ureF* e *ureH* (DCP41239-40), dois genes hipotéticos uréia carboxilase (DCP41235-36) e cinco putativos transportadores tipo ABC para composto de uréia (DCP41241-45).

6.12 – Assimilação de compostos inorgânicos

6.12.1 Enxofre – Além do sistema de aquisição de compostos de enxofre a partir de alcanosulfonatos pelo complexo SsuED descrito no subitem 6.7.1, a anotação do genoma da estirpe P4 permitiu também detectar os genes putativos *tauABC* (DCP41270-72) e, em uma outra posição do genoma, o gene hipotético *tauD* (DCP41254). Além disso, foi detectada a presença de mais um possível regulador transcricional da família TetR (DCP41253) adjacente ao *tauD*. A estirpe *D. cinnamea* P4 provavelmente utiliza o sistema Tau para a assimilação de enxofre a partir de compostos de taurina, especificamente o ácido 2-aminoetanosulfônico (Figura 31). Para isso, os compostos de taurina são primeiramente captados do meio extracelular através de um sistema de transporte tipo ABC codificado pelos genes *tauABC*. Em seguida, a proteína TauD, uma dioxigenase dependente de α -cetogluturato, realiza uma clivagem oxigenolítica liberando sulfito e taurina. Também foram identificados dois possíveis genes que codificam a enzima sulfito oxidase (SO) (DCP41266, DCP41275), que catalisa a conversão do sulfito a sulfato. O produto desta conversão, o sulfato, é então direcionado para ser assimilado como enxofre pela célula através do sistema Cys descrito abaixo.

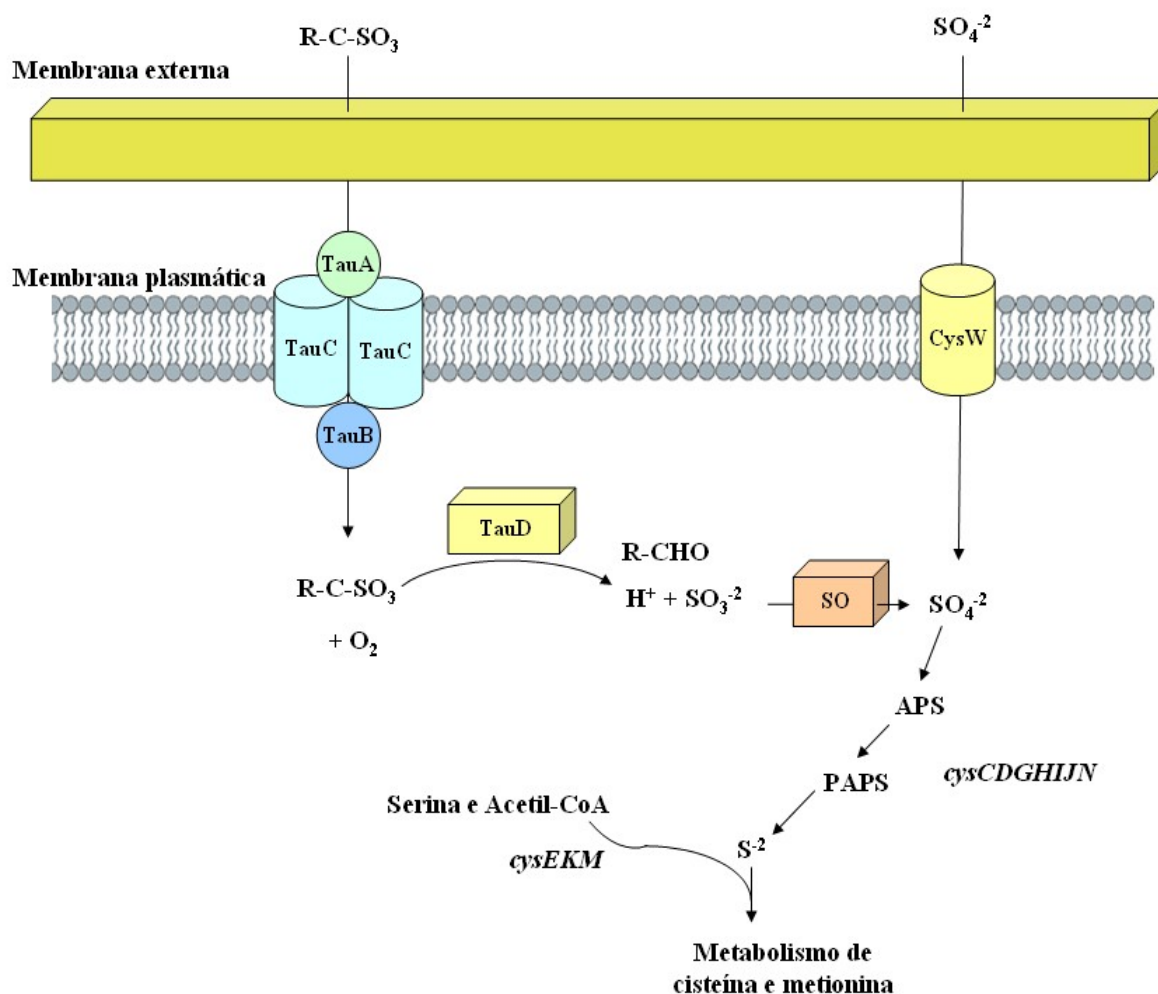


Figura 31: Esquema representativo da assimilação de enxofre utilizando os sistemas Tau e Cys, possivelmente presentes no genoma de *D. cinnamea* P4 (van der PLOEG, EICHHORN & LEISINGER, 2001).

A via assimilatória de redução de sulfato parece existir na estirpe P4 já que foram anotados oito possíveis genes que compõem o *operon cysDHIJKNW*. Hipoteticamente, o sulfato pode ser utilizado pela célula de *D. cinnamea* P4 após ser obtido do meio extracelular através do transportador tipo ABC CysW (DCP41268) e pela ação do produto das reações catalisadas pelas vias metabólicas Ssu e Tau. Em seguida, o sulfato é provavelmente ativado em uma reação de adenilação catalisada pelo complexo bifuncional CysCDN (DCP41256-57). O produto desta reação, o adenilil-sulfato (APS), é novamente reduzido a sulfito pela enzima CysH (DCP41252) e as enzimas CysIJ (DCP41251, DCP41279) catalisam a redução do sulfito a sulfeto de hidrogênio (H_2S). O H_2S é então direcionado para ser assimilado pelo metabolismo de cisteína e metionina, através da enzima CysK (DCP41267, DCP41278).

6.12.2 *Fósforo* – A anotação do genoma da estirpe P4 identificou genes hipotéticos da família Pho e Pst, os quais são responsáveis pela resposta da célula às condições de escassez de fósforo. O fósforo extracelular pode ser assimilado através dos genes hipotéticos *phoB* (DCP41290) e *phoR* (DCP41291), que compõem um sistema de sinalização dois-componentes, e um provável regulador transcricional (DCP41292) responsável pela expressão do gene *phoB*. A incorporação para o meio intracelular dos compostos de fosfato parece acontecer pela presença dos possíveis genes *pstA* (DCP41294-95) e *pstS* (DCP41293), que codificam proteínas que compõem um sistema de transporte tipo ABC. Também foi detectado um provável gene regulador *phoU* (DCP41296), sugerido pela literatura como um possível regulador da expressão dos genes *pst*. Além disso, foi identificado o gene putativo *phoD* (DCP41297) que apresenta atividade fosfatase alcalina e pode estar envolvido na hidrólise de ligações ésteres fosfato quando em condições limitantes de fosfato.

6.10.3 *Ferro* – Mesmo com uma grande quantidade de ferro disponível no nicho onde foi isolada a estirpe P4, a anotação do seu genoma revelou um elaborado sistema para a aquisição de ferro a partir do meio extracelular. Foram detectados sete genes hipotéticos do sistema de transporte do tipo ABC envolvidos na aquisição de ferro complexado com outros compostos como sideróforos, cobalamina e dicitrato (DCP40683-89). Também foi identificado um provável regulador transcricional da família Fur (DCP40681), que pode ser relacionado com a aquisição de ferro (HÉCTOR-OSORIO *et al.*, 2008), e uma proteína hipotética envolvida no efluxo de ferro ferroso FieF (DCP40682).

6.13 – Enzimas celulolíticas

A análise do genoma de *D. cinnamea* P4 utilizando o banco de dados CAZy MdoO disponível no sítio <http://www.cazy.org> resultou na identificação provável de 25 genes que codificam enzimas hidrolíticas de carboidratos (Tabela 5). Entre estas 25 prováveis enzimas hidrolíticas, oito foram identificadas em grupamentos gênicos, enquanto as outras 17 restantes foram encontradas espalhadas por todo o genoma. A classificação das 25 CDS em famílias de enzimas hidrolíticas de carboidratos identificou 21 CDS como pertencentes ao grupo glicosídeo hidrolases (GH) e quatro como pertencentes ao grupo carboidrato esterases (CE). Os prováveis genes pertencentes ao grupo glicosídeo hidrolases estão distribuídos entre 10 famílias GH diferentes, sendo a família GH13 a mais representativa (com 9 representantes).

Sete destes apresentaram o domínio CBM (*carbohydrate-binding modules*). O programa SignalP, disponível no sítio <http://www.cbs.dtu.dk/services/SignalP/output.php>, permitiu a identificação, entre as prováveis 25 CDS com homologia com as enzimas hidrolíticas de carboidratos, de sete com peptídeo sinal. Este fato sugere uma possível ação extracelular dessas sete CDs.

Tabela 5: Resultado da análise de similaridade entre as sequências deduzidas de aminoácidos de *D. cinnamea* P4 e as sequências depositadas no banco de dados CAZy, disponível no sítio <http://www.cazy.org>.

CDS	Melhor índice de similaridade	Número E.C.	Classe	Peptídeo sinal ^a	Processo
DCP47001	AmyA, α -amilase	3.2.1.1	GH13	-	Endo-hidrólise de ligações 1→4 α -D-glicosídica em polissacarídeos
DCP47002	GlgX, transglicosilase	2.4.1.18	GH57 (CBM)	ATA-AP	Transfere um segmento de uma cadeia 1→4 α -D-glicano para um grupo hidroxil
DCP47003	AmyA, α -amilase	3.2.1.1	GH13	-	Endo-hidrólise de ligações 1→4 α -D-glicosídica em polissacarídeos
DCP47004	GlgX, transglicosilase	2.4.1.18	GH57 (CBM)	ATA-AP	Transfere um segmento de uma cadeia 1→4 α -D-glicano para um grupo hidroxil
DCP47005	GlgX, transglicosilase	2.4.1.18	GH57 (CBM)	ATA-AP	Transfere um segmento de uma cadeia 1→4 α -D-glicano para um grupo hidroxil
DCP47006	TreY, malto-oligosil trealose sintase	5.4.99.15	GH13	-	
DCP47007	Malto-oligosil trealose-trealo hidrolase	3.2.1.141	GH13	VPA-VP	Hidrólise de ligações 1→4 α -D-glicosídica em polissacarídeos para produzir trealose e 1→4 α -D-glicano
DCP47008	TH1, trealose e maltose hidrolase		GH13 (CBM)	-	
DCP47009	Trealose sintase	5.4.99.16	GH13	-	Maltose→ α , α -trealose
DCP47010	α -trealose-fosfato sintase	2.4.1.245	GH13	-	ADP-glicose + D-glicose = α , α -trealose + ADP
DCP47013	Malto-oligosil trealose-trealo hidrolase	3.2.1.141	GH13	VPA-VP	Hidrólise de ligações 1→4 α -D-glicosídica em polissacarídeos para

					produzir trealose e 1→4 α-D-glicano
DCP47014	Quitinase	3.1.1.74	CE5	-	Quitina + H ₂ O = Monômero de quitina
DCP47015	Celobiohidrolase A		GH6 (CBM)	-	Hidrólise de ligações 1→4 β-D-glicosídica em celulose liberando celobiose
DCP47016	Celobiohidrolase A		GH6 (CBM)	-	Hidrólise de ligações 1→4 β-D-glicosídica em celulose liberando celobiose
DCP47017	Precursor da lisozima		GH24	ASA-AA	
DCP47018	N-acetil glicosaminil fosfatidil inositol deacetilase	3.5.1.89	CE14	VLA-AA	6-(N-acetil-α-D-glicosaminil)-1-fosfatidil-1D-mio-inositol + H ₂ O = 6-(α-D-glicosaminil)-1-fosfatidil-1D-mio-inositol + acetate
DCP47019	BglX, β-glicosidase-	3.2.1.21	GH3	-	Hidrólise terminal de β-glicosil
DCP47020	ManA, Mannan endo-1,4-β-mannosidase	3.2.1.78	GH26 (CBM)	-	Hidrólise randômica de ligações 1→4 β-D-manosídica em mananas
DCP47021	malQ, 4-α-glicano transferase	2.4.1.25	GH77	-	Transferência de segmentos 1→4 α-D-glicano em uma nova posição em um acceptor de glicose
DCP47022	1D-mio-inositol-2-acetamido-2-deoxi-α-D-glicopiranosídeo deacetilase	3.5.1.89	CE14	-	6-(N-acetil-α-D-glicosaminil)-1-fosfatidil-1D-mio-inositol + H ₂ O = 6-(α-D-glicosaminil)-1-fosfatidil-1D-mio-inositol + acetate
DCP47023	AmyA, α-amilase	3.2.1.1	GH13	-	Endo-hidrólise de ligações 1→4 α-D-glicosídica em polissacarídeos
DCP47024	OtsA, α,α-trealose-fosfato sintase	2.4.1.15	GT20	-	UDP-glicose + D-glicose 6-fosfato = UDP + α, α-trealose 6-fosfato
DCP47025	Arabino endo-1,5-α-L-arabinosidase	3.2.1.99	GH43	ATA-AA	Endo hidrólise de ligações 1→5 α-arabinofuranosídica em 1→5-arabinanos
DCP47026	α-L-fucosidase	3.2.1.51	GH29	ARR-CR	α-L-fucosídeo + H ₂ O = L-fucose + álcool

^a Sequência do sítio de clivagem

Doze das 25 prováveis enzimas hidrolíticas de carboidratos estão envolvidas em atividades celulolíticas: três prováveis α -amilases (GH13) (DCP47001, DCP47003, DCP47023), uma carboidrato esterase quitinase (DCP47014), duas celobiohidrolases da família GH6 (DCP47015), sendo uma delas apresentando a sequência de aminoácidos CBM, duas enzimas envolvidas na degradação de compostos de maltose (DCP47007-08), ambas da família GH13 e uma delas com o domínio CBM, uma manose hidrolase (DCP47020) da família GH26, apresentando o domínio CBM e uma arabinosidase da família GH43 (DCP47026). Além disso, duas prováveis deacetilases de compostos de *N*-acetil-glucosamina, DCP47018 e DCP47022, sugerem a possibilidade de uso destes compostos na produção de energia pela célula (LYKIDIS *et al.*, 2007).

6.14 – Metabolismo secundário

A identificação de 42 CDS com similaridade a genes relacionados à biossíntese de metabólitos secundários (também denominados de produtos naturais) sugere que a estirpe P4 seja um importante produtor destes compostos. Foram identificados genes envolvidos na biossíntese de compostos terpenóides, β -caroteno, ramanose e genes putativos codificando policetídeos sintases.

A utilização da via metabólica do esqueleto terpenóide pela estirpe P4 é sugerida pela presença de 16 CDSs, das quais oito participariam da via 2-*C*-metil-D-eritritol 4-fosfato/1-deoxi-D-xilulose 5-fosfato (MEP/DOXP), responsável pela síntese de geranil a partir de piruvato e gliceroldeído 3-fosfato (WANKE, SKORUPINSKA-TUDEK & SWIEZEWSKA, 2001). A conversão de geranil a undecaprenil difosfato é indicada pela presença dos genes hipotéticos *cis*-(*Z*)-isoprenil difosfato sintase (DCP41427) e undecaprenil pirofosfato sintase (DCP41416). O composto geranil também pode ser direcionado para a via de biossíntese de compostos monoterpénoides. Em *D. cinnam*e P4, embora não tenha sido possível identificar a presença da provável enzima limoneno monoxigenase, a presença do gene hipotético carveol desidrogenase (DCP41429) sugere uma possível conversão de geranil a (*S*)-carvone. Outro destino possível para o geranil é a sua conversão a β -caroteno. Após o geranil ser convertido a farnesil difosfato pela ação da CDS DCP41427, o gene hipotético geranil-geranil pirofosfato sintase (DCP41406) o converte a geranil-geranil difosfato, o qual em seguida é convertido a fitoeno pela enzima hipotética fitoeno sintase, *crtB* (DCP41407). Em seguida a molécula de fitoeno é convertida a licopeno por uma série de reações catalisadas pela provável enzima fitoeno desidrogenase, *crtI* (DCP41402, DCP41404) e, finalmente, convertido a β -caroteno pela ação da hipotética licopeno β -ciclase, *crtLb* (DCP41400) (Figura

32).

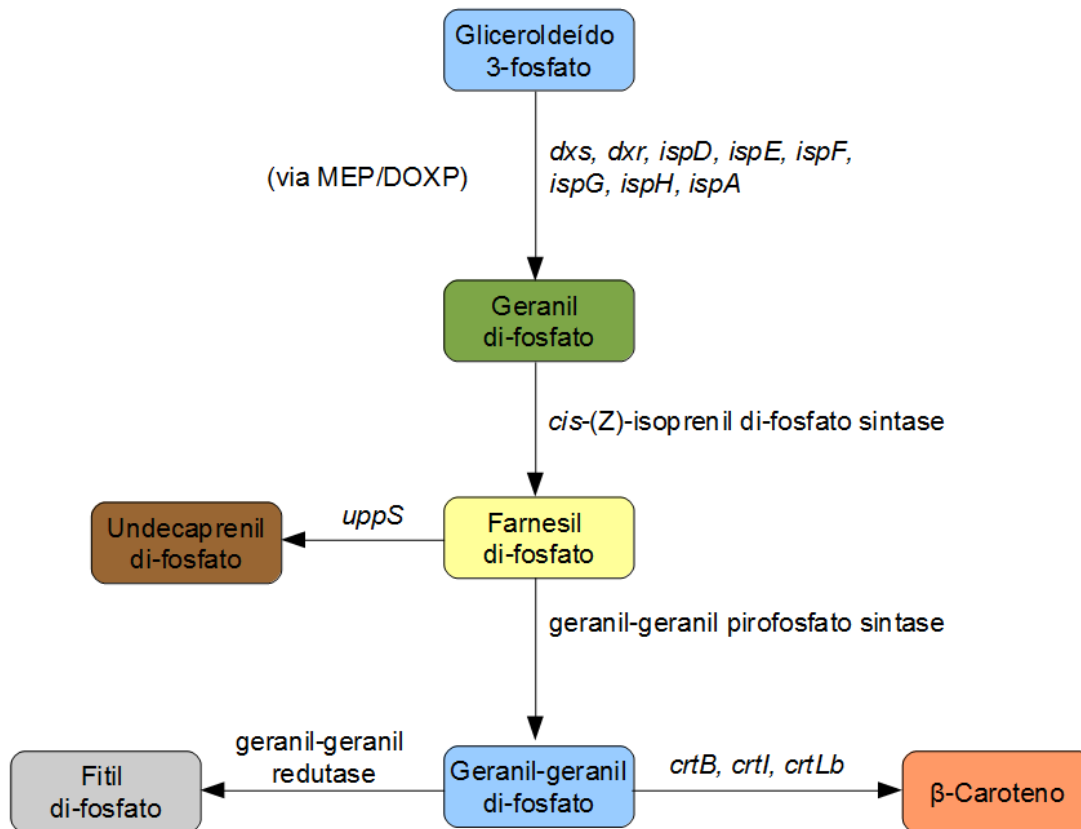


Figura 32: Esquema representativo da via biossintética dos compostos de geranil, undecaprenil, fitil e β-caroteno que pode estar presente no genoma de *D. cinnamea* P4 (WANKE , SKORUPINSKA-TUDEK & SWIEZEWSKA, 2001).

Com a anotação do genoma da estirpe P4 foram identificados cinco possíveis genes envolvidos na síntese da molécula de dTDP-ramnose a partir de glicose 1-fosfato, distribuídos em três diferentes regiões do genoma de *D. cinnamea* P4. A dTDP-ramnose é um importante componente da membrana celular e faz parte do esqueleto carbônico da estrutura do antibiótico novobiocina (BOELS *et al.*, 2004; THUY *et al.*, 2005). Os genes putativos dTDP-glicose desidratase, *rfbB* (DCP41411) e dTDP-dehidro-ramnose epimerase, *rfbD* (DCP41412) foram identificados juntamente com outros genes envolvidos na síntese de polissacarídeos capsulares (subitem 6.19). Em outro local do genoma foram detectados uma segunda cópia da dTDP-glicose desidratase, *rfbB* (DCP41413) além do provável gene glicose 1-fosfato transferase, *rfbA* (DCP41413). Foi identificado também o gene hipotético dTDP-dehidro-

ramnose epimerase, *rfbC* (DCP41414), mas em outra região.

Entre os sete genes putativos que codificam policetídeos sintases detectados no genoma da estirpe P4, dois estão envolvidos na biossíntese de antibióticos aromáticos: o gene hipotético *frnE* (DCP41441), associado à síntese dos antibióticos frenolicina e nanaomicina (LI *et al.*, 2003). Também foram identificados o gene hipotético *aviGT4* (DCP41447), relacionado à síntese do antibiótico avilamicina A (MARTINEZ-FLEITES *et al.*, 2006) e um provável gene que codifica a proteína AziB (DCP41449). Esta proteína juntamente com as prováveis enzimas metiltransferase dependente de S-adenosilmetionina (DCP41450) e citocromo P450 hidroxilase (DCP41451) são descritas como componentes da via de síntese do potente anti-tumor Azinomicina B (ZHAO *et al.*, 2008). Além disso, foi anotado um policetídeo sintase com domínio β -cetoacil (DCP41448), envolvido na síntese de ácido micólico (importante componente da parede celular de bactérias Gram-positivas) e mais dois genes hipotéticos que codificam policetídeos sintases sem função ainda conhecida (DCP41445-46). A anotação do genoma de *D. cinnamea* P4 permitiu ainda identificar apenas um peptídeo sintetase não ribossomal (NRPS), a CDS DCP41443, que apresentou 55% de similaridade com o NRPS depositado no banco de dados CDD com o número de referência 131406.

6.15 – Transporte transmembrânico

As análises do genoma de *D. cinnamea* P4 empregando-se as ferramentas BLAST, BALSTALL, TMHMM e TransportDB levaram à identificação de 312 prováveis proteínas associadas ao transporte transmembrânico de solutos, correspondendo a cerca de 9% das CDS descritas. Entre as sequências deduzidas de aminoácidos a partir da estirpe P4 que apresentaram homologia com as proteínas transportadoras, a classe mais representativa foi o transporte tipo ABC (49% das CDS) (Figura 33). Entre os diversos tipos de solutos transportados pela classe de transportador tipo ABC destaca-se os envolvidos no transporte de carboidratos, e.g. glicose, maltose, xilose e ribose. Também se destacaram o transporte de solventes orgânicos (subitem 6.18.5), compostos nitrogenados como uréia, amônia, glutamato e glutamina, peptídeos e aminoácidos, vitaminas B12 e compostos de policetídeos (DCP43132) e compostos inorgânicos de manganês, zinco, arsênio e cobalto, normalmente associados à resistência a metais pesados (subitem 6.18.2). A segunda classe de transportadores mais representada é a classe MFS (*major facilitator superfamily*), com 49 prováveis proteínas presentes no genoma de *D. cinnamea* P4. O principal tipo de transporte realizado pelas proteínas da classe MFS é o envolvido no efluxo de drogas (subitem 6.18.4).

Além disso, o transporte tipo MFS também é associado ao transporte de compostos de benzeno (subitem 6.7.2), glicerol (DCP43145), acetato (DCP43150), formato (subitem 6.6.3) e alguns aminoácidos (Anexo II).

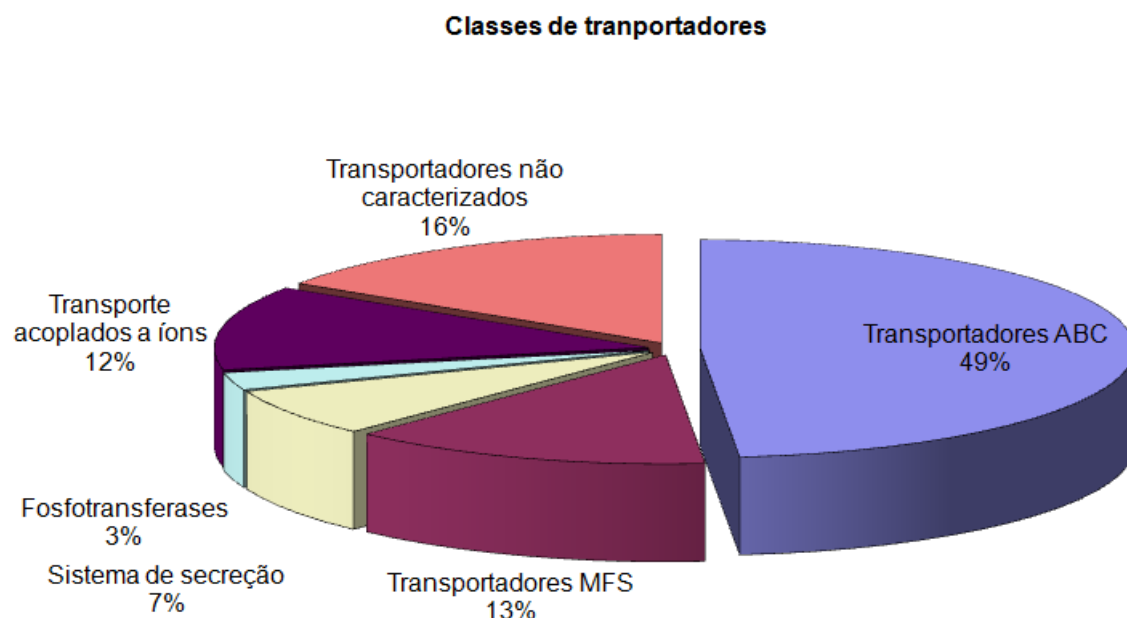


Figura 33: Distribuição das classes de prováveis transportadores presentes no genoma de *D. cinnamea* P4, classificados através das ferramentas BLAST, BALSTALL, TMHMM e TransportDB (veja material e métodos).

O sequenciamento do genoma da estirpe P4 levou à identificação de 21 prováveis CDS envolvidas na secreção de proteínas. Entre elas, 11 foram identificadas como prováveis proteínas pertencentes ao sistema de secreção Sec: duas prováveis subunidades SecA (DCP43195-96), uma subunidade SecD (DCP43190), três subunidades SecF (DCP43191-92), uma SecY (DCP43193), duas proteínas hipotéticas YidC (DCP43205; DCP43208) e duas partículas de reconhecimento, FtsY (DCP43209) e Ffh (DCP43250).

Braunstein e colaboradores (2003) haviam relatado que as duas cópias do gene *secA* presentes em *M. tuberculosis* apresentam funções distintas na célula, enquanto a translocase SecA1 está relacionada com as atividades de transporte de diversas proteínas como chaperonas, fosfolipídeos e proteínas transmembrânicas. Uma segunda proteína SecA2 estaria envolvida na exportação de proteínas da família super óxido dismutase (Sod) (subitem 6.17.1), que tem papel importante na virulência durante a infecção de *M. tuberculosis* em

macrófagos. Com base nessas observações foram realizadas então análises de similaridade entre as sequências DCP43195 e DCP43196 de *D. cinnamea* P4 e outras sequências deduzidas de aminoácidos de organismos relacionados. Os resultados das comparações de similaridade entre DCP43195 e DCP43196 foram de aproximadamente 50%, enquanto que DCP43195 apresentou cerca de 80% de similaridade com as sequências deduzidas de aminoácidos de SecA1 das estirpes de *M. tuberculosis* e DCP43196 apresentou a mesma média de 80% de similaridade com as sequências deduzidas de aminoácidos de SecA2. A construção da árvore filogenética utilizando-se as mesmas sequências empregadas nas análises de similaridade demonstrou a proximidade da CD DCP43195 com SecA1 e da CD DCP43196 com SecA2 (Figura 34).

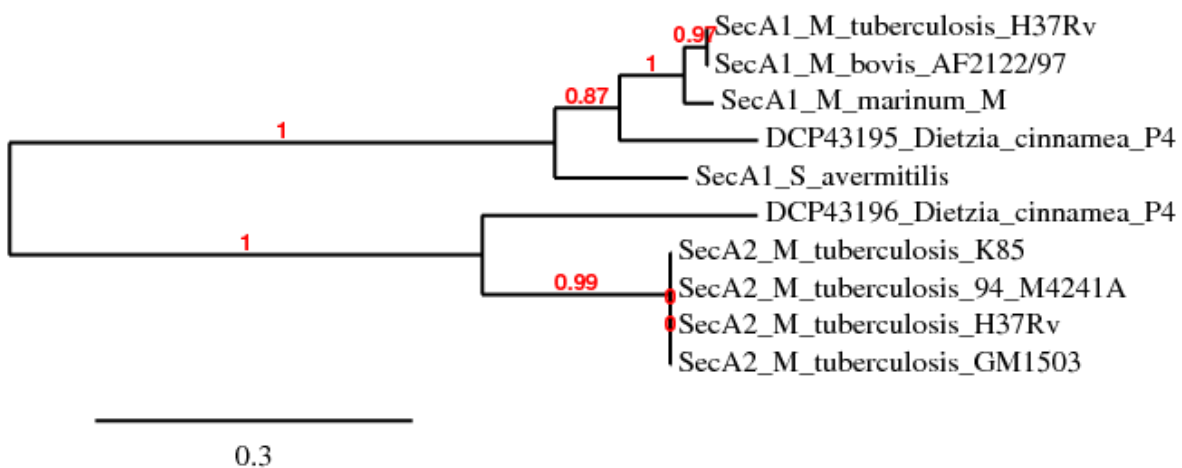


Figura 34: Análise filogenética entre as sequências de aminoácidos DCP43195 e DCP43196 de *D. cinnamea* P4 com sequências deduzidas de aminoácidos de SecA1 e SecA2 de organismos relacionados.

Os genes hipotéticos *tatB* (DCP43106), *tatC* (DCP43197) e *tatE* (DCP43199, DCP43107) que compõem o sistema de secreção tipo *Tat independente de Sec* estão presentes no genoma da estirpe P4. O sistema de secreção *Tat* está envolvido no transporte de proteínas em sua conformação final, normalmente associado na translocação de proteínas de ação extracelular, como as enzimas hidrolíticas de carboidratos, descritas no subitem 6.13. Foram detectadas quatro prováveis proteínas das 12 que compõem o sistema de secreção tipo *Tat*, o qual compartilha componentes da maquinaria de biogênese do complexo pili: *TadB* (DCP43203), *CpaE* (DCP43251), *CpaF* (DCP43203) e uma proteína que compõe o sistema

pili sem função definida (DCP43201).

Finalmente, foram identificadas 36 prováveis proteínas com atividade de transporte acoplado a íons (*ions-coupled transport*, ICT) e nove fosfotransferases putativas. Entre as prováveis ICTs, destacam-se as com atividade relacionada à resistência de metais pesados. Entre as fosfotransferases, foram identificadas quatro prováveis aminoglicosídeo 3'-fosfotransferases (DCP45098-99, DCP45101, DCP45105), uma provável fosfotransferase envolvida na resistência a antibióticos (DCP45104) e duas putativas fosfotransferases do tipo Pst que participa no transporte de carboidratos de frutose e manitol (DCP45103 e DCP43200).

6.16 – Duplicação, reparo e recombinação do DNA

Em relação à maquinaria necessária para a duplicação do DNA de *D. cinnamea* P4 foram detectados os prováveis genes que codificam a DNA primase, *dnaG* (DCP49050); DNA girase, *gyrA* (DCP49056, DCP49058); helicase, *dnaB* (DCP49060); DNA polimerase I, *polA* (DCP49064); DNA polimerase IV (DCP49053); DNA ligase, *ligD* (DCP49066-67); além das putativas subunidades que compõem a DNA polimerase III: subunidades alfa (DCP49052, DCP49061), beta (DCP49055), delta (DCP49057, DCP49065), épsilon (DCP49051, DCP49063) e gama/tau (DCP49059, DCP49062). Não foi possível identificar nenhuma CDS com homologia às proteínas SSBs, envolvidas na manutenção das fitas simples de DNA durante a replicação, e nem a proteína carreadora de helicase *dnaC*. As CDSs com similaridade com as proteínas componentes da maquinaria de replicação estão dispostas por todo o genoma de *D. cinnamea* P4, não parecendo estar organizadas na forma de *operon*.

A anotação do genoma da estirpe P4 também permitiu identificar prováveis proteínas envolvidas no reparo e recombinação do DNA. O sistema de recombinação homóloga é composto das putativas proteínas dos sistemas Rec e Ruv. O mecanismo de reparo do DNA da estirpe P4 deve utilizar, além da proteína hipotética RecA (DCP49037), as proteínas RadaA (DCP49015), DnaE2 *prone* (DCP49008), 3-metiladenina DNA glicosilase (DCP49006) envolvida na excisão de bases alquiladas, deoxiribodipirimidina foto-liase tipo I (DCP49014) envolvida no reparo de dímeros de pirimidina, as proteínas MutM (DCP49007) envolvidas no reparo de bases oxidadas e MutY (DCP49009-10) envolvida no reparo de pareamentos incorretos entre bases guanina e adenina. Também foram identificadas as putativas proteínas que compõem o sistema de reparo Uvr, composto pelas enzimas UvrABCD em *D. cinnamea* P4.

6.17 – Transcrição e tradução

Os prováveis genes que codificam as subunidades RNA polimerase alfa, *rpoA* (DCP49102); alfa', *rpoS* (DCP49101); beta, *rpoB* (DCP49109) e beta', *rpoC* (DCP49110) e três prováveis fatores sigma 70, *rpoE* (DCP49100, DCP49107, DCP49111) e um putativo fator sigma B (DCP49103) foram detectados no genoma de *D. cinnamea* P4. A anotação permitiu a identificação de cinco proteínas hipotéticas componentes da subunidade maior do ribossoma e quatro da subunidade menor (DCP49120-29), além dos putativos fatores de iniciação IF-2 (DCP49130) e IF-3 (DCP49131), fatores de elongação Ts, Tu e G (DCP49135-39) e os fator de terminação Rho (DCP49145).

Além disso, foram encontrados 72 prováveis reguladores transcricionais. Sessenta e dois dos prováveis reguladores transcricionais estão divididos em 27 famílias diferentes. Dez dos reguladores transcricionais encontrados não puderam ser agrupados em família A família de reguladores transcricionais mais comum entre todas foi a TetR com 13 CDSs representantes, seguida de LysR, IclR e LuxR com 5 representantes cada. Além das funções prováveis de manutenção celular, foi possível relacionar algumas funções com determinados reguladores transcricionais, como por exemplo, regulação da expressão de genes envolvidos na degradação de hidrocarbonetos alifáticos (subitem 6.5.1, DCP451331) e aromáticos (subitem 6.5.2, DCP451314), biossíntese de policetídeos (subitem 6.12, DCP451444), respostas a estresse oxidativo (subitem 6.18.1, DCP45058), metais pesados (subitem 6.18.3, DCP45005, DCP45018) e osmótico (subitem 6.18.3, DCP45113).

6.18 – Respostas a estresses

6.18.1 Estresse oxidativo

O sequenciamento identificou diversas proteínas no genoma de *D. cinnamea* P4 com algum provável papel nas respostas a estresses ambientais. A tabela 6 apresenta os resultados da análise de similaridade entre as CDSs da estirpe P4 com enzimas envolvidas nas respostas ao estresse oxidativo proveniente de organismos relacionados. Foram identificados três genes envolvidos na reação antioxidante em resposta ao oxigênio, sendo uma provável superóxido dismutase SodC (DCP45051) e duas superóxido dismutase SodA (DCP45063, DCP45064). Apesar do baixo índice de similaridade apresentada pela CDS DCP45041, não é comum a identificação de superóxido dismutase com centro reativo Cu/Zn (SodC) em procaríotos e sim a presença de superóxido dismutase Fe/Mn citosólica (SodA). O peróxido de hidrogênio (H₂O₂), possivelmente originado da reação catalisada pela superóxido dismutase, pode ser

utilizado como aceptor de elétrons na reação de oxidação de um substrato orgânico, como NADH ou glutatona, através da reação catalisada pelas enzimas peroxidase (DCP45042, DCP45057, DCP45071) e glutatona peroxidase (DCP45060), respectivamente. Outra enzima envolvida na destoxificação de composto H_2O_2 na estirpe P4 é a catalase, a qual está presente em seis prováveis cópias, sendo três do tipo catalase contendo heme (DCP45043-45, DCP45055-56, DCP45073) e três catalases contendo manganês (DCP45046-47). Também foram detectadas duas peroxidases hipotéticas não-heme designadas de peroxidoxina, *bcp* (DCP45048) e alquil peroxidase, *ahpD* (DCP45059), ambas contendo um centro catalítico com resíduos de cisteína, onde é catalisada a reação de decomposição de H_2O_2 , hidroperóxido orgânicos e peroxinitrito (POOLE, 2005). Um putativo regulador transcricional OxyR (DCP45058) envolvido em respostas aos altos níveis de H_2O_2 foi detectado no genoma localizado próximo da CDS DCP45059.

Um possível controle do balanço tiol-redox é realizado na célula de *D. cinnamea* P4 através das prováveis enzimas tioredoxina (DCP45067), tioredoxina redutase (DCP45069, DCP45062) e glutatona redutase (DCP45049). Uma provável enzima óxido nítrico dioxigenase (DCP45050) também foi detectada no genoma da estirpe P4, sugerindo uma provável ação na redução de óxido nítrico (GARDNER, 2008). Um provável *operon* contendo seis genes putativos do sistema de transporte Suf foi identificado no genoma da estirpe P4 (DCP45074-79). Além disso, foram detectados mais dois genes hipotéticos, *sufC* (DCP45080) e *sufI* (DCP45081), em diferentes localidades do genoma. Este sistema de transporte transmembrânico é expresso durante condições de estresse oxidativo e privação de níveis de ferro intracelular (OUTTEN *et al.*, 2003). O controle provável dos altos níveis de ferro intracelular é realizado pelo regulador transcricional hipotético da família Fur (DCP45054).

Tabela 6: Principais processos de defesa ao estresse oxidativo sugeridos pela anotação do genoma de *D. cinnamea* P4.

CDS	Melhor índice de similaridade ^a	Processo
DCP45065	Metionina sulfoxido redutase	
DCP45063, DCP45064	SodA, Fe/Mn superóxido dismutase	Dismutação do superóxido em O ₂ e H ₂ O ₂
DCP45041	SodC, Cu/Zn superóxido dismutase	Dismutação do superóxido em O ₂ e H ₂ O ₂
DCP45042, DCP45057, DCP45071	Peroxidase	Oxidação de NADH e H ₂ O ₂ em H ₂ O
DCP45043, DCP45044, DCP45045, DCP45073	KatE, Catalase contendo ferro	Decomposição de H ₂ O ₂ em H ₂ O
DCP45046, DCP45047, DCP45055, DCP45056	Catalase contendo heme	Decomposição de H ₂ O ₂ em H ₂ O
DCP45048	Peroxidoxina	Decomposição de H ₂ O ₂ em H ₂ O
DCP45049	Glutationa redutase	Redução da glutatona a sulfridil
DCP45053	Glutaredoxina	
DCP45060	Glutationa peroxidase	
DCP45068	Glutationa S-transferase	
DCP45061, DCP45062, DCP45069	Tioredoxina redutase	
DCP45072	Tioredoxina	
DCP45050	Óxido nítrico dioxigenase	
DCP45051	Tioredoxina	
DCP45052, DCP45070	Hidrolase NUDIX	Proteína envolvida nas respostas globais a estresses
DCP45059	AhpD, aquil peroxidase	
DCP45074-81	Suf, Transportadores tipo ABC	Transporte de Fe ⁽³⁺⁾
DCP45054	Fur, regulador transcricional	Regulador transcricional
DCP45058	OxyR, regulador transcricional	Regulador transcricional

6.18.2 Metais pesados

A capacidade da *D. cinnamea* P4 em se proteger contra os efeitos tóxicos dos metais pesados é evidenciada pela identificação de 33 CDS potencialmente envolvidas na resistência a estes compostos. Genes possivelmente envolvidos no efluxo de compostos de arsênio foram localizados em dois diferentes pontos do genoma. O primeiro local é uma região composta por uma proteína hipotética exportadora de arsenito (DCP45011) e uma enzima hipotética arsenato redutase (DCP45012). O segundo ponto contém dois putativos transportadores *em tandem*, os quais estão provavelmente envolvidos na exportação de compostos de arsênio (DCP45015-16). *D. cinnamea* P4 mostrou ser também resistente ao mercúrio. Uma enzima hipotética, a mercúrio redutase (DCP45016), foi detectada junto de um transportador putativo envolvido no transporte de metais pesados (DCP45015). Além disso, dois prováveis reguladores transcricionais da família MerR (DCP45005, DCP45018) foram identificados em diferentes localizações do genoma, sendo que a CDS DCP45018 está adjacente a uma provável proteína transportadora envolvida no efluxo de compostos de mercúrio, além de cádmio, zinco e cobalto. Embora tenham sido detectados três prováveis cópias dos genes *terC* (DCP45026-27, DCP45032), componente do sistema de resistência a íons de telúrio, não foi possível localizar outros genes, e.g. *terBDE*. Uma provável resistência a cobre é associada com a presença de prováveis genes sensores *copR* (DCP45014) e *copS* (DCP45014), três cópias do gene hipotético *copA* que codifica transportadores de cobre ATP dependentes (DCP45008, DCP45021, DCP45023) e uma cópia do provável gene *copC* que codifica também um transportador de compostos de cobre (DCP45004). O gasto de energia dispensado no efluxo de compostos de cobre pelo transportador CopA sugere a importância dada pela estirpe P4 na manutenção do nível intracelular de cobre. Apesar da importância de íons metais na homeostase intracelular, foi detectada uma proteína hipotética envolvida no efluxo de íons ferro, quando em altos níveis de concentração. Além disso, a presença de mais sete CDS com homologia a proteínas ATP dependentes com funções de efluxo de íons de cádmio, zinco e cobalto corrobora a importância da resistência de metais pesados para a ecologia de *D. cinnamea* P4.

6.18.3 Estresse osmótico

A análise do genoma permitiu a identificação de diversos genes envolvidos na resposta a estresse osmótico através do acúmulo de solutos intracelular por meio das vias biossintéticas de compostos de colina, betaina e trealose, além da aquisição extracelular de compostos de ectoine. A via biossintética de colina é sugerida através da presença de uma provável colina desidrogenase (DCP45111) e duas betaína aldeído desidrogenase hipotéticas (DCP45110, DCP45112), além de um possível transportador envolvido na obtenção destes compostos do meio extracelular (DCP45114-20). Também foi detectado outro sistema de transporte para a obtenção de compostos de ectoine (DCP43050-51). O acúmulo de compostos de ectoine, assim como para os compostos de colina e betaina, é relatado como um mecanismo para a tolerância ao estresse osmótico (VARGAS *et al.*, 2008). O acúmulo intracelular de trealose é outro mecanismo possível utilizado pela bactéria *D. cinnamea* P4 na tolerância ao estresse osmótico. Para isso, a estirpe P4 utiliza os depósitos de glicogênio disponíveis para a síntese deste composto, através da ação da provável enzima desramificadora de glicogênio TreX, que libera compostos de 1,4- α -glucano. Em seguida, as enzimas hipotéticas TreY e TreZ liberam os monômeros de trealose. Os genes que participam desta via metabólica estão arranjados na forma de *operon* no genoma (DCP40100-02). Além disso, foram anotados um possível sensor osmótico OmpR (DCP45113) que pode estar envolvido na sinalização em condições de estresse osmótico e uma aquaporina Z hipotética (DCP45121), sugerindo um papel no controle do movimento da água no esforço para a tolerância ao estresse osmótico.

6.18.4 Resistência a drogas

Entre os mecanismos envolvidos na sobrevivência de *D. cinnamea* P4 no solo, foram encontrados uma ampla variedade de genes de resistência a drogas em seu genoma (Tabela 7). Foi identificado um total de 12 prováveis β -lactamases, que podem estar envolvidas na resistência aos antibióticos β -lactâmicos, como penicilina, cefamicina e carbapenemas. Também foram detectados 13 transportadores hipotéticos tipo ABC e 10 transportadores tipo MFS envolvidos no efluxo de drogas, dois exportadores de antibióticos policetídeos e quatro fosfotransferases putativas com ação sobre os antibióticos canamicina e neomicina. Um gene que codifica a proteína denominada de Radical SAM 23-S rRNA Metiltransferase da família Cfr (DCP45088), também detectado no genoma da estirpe P4, sugere que essa estirpe seja capaz de realizar uma modificação pós-transcricional sobre o rRNA 23S, impedindo a ação de cinco classes de antibióticos (KAMINSKA *et al.*, 2010).

Tabela 7: Classificação das prováveis proteínas envolvidas em resistência a drogas presentes no genoma de *D. cinnamea* P4.

Provável classe de proteínas	Tipo de resistência conferida
β -lactamase	Hidrolisa ligações carbono-nitrogênio em amidas cíclicas
Transportador tipo ABC	Transporte de drogas para o meio extracelular
Transportador tipo MFS	Transporte de drogas para o meio extracelular
TnrB3	Transporte de antibióticos policetídeos
Canamicina e neomicina fosfotransferases	
Radical SAM Metiltransferase	Modificação pós-transcricional do rRNA 23S
Eritromicina esterase	Proteção contra antibióticos de produção própria

6.18.5 Resistência a outros sinais ambientais

Parece existir um sistema de tolerância a solventes orgânicos no genoma de *D. cinnamea* P4. Isso porque foram detectados nove prováveis genes que codificam transportadores transmembrânicos localizados em um único ponto do genoma. Além disso, uma única cópia do mesmo tipo de transportador foi detectada em outro local do genoma. Este sistema é composto pelos transportadores tipo ABC da família Ttg envolvidos no efluxo de solventes orgânicos, principalmente tolueno (KIM *et al.*, 1998).

A presença de genes da via metabólica do β -caroteno (descritos no subitem 6.14) segrega a resistência à radiação U.V. pela estirpe P4. Também foram identificadas diversas proteínas hipotéticas envolvidas em respostas a sinais de estresses, como o regulador transcricional de choque térmico HcrA (DCP45127), três prováveis chaperoninas ClpBCX (DCP45128, DCP45139-40, DCP45145), uma proteína de choque térmico hipotética Hsp90 (DCP45141) e uma proteína putativa de choque ao frio (DCP45132).

6.19 – Biossíntese da parede celular

Por fim, a análise do genoma de *D. cinnamea* P4 também permitiu localizar genes responsáveis pela biossíntese de peptídeoglicanos, componentes da parede celular. A via metabólica é iniciada pela construção de uma molécula de UDP-*N*-acetilmuramoil-L-alanil-D-glutamato, a partir de UDP-*N*-acetilglucosamina, em reações de adições de resíduos de aminoácidos e di-açúcares pelas enzimas MurA (DCP45200), MurB (DCP45213), MurC (DCP45209) e MurD (DCP45207). Em seguida, ocorrem uma série de reações de ligações e rearranjos catalisados pelas enzimas hipotéticas MurE (DCP45201, DCP45203), Ddl (DCP45212), MraY (DCP45206), MurG (DCP45208) e peptídeoglicano glicotransferase (DCP45211), dando origem ao peptídeoglicano que constitui a parede celular de *D. cinnamea* P4. Foram também identificadas três prováveis proteínas de ligação à penicilina (DCP45100, DCP45106, DCP45108), que participam no estágio final de síntese de peptídeoglicanos. Embora tenha sido detectada a hipotética proteína lauroil acil-transferase lipídio A (DCP45215), não foi identificada nenhuma outra enzima envolvida na síntese de lipopolissacarídeos no genoma da estirpe P4.

II - Evolução experimental de *D. cinnamea* P4 em condições de biorreator

6.20 – Crescimento em diferentes meios de cultura

Para se analisar a possibilidade de manutenção da célula de *D. cinnamea* P4 em condições de biorreator por longo período de incubação, esta estirpe foi crescida em batelada em meio de cultura líquido (LB e BH) contendo C-16 a 1% (v/v). O perfil de crescimento celular em meio LB em função do tempo de cultivo (em dias) é apresentado na figura 35. A cultura apresentou uma fase inicial de adaptação, a fase lag, de aproximadamente 3 dias. Após esse período, a cultura apresentou um crescimento exponencial de aproximadamente 3 dias, atingindo a fase de crescimento estacionária no 6º dia de cultivo. O início da fase de declínio tornou-se clara no 12º dia, com a diminuição do crescimento celular.

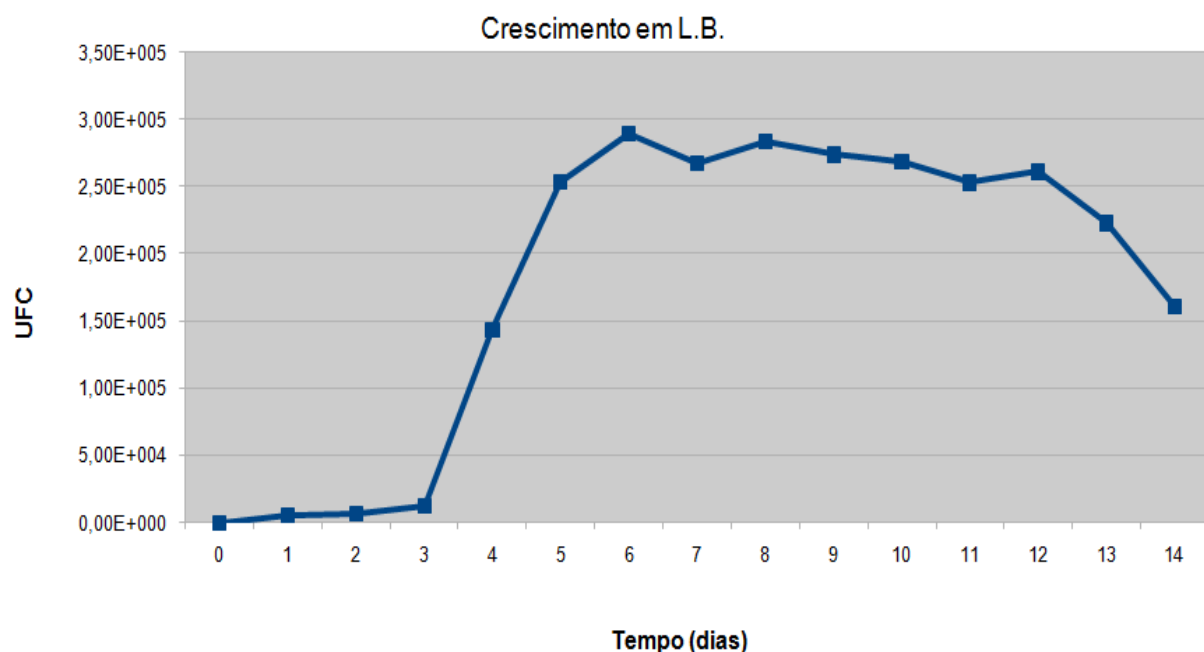


Figura 35: Curva de crescimento celular de *D. cinnamea* P4 em meio líquido LB.

O perfil de crescimento celular em meio líquido BH contendo C16 em função do tempo de cultivo em dias é apresentado na figura 36. A curva de crescimento nessas condições foi muito semelhante à obtida quando o meio utilizado foi o LB. Porém, a fase inicial de adaptação foi um pouco maior, com o início do crescimento exponencial a partir do 4º dia de cultivo. A fase de crescimento estacionária teve início no 5º dia, enquanto a fase de morte celular no 8º dia.

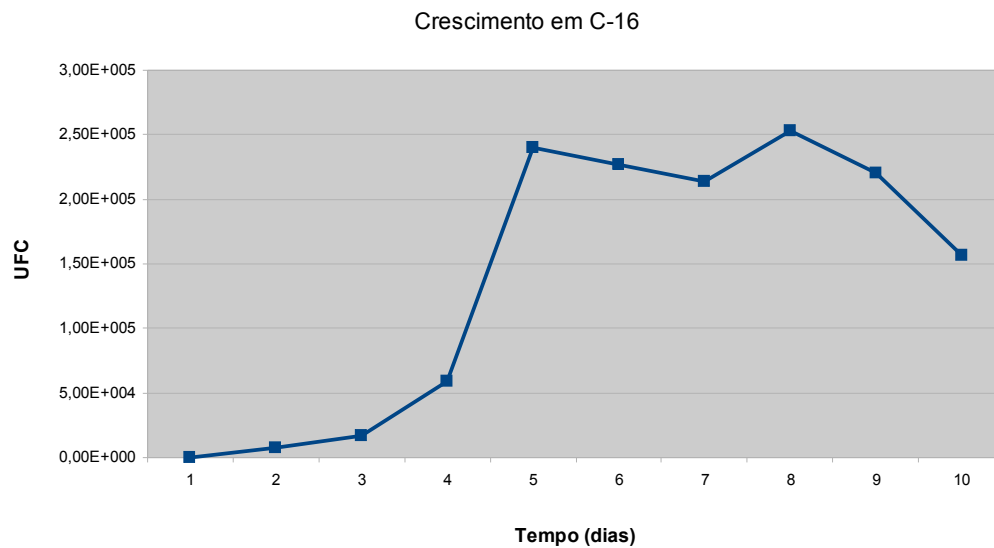


Figura 36: Curva de crescimento celular de *D. cinnamea* P4 em meio líquido BH contendo C-16.

6.21 – Crescimento da estirpe P4 em meio contendo C-16 como fonte de carbono em diferentes condições de cultivo

Já foi descrito anteriormente que a estirpe P4 é capaz de crescer em meios apresentando diferentes valores de pH (von der WEID *et al.*, 2007). Baseado nestes dados, foram realizados experimentos de crescimento em biorreator contendo meios cujos valores de pH foram ajustados para 7 e 9 e as temperaturas para 35°C e 45°C (Figura 37). Quando incubada à temperatura de 35°C, a estirpe P4 apresentou melhor crescimento quando cultivada em pH 9, e um crescimento moderado em pH 7. Por outro lado, quando a estirpe P4 foi crescida em meio contendo glicose na concentração final de 1% como única fonte de carbono (Figura 37, pH 7*) esta apresentou um crescimento mais acentuado quando comparado com o crescimento observado no meio contendo C-16 como única fonte de carbono. A fase de morte foi observada em menor tempo de incubação da P4 em meio com glicose. A estirpe P4 não foi capaz de crescer em pH 4 e 5 nas condições utilizadas (Figura 37).

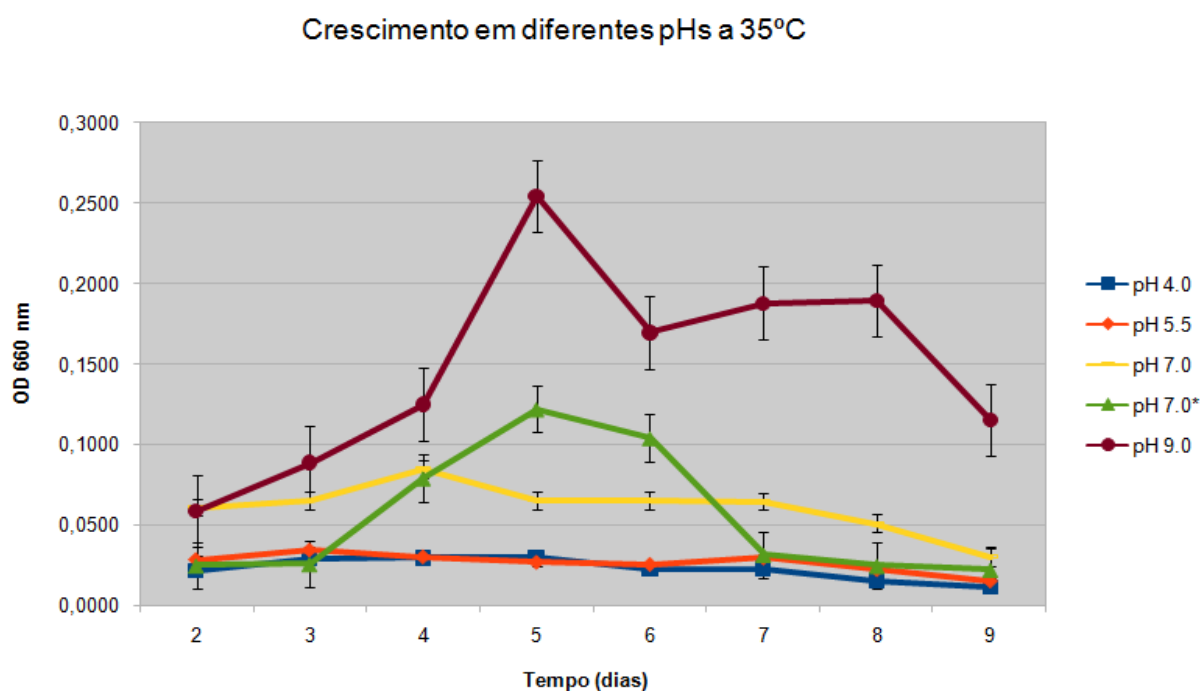


Figura 37: Curvas de crescimento de *D. cinnamea* P4 em meio líquido BH contendo C-16 em biorreator de 2 litros em diferentes pH a 35°C.

As curvas de crescimento resultantes da incubação de P4 a 45°C foram bastante semelhantes às obtidas a 37°C (Figura 38). O pH 9 mostrou ser o melhor pH para o seu crescimento, seguido do pH 7. Novamente, a estirpe P4 não foi capaz de crescer em meio com pH 4 e 5, nas condições testadas (Figura 37).

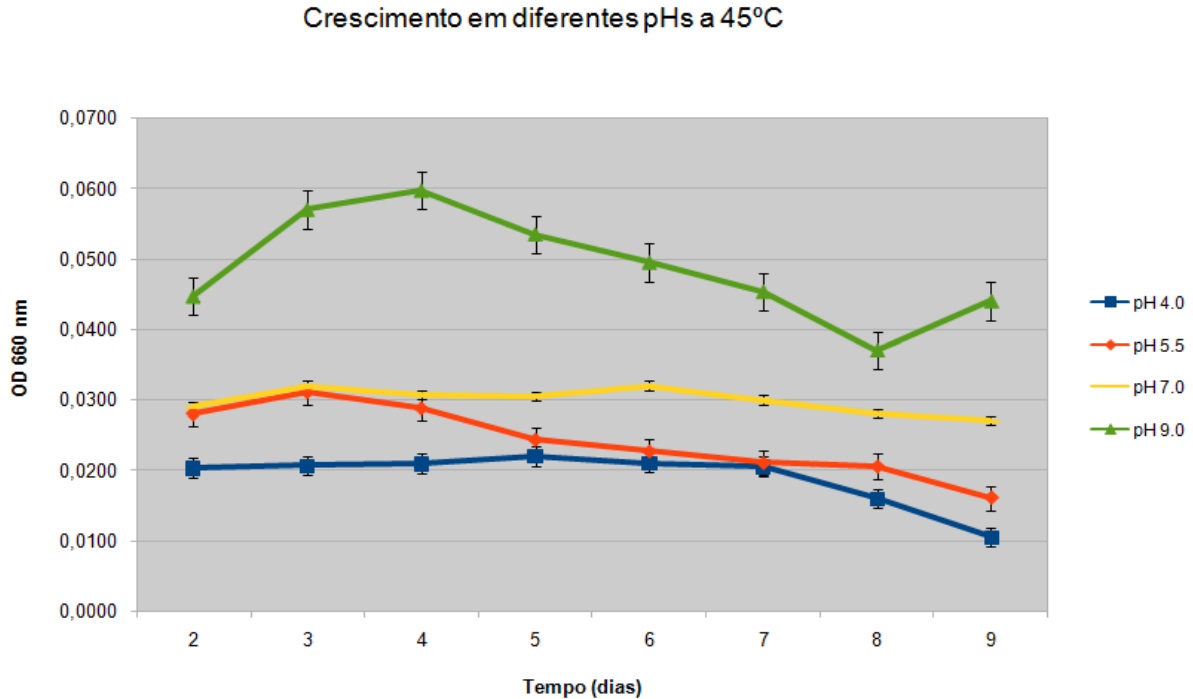


Figura 38: Curvas de crescimento de *D. cinnamea* P4 em meio líquido BH contendo C-16 em biorreator de 2 litros em diferentes pHs a 45°C.

6.22 – Crescimento da estirpe P4 em meio contendo C-16 como fonte de carbono por um longo período de cultivo

Baseado nos resultados obtidos nos diferentes experimentos de crescimento realizados anteriormente (item 6.21) foi escolhida a condição ótima de cultivo de *D. cinnamea* P4 (pH 7 e temperatura de 35°C) para sua manutenção por longo período de cultivo em biorreator de 2 litros. O método de suprimento da fonte de carbono escolhido foi o de batelada em série. Como pode ser observado na figura 39, a estirpe de P4 apresentou um comportamento semelhante aos resultados obtidos anteriormente durante a fase inicial de cultivo: um período de adaptação de aproximadamente quatro dias, seguida de uma fase exponencial de aproximadamente dois dias. Contudo, ao contrário dos experimentos anteriores, não foi observada uma fase estacionária longa e sim uma diminuição do número de células 24 horas após ter sido alcançada a fase estacionária, culminando na grande diminuição da população após 10 dias de incubação. No décimo dia, foi adicionado C-16 na concentração final de 1% (v/v) acarretando um novo crescimento exponencial de aproximadamente dois dias. Este comportamento de declínio e recuperação dos níveis celulares se repetiu mais duas vezes, embora a figura 39 mostre apenas o resultado obtido até o 31º dia.

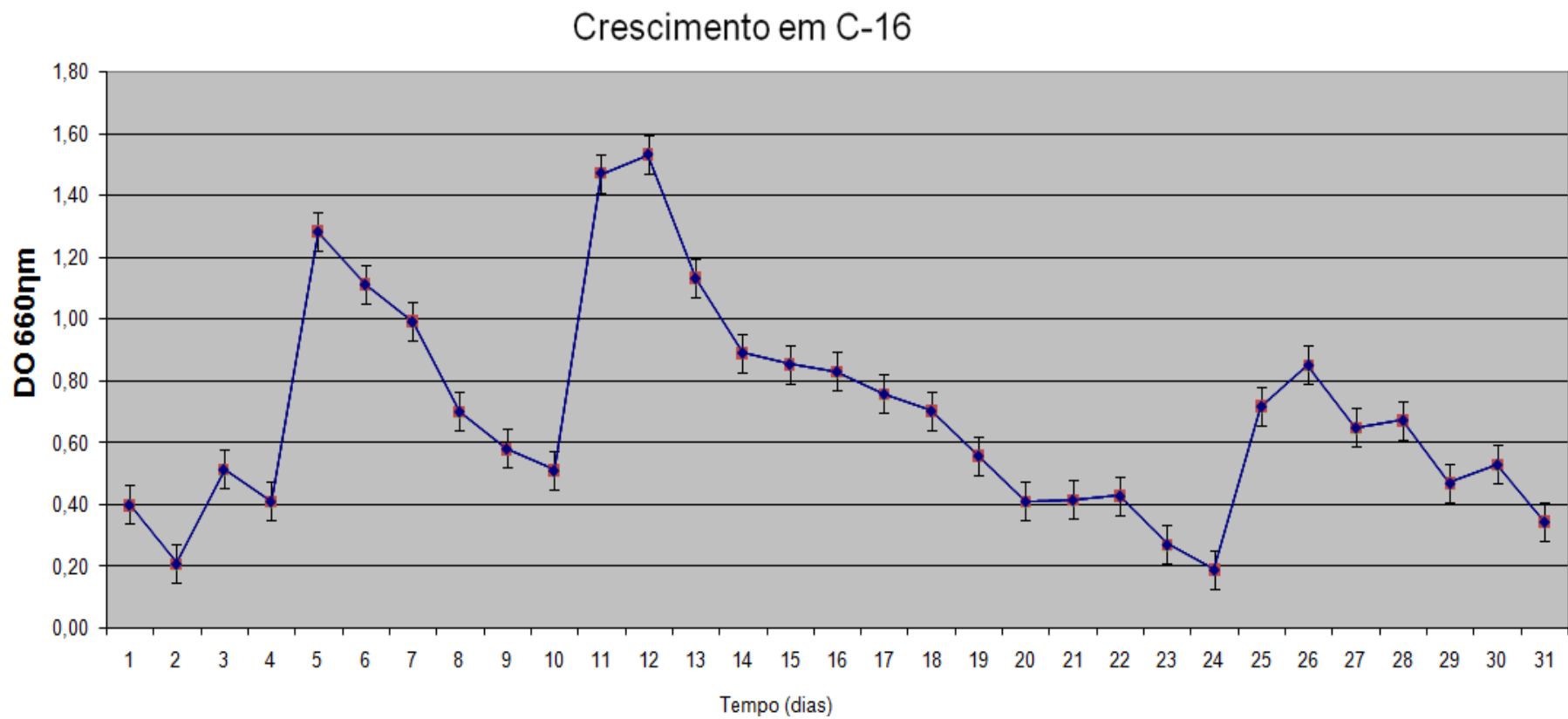


Figura 39: Manutenção do crescimento celular de *D. cinnamea* P4 utilizando C-16 como única fonte de carbono em biorreator de 2 litros.

6.23 – Análise das alterações morfológicas e genômicas

Uma cultura de *D. cinnamea* P4 foi mantida em biorreator por 45 dias, o correspondente a cerca de 300 gerações. Após esse tempo, uma alíquota da cultura foi plaqueada em LB-ágar e as colônias crescidas foram comparadas com as plaqueadas no mesmo meio antes da incubação em biorreator. Não foram observadas diferenças na morfologia e na pigmentação das colônias antes e depois do crescimento em biorreator.

Também foram analisadas possíveis alterações genéticas na cultura de P4 após seu cultivo em biorreator., conforme descrito nos materiais e métodos. Oito colônias foram escolhidas aleatoriamente a partir do crescimento em meio LB-ágar após 45 dias de cultivo em biorreator. Seu DNA foi extraído e amplificado por BOX-PCR. Na figura 40 é apresentado o perfil de amplificação do DNA das 8 colônias, uma do *mix* (amostra referente ao 45º dia, retirada diretamente do biorreator) e de uma colônia de *D. cinnamea* P4 antes da incubação em biorreator (ISW). É possível observar uma diferença no perfil de amplificação da colônia 6 (IS6) e do *mix* quando comparados com o perfil de amplificação de ISW.

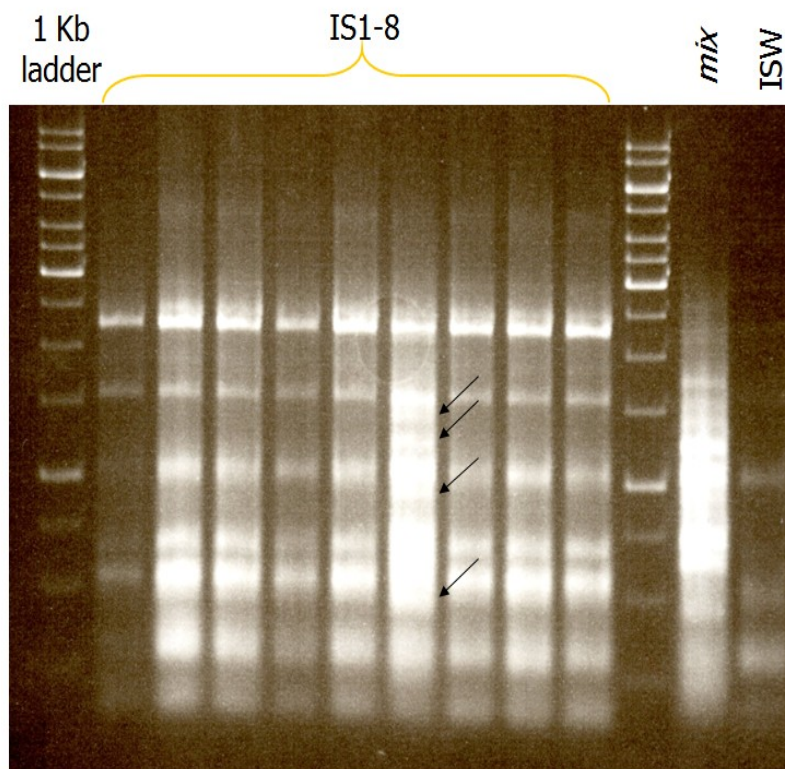


Figura 40: BOX-PCR de oito colônias da estirpe P4 obtidas a partir do experimento de cultivo em biorreator de 2 litros por longo período de cultivo, da amostra referente ao 45º dia, retirada diretamente do biorreator (*mix*) e mais a colônia controle (ISW). As setas indicam as prováveis duplicações gênicas.

Para que fosse confirmada a alteração genética observada em IS6 e no DNA obtido a partir das populações prováveis no final do experimento (*mix*), foram realizadas ampliações do DNA de IS6, *mix* e ISW utilizando quatro concentrações diferentes de DNA como molde, conforme descrito nos materiais e métodos. A figura 41 mostra o resultado desse segundo BOX-PCR confirmando os resultados observados na figura 40, isto é, os DNAs de IS6 e *mix* realmente apresentaram diferenças nos seus perfis de amplificação quando comparados com o perfil obtido para ISW.

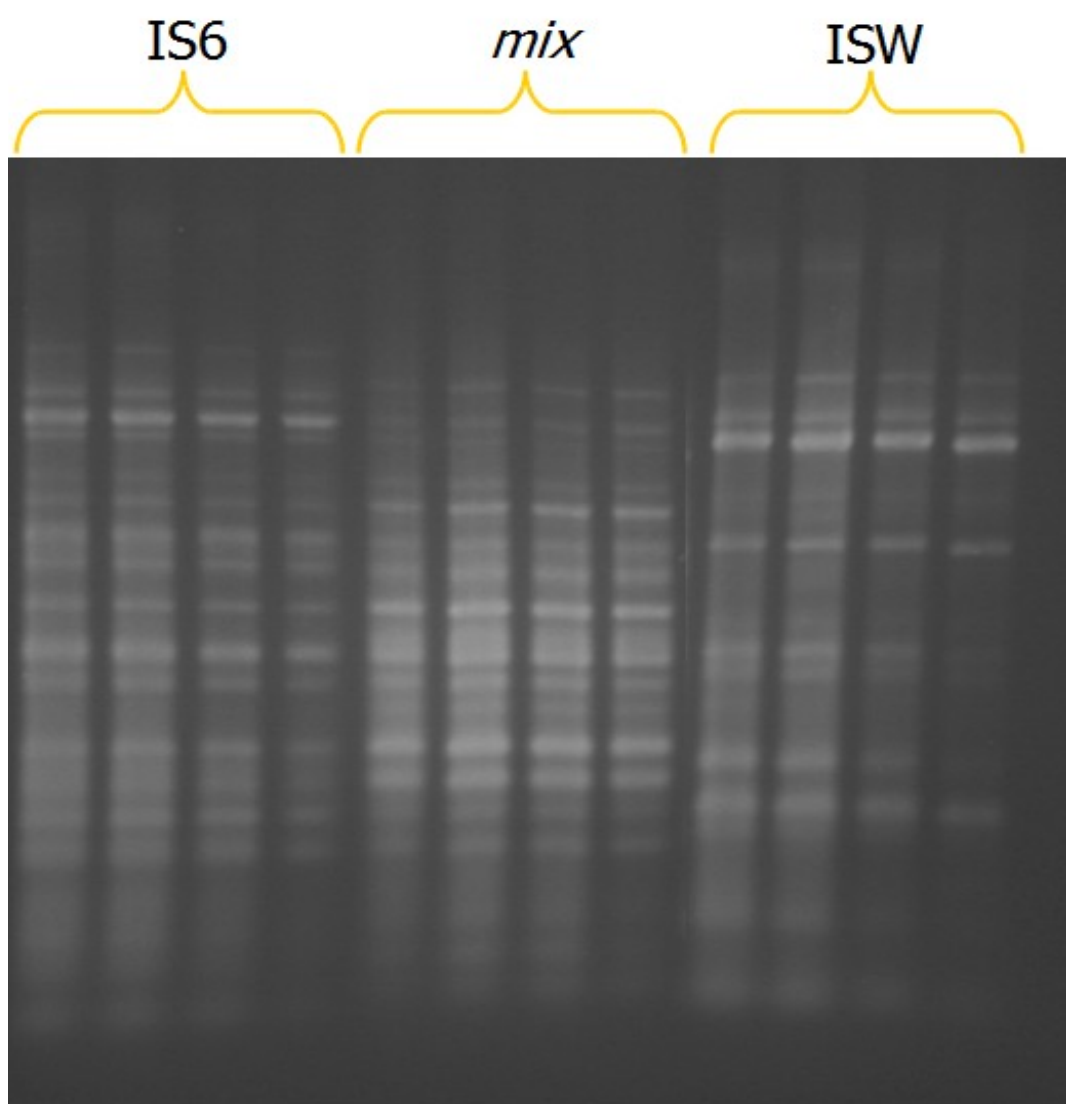


Figura 41: Perfil da amplificação por BOX-PCR utilizando-se quatro concentrações diferentes dos DNAs de IS6, *mix* e ISW. As concentrações finais de DNA genômico dos

isolados foram 10 ng, 1 ng, 0,5 ng e 0,1 ng.

III - Análise dos níveis de expressão dos prováveis genes envolvidos na degradação de *n*-alcanos

6.24 – Análise dos níveis de expressão dos prováveis genes *alkS*, *alkB* e *alkTrans*

Em experimentos de degradação de compostos de petróleo, *D. cinnamea* P4 mostrou ser capaz de degradar uma ampla faixa de *n*-alcanos (C₁₁-C₃₆) sem a presença de outra fonte de carbono (von der WEID *et al.*, 2007). Baseado nesses resultados, foram realizados experimentos de análise dos níveis de expressão dos prováveis genes *alkS*, *alkB* e *alk-trans*, (envolvidos na degradação de *n*-alcanos) através de PCR em tempo real.

Na figura 42, encontram-se apresentadas a relação entre a expressão do provável gene *alkS* de *D. cinnamea* P4 nas fases lag, log e estacionária do crescimento em biorreator de 2 litros quando C-16 ou a glicose foram utilizados como única fonte de carbono. O provável gene *alkS* foi mais expresso durante a fase lag (mais de 5 vezes). Nas fases log e estacionária, sua expressão foi 5 e 2,2 vezes maior que em glicose, respectivamente.

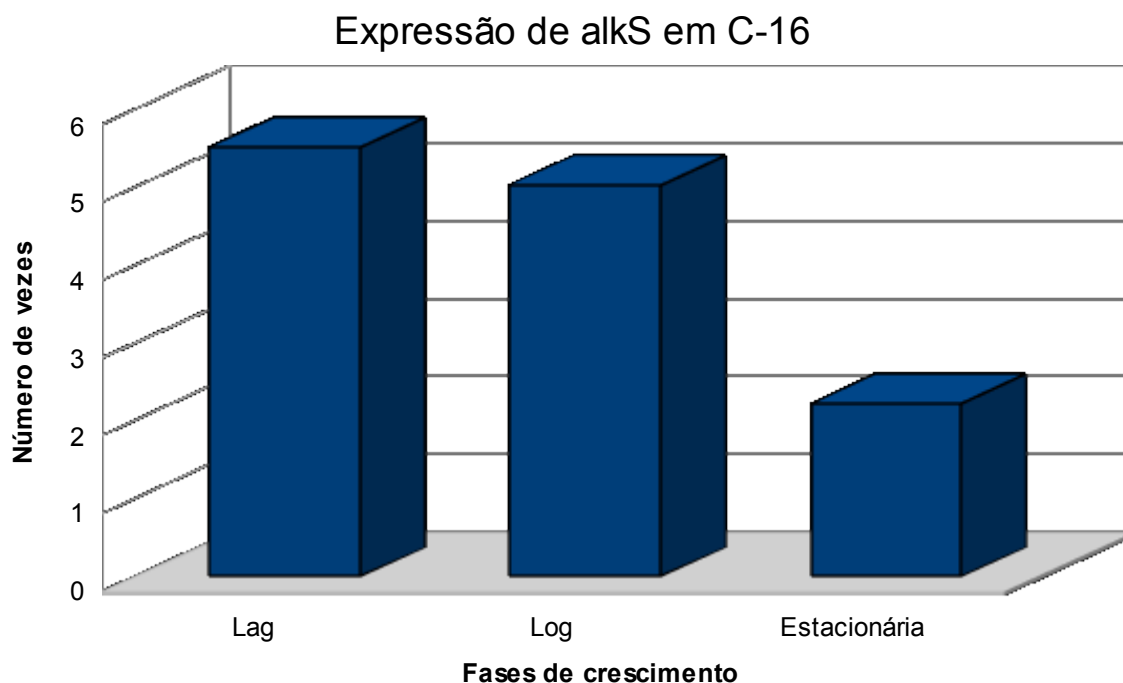


Figura 42: Níveis de expressão do provável gene *alkS* de *D. cinnamea* P4 crescida em C-16, em relação ao nível de expressão durante crescimento em LB.

A relação entre os níveis de expressão do provável gene *alkB* da estirpe P4 foi determinada utilizando-se as mesmas condições descritas para o provável *alkS* e o resultado obtido está apresentado na figura 43. Essa relação foi alta tanto na fase lag quanto log, de 3,8 e 3,7 vezes, respectivamente. Porém, uma forte queda dessa relação no seu nível de expressão foi observada durante a fase estacionária, onde esse valor foi de apenas 3 vezes.

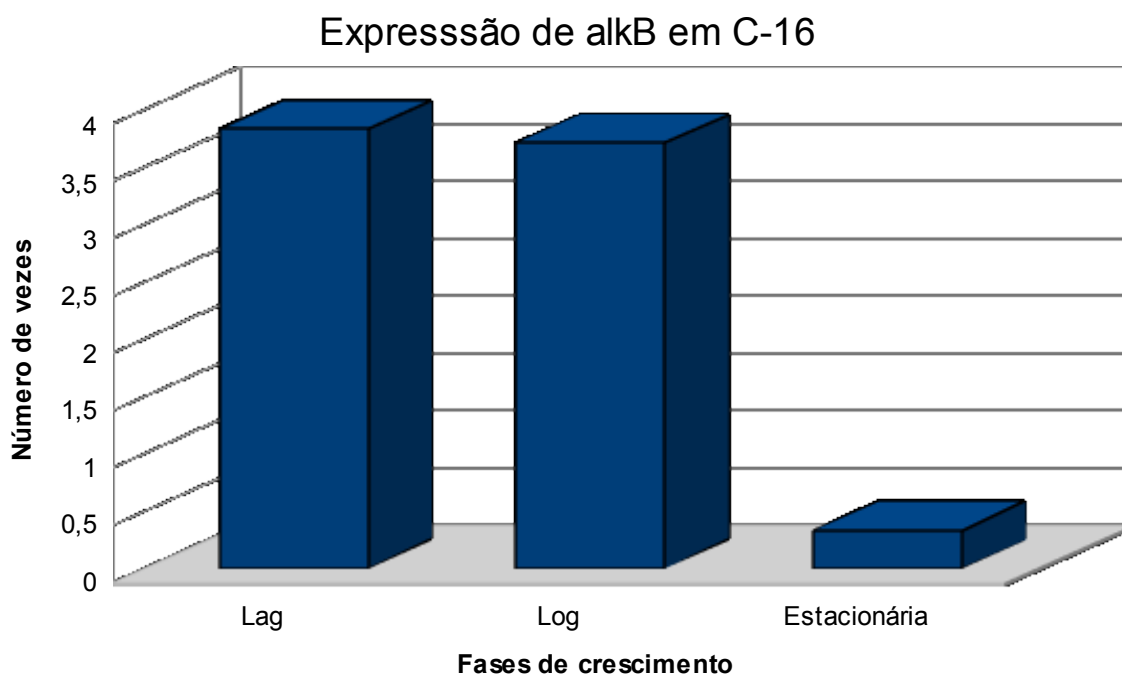


Figura 43: Níveis de expressão do provável gene *alkB* de *D. cinnamea* P4 crescida em C-16, em relação ao nível de expressão durante crescimento em LB.

O provável gene *alkTrans*, identificado juntamente com os genes putativos da via metabólica Alk, também foi utilizado na análise de sua expressão conforme realizado para os genes *alkS* e *alkB*. Os resultados obtidos mostraram-se semelhantes aos do gene putativo *alkB*. Durante as fases lag e log, a expressão do *alkTrans* foi de 3,8 vezes maior, enquanto na fase estacionária esse valor foi de apenas 0,05 vezes em relação ao seu nível de expressão durante o crescimento com glicose (Figura 44).

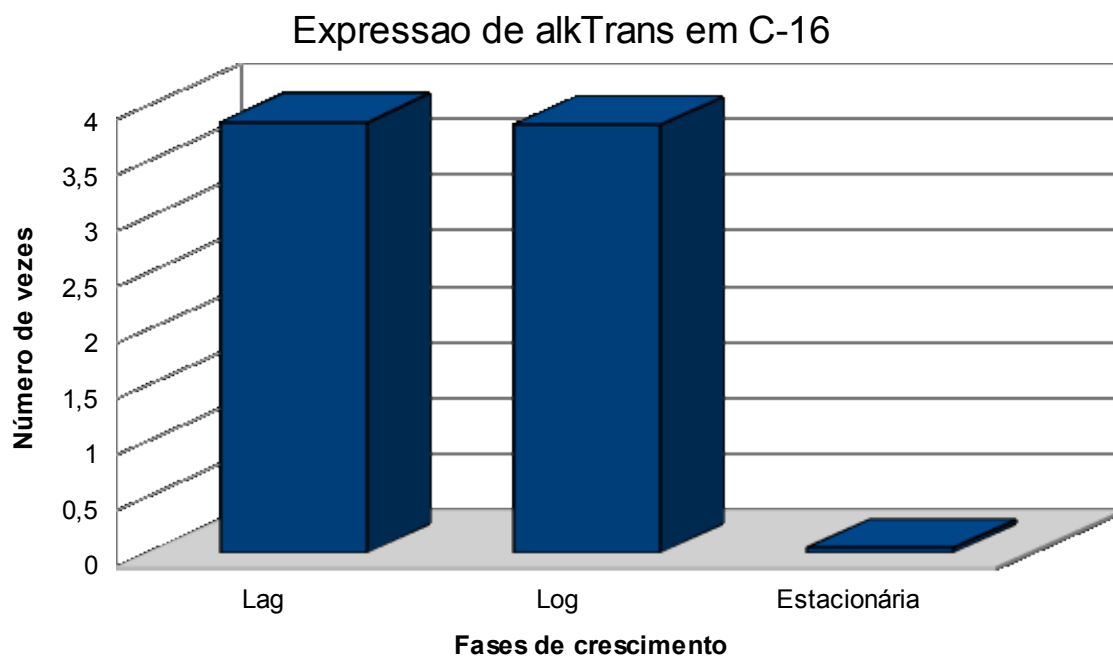


Figura 44: Níveis de expressão do provável gene *alkTrans* de *D. cinnamea* P4 crescida em C-16, em relação ao nível de expressão durante crescimento em LB.

7 - Discussão

Os dados acumulados em bancos de dados, como GenBank NR, EMBL, KEGG e CAZy, associados aos métodos computacionais permitiram não apenas apresentar um conjunto de genes e suas localizações nos diferentes genomas já estudados, mas também estabelecer relações filogenéticas entre diferentes organismos e possíveis predições de estruturas moleculares (KANEHISA & BORK, 2003). Já os programas de busca por similaridade, como o BLAST (ALTSCHUL *et al.*, 1990) e BLASTALL, permitiram a obtenção de respostas mais rápidas, de pareamento a partir de pequenas sequências, de valores estatísticos de alinhamento (e-value), além da possibilidade do uso da interpretação humana ao comparar duas sequências (McGINNIS & MADDEN, 2004).

Os diferentes projetos de sequenciamento que empregam as metodologias disponibilizadas pela terceira geração de sequenciadores normalmente resultam em grandes sequências de *contigs* que, apesar de cobrir a maior porção do genoma, deixam lacunas a serem preenchidas. A este tipo de sequenciamento é dado o nome de *draft genome*. Diversos projetos de sequenciamento de genomas bacterianos empregam na maioria das vezes a técnica de *draft genome* em conjunto com a técnica de sequenciamento automático por capilaridade, para fechar estas lacunas. Contudo, atualmente, cada vez mais os sequenciamentos de genomas microbianos têm utilizado apenas a metodologia de *draft genome*. Isso se deve principalmente aos elevados custos necessários para o sequenciamento automático de Sanger, além da expectativa de que um *draft genome* seja suficiente para responder as questões levantadas pelos projetos de sequenciamento.

A motivação para o sequenciamento do genoma da estirpe P4 de *D. cinnamea* foi, além da busca por novas informações a respeito de seu potencial catabólico de poluentes do petróleo, a tentativa de compreender a relação entre os genes presentes em seu genoma e as características do ambiente em que esta estirpe foi isolada. Desta forma, a metodologia de *draft genome*, escolhida neste projeto, permitiu o melhor entendimento sobre a capacidade de utilização de fontes orgânicas e inorgânicas e até levantar hipóteses sobre esses tipos de metabolismo. Além disso, permitiu fazer suposições a respeito das estratégias empregadas pela estirpe P4 visando sua sobrevivência e competição frente a outros microrganismos presentes no ambiente em que ela foi isolada.

A análise do genoma de *D. cinnamea* P4 confirmou a presença de todas as vias metabólicas centrais, e.g. via glicolítica, gliconeogênese, TCA, metabolismo de ácidos graxos e biossíntese dos aminoácidos e nucleotídeos, através da identificação da maioria dos genes hipotéticos. Também foram encontrados os genes hipotéticos pertencentes à via de biossíntese do aminoácido selenocisteína, componente este de proteínas com atividade redox em reações de destoxificação celular (WU *et al.*, 2005). A presença do ciclo do glioxilato, identificado também no genoma de *D. cinnamea* P4, vem sendo relatada em bactérias, fungos e plantas e esse ciclo é responsável pela utilização de compostos de dois carbonos como fonte para síntese de glicose (LORENZ & FINK, 2001). Esta via é comumente descrita em fungos e bactérias patogênicas, onde desempenha importante papel na aquisição de fonte de carbono durante a infecção em células de macrófagos. Embora a espécie *Dietzia cinnamea* tenha sido anteriormente descrita como patogênica ao homem, o papel provável desta via seria na aquisição de fonte de carbono e energia através de diversos substratos presentes no ambiente. Outras vias descritas no genoma da estirpe P4, como por exemplo, o ciclo do 3-hidroxiopropionato (HERTER *et al.*, 2001) e o da degradação de nitriloacetato (KIM *et al.*, 2006), utilizam o ciclo do glioxilato como intermediário ou produto final em suas vias. A via de β -oxidação descrita no genoma da P4 possui um papel central para diversas outras vias catabólicas secundárias envolvidas na degradação de hidrocarbonetos alifáticos, cíclicos e aromáticos, desta forma possuindo grande importância na fisiologia de bactérias degradadoras de hidrocarbonetos.

Embora a anotação de genes necessariamente não implique na determinação da sua funcionalidade, os projetos de sequenciamento de genomas têm revelado novas informações sobre microrganismos que oxidam CO. A utilização de compostos de CO como fonte de carbono e energia provavelmente vem ocorrendo desde o início da evolução dos microrganismos. Comunidades de bactérias e archaeas quimiolitotróficas habitantes de chaminés oceânicas são comumente descritas como utilizadores de CO como fonte de carbono e energia em condições de anaerobiose (KING & WEBER, 2007). Mais recentemente, a bactéria *Caboxydotherrnus hydrogenoformans* vem sendo empregada como modelo da reação de oxidação de CO, provavelmente utilizada pelos microrganismos ancestrais. A presença da via metabólica de oxidação de CO a CO₂ por *D. cinnamea* P4 reforça a proposição de King e Weber (2007), de que cada vez mais projetos de sequenciamento de genomas microbianos revelam espécies fora do grupo das archaeas e

bactérias extremófilas, como os actinomicetos do solo, como capazes de utilizarem CO como única fonte de carbono e energia. Outros benefícios proporcionados pela anotação genômica é a identificação de novas formas da subunidade catalítica da monóxido de carbono desidrogenase (CODH). Na estirpe P4, foram detectados 2 CDSs com similaridade com a subunidade catalítica de CODH, as quais mostraram-se distintas após análises filogenéticas. Dados da literatura também sugerem a presença de uma segunda cópia da subunidade catalítica, embora ainda não tenha sido determinada sua possível função bioquímica. Na análise do genoma de *Thermomicrobium roseum*, uma bactéria Gram-negativa termofílica descrita como oxidadora de CO, também foram descritas duas prováveis CODH, pertencentes às classes I e II com funções distintas (WU *et al.*, 2009).

A identificação dos genes envolvidos na redução dissimilatória de nitrato durante a anotação do genoma de *D. cinnamea* P4 não era esperada. Isso porque essa via é mais comumente descrita em bactérias denitrificantes, como por exemplo, a bactéria quimiolitotrófica *Nitrobacter winogradskyi* Nb-255 (STARKENBURG *et al.*, 2006). Além disso, o crescimento anaeróbico utilizando nitrato como aceptor de elétrons não foi anteriormente relatado por von der Weid e colaboradores (2007). Entretanto, a identificação de todas as subunidades prováveis que compõem o complexo Nar, aliada a identificação de diversos possíveis transportadores de fontes de compostos nitrogenados, sugerem a possibilidade da estirpe P4 em utilizar a redução dissimilatória de nitrato como fonte de energia alternativa. Situação parecida também é descrita na bactéria *Polaromonas naphthalenivorans* CJ2 que, apesar de possuir todas as enzimas do complexo Nar, não apresenta crescimento em condições de anaerobiose na presença de nitrato (YAGI *et al.*, 2009).

Os dados disponíveis na literatura apontam a prevalência dos mecanismos de autotrofia em bactérias anaeróbicas como, por exemplo, nas bactérias que habitam as chaminés oceânicas. Contudo, os projetos de sequenciamentos têm levado à descrição de novas espécies de actinomicetos de solos possuindo enzimas necessárias para a fixação de CO₂ e vivendo em condições de microaerofilia, algumas vezes utilizando o sulfeto ou hidrogênio como doador de elétrons (ROTHSCHILD, 2008). A anotação do sequenciamento do genoma de *D. cinnamea* P4 permitiu a identificação da maioria dos genes pertencentes à via do ciclo do ácido cítrico reverso (rTCA) e à via do 3-hidroxipropionato, ambas relacionadas à assimilação de compostos C1, como CO₂ e formaldeído. O ciclo rTCA é

encontrado em bactérias e archaeas autotróficas, sendo normalmente descrito junto com o ciclo de *Calvin Benson Bassham*, o qual emprega a reação chave catalisada pela enzima ribulose 1,5-bifosfato carboxilase/oxigenase (Rubisco) (ROTHSCHILD, 2008). A primeira vez descrita em *Chloroflexus aurantiacus* (HERTER *et al.*, 2001), o ciclo do 3-hidroxiacetonato recebeu esta denominação devido à presença de hidroxiacetonato, um composto com três carbonos e intermediário chave nesta via. A fixação de CO₂ por meio do ciclo do hidroxiacetonato é possível através da redução de duas moléculas de CO₂ em glioxilato (HUGLER *et al.*, 2003). Presente principalmente em bactérias anoxigênicas fototróficas, o ciclo do 3-hidroxiacetonato é sugerido como um dos primeiros mecanismos de autotrofia durante a história da evolução (ROTHSCHILD, 2008). Além do gênero *Chloroflexus*, as archaeas hipertermófilas não fototróficas dos gêneros *Metallosphaera*, *Acidianus* e *Sulfolobus* também foram descritas como capazes de utilizar o ciclo do 3-hidroxiacetonato para fixar CO₂. A estirpe *D. cinnamea* P4 é o primeiro actinomiceto descrito como capaz de empregar o ciclo 3-hidroxiacetonato como um mecanismo de fixação de carbono. Semelhante ao gênero *Sulfolobus*, a estirpe P4 provavelmente habita nichos em condições de microaerofilia, o que poderia contribuir em uma adaptação de seu genoma para sobrevivência nestas condições de níveis flutuantes de oxigênio.

A metilotrofia é a capacidade de utilizar compostos C1, ou outros compostos orgânicos desprovidos de ligações C-C, como doadores de elétrons e fontes de carbono. Este processo propicia uma grande economia de energia quando comparado à assimilação de carbono pelos outros meios de autotrofia que utilizam o CO₂ ao invés de compostos orgânicos. Isso se deve principalmente pelo fornecimento do poder redutor de C1, que ocorre através de etapas de oxidações posteriores à mesma via de assimilação de C1. A fixação de compostos de C1 foi inicialmente descrita apenas em espécies metilotróficas. Contudo, abordagens genômicas têm sugerido uma distribuição mais ampla dos genes que participam destas vias metabólicas em organismos não-metilotróficos, como previamente relatado em espécies do gênero *Streptomyces* (HAN & REYNOLDS, 1997). Embora os genes responsáveis pela oxidação de metano a formaldeído não tenham sido detectados, os genes envolvidos nas reações subsequentes para a assimilação de compostos de C1 estão presentes no genoma da estirpe de P4. Não foi possível detectar um transportador de compostos de formaldeído provável, o que deixa em aberto ainda uma possível fonte deste composto. Contudo, a presença de todos os genes hipotéticos responsáveis pela conversão de formaldeído/formato a CO₂ indica a

capacidade de assimilação de compostos de C1 como fonte de carbono e energia pelas células de *D. cinnamea* P4. Além disso, a anotação do genoma da estirpe P4 sugere o redirecionamento de um intermediário da via de conversão de formaldeído a CO₂, o Metileno-H₄F, para o ciclo da serina. Este último, juntamente com os ciclos de regeneração do glioxilato, ciclo da serina e PHB, são descritos como responsáveis pela assimilação de compostos C1 em bactérias metilotróficas do gênero *Methylobacterium* (CHISTOSERDOVA *et al.*, 2004).

Experimentos de crescimento da estirpe P4 em compostos de hidrocarbonetos alifáticos e aromáticos mostraram que essa estirpe possui a capacidade de utilizar *n*-alcanos de diversos tamanhos, além de alguns tipos de hidrocarbonetos aromáticos, como única fonte de carbono (von der WEID *et al.*, 2007). A anotação do genoma de *D. cinnamea* P4 corrobora esta habilidade em utilizar *n*-alcanos como única fonte de carbono e energia, através da presença do *operon* composto pelos prováveis genes alcano monoxigenase (*alkB*), rubredoxina (*rubA*), uma transportador de compostos hidrofóbicos (*alkTrans*) e um provável regulador transcricional envolvido na expressão destes genes (*alkS*). Além disso, a identificação de outras três monoxigenases hipotéticas com homologia com hidroxilases envolvidas na degradação de hidrocarbonetos alifáticos permitiu estabelecer uma relação entre a presença destas monoxigenases putativas e a aptidão da estirpe P4 em utilizar cadeias de *n*-alcanos de diferentes tamanhos. A capacidade de utilização de *n*-alcanos de cadeia longa é sugerida pela presença de genes específicos presentes no genoma de *Geobacillus thermodenitrificans* NG80-2 (FENG *et al.*, 2007). Embora não tenha sido identificada nenhuma provável CDS com homologia com a AlkB hidroxilase, os resultados dos experimentos de cromatografia gasosa confirmaram a conversão de hexadecano a hexadecanol, reação esta descrita como dependente da enzima AlkB. O provável gene que codifica o transportador de compostos hidrofóbicos *alkTrans* foi identificado junto aos genes hipotéticos *alkB*, *rubA* e *alkS* e levanta a questão sobre o seu provável papel na aquisição de compostos de *n*-alcanos pelas células de *D. cinnamea* P4. Embora a existência de um provável transportador de *n*-alcanos ainda não tenha sido confirmada na literatura, Vaysse e colaboradores (2009) demonstraram algumas evidências da presença de uma provável proteína desempenhando este papel em uma análise proteômica da bactéria *Marinobacter hydrocarbonoclasticus* SP17. Dados obtidos no presente trabalho, através de experimentos de PCR em tempo real e utilizando os hipotéticos genes *alkB*, *alkTrans* e *alkS* de *D. cinnamea*

P4, corroboram a participação das prováveis enzimas codificadas por estes genes na via metabólica de degradação de *n*-alcanos.

Os hidrocarbonetos aromáticos são amplamente distribuídos na natureza, podendo ser de origem de material orgânico em decomposição, fontes antropológicas como pirólise de material vegetal ou combustíveis fósseis, mas sua principal fonte é o petróleo. Entre os hidrocarbonetos aromáticos, os compostos de bifenil e benzeno são um grande problema ambiental, devido a sua estabilidade química, baixa solubilidade em água e alta toxicidade (van HAMME, SINGH & WARD, 2003). Apesar de von der Weid e colaboradores (2007) terem sugerido a capacidade da estirpe P4 em degradar compostos de tolueno, xileno e naftaleno, não foi possível detectar no presente estudo as prováveis enzimas responsáveis pelas reações de hidroxilação que ocorrem durante a fase inicial de degradação destes compostos. Também não foram detectados os genes hipotéticos que codificam a enzima bifenil dioxigenase, embora tenham sido identificados todos os prováveis genes subsequentes que codificam as proteínas que constituem as vias superior bifenil, degradação de benzoato via ligação de CoA e via central β -cetoadipato. Duas possíveis vias de catabolismo diferentes para compostos de bifenil também foi descrita no acinomiceto *R. jostii* RHA1 (McLEOD *et al.*, 2006) e na bactéria Gram-negativa *Burkholderia xenovorans* LB400 (CHAIN *et al.*, 2006). O sequenciamento do genoma de *D. cinnamea* P4 revelou a presença de prováveis enzimas envolvidas na degradação de compostos de ciclohexanona, um hidrocarboneto cíclico. A identificação da provável via catabólica responsável em converter ciclohexanona em ácido adípico sugere a possibilidade do emprego da estirpe P4 em processos biotecnológicos, uma vez que Thomas e Raja (2005) relatam a aplicação da via metabólica de ciclohexanona na biossíntese de náilon.

A anotação do genoma de *D. cinnamea* P4 revelou a presença de diversos genes hipotéticos envolvidos na degradação de compostos nitrogenados, como nitrilas, nitriloacetato e nitropropano. Os compostos de nitrilas possuem a fórmula geral R-CN e são amplamente distribuídos na natureza. Sua principal origem é a partir dos cianoglicosídeos produzidos pelas plantas, mas também podem ser oriundos de indústrias químicas, através da manufatura de uma grande variedade de polímeros (como solventes e herbicidas aromáticos) e, desta forma, podendo ser um grave problema ambiental (BRENNER, 2002). As enzimas da família nitrilase/cianida hidratase encontradas na estirpe *Rhodococcus rhodocorus* M8 apresentaram similaridade com as CDSs da estirpe P4 e possuem aplicação biotecnológica em potencial na

produção de acrilamida e nicotinamida (KOBAYASHI *et al.*, 1993). Além disso, essas enzimas apresentam um papel relevante na remediação ambiental, onde elas têm sido empregadas na remoção de compostos de nitrilas durante o tratamento de águas (WYATT & KNOWLES, 1995). Outra via de importância na biorremediação de compostos nitrogenados cujos genes foram encontrados no genoma da estirpe P4 é a via de degradação de compostos de aminoácido carboxílico nitriloacetato (NTA). A enzima chave desta via é a nitriloacetato monoxigenase, amplamente distribuída entre as proteobactérias e os actinomicetos (EGLI, 2001). Os compostos de nitriloacetato são empregados em processos de remediação de radionucleotídeos, atuando como agente quelante de íons metais. Contudo, a intensa utilização destes compostos tem provocado inúmeras contaminações de aquíferos e solos (BUCHELI-WITSCHER & EGLI, 2001). Finalmente, a presença da provável nitropropano dioxigenase, representada no genoma de *D. cinnamea* P4 com cinco prováveis CDSs, também sugere sua possível aplicação em processos de remediação. Os compostos de nitroalcanos são amplamente distribuídos no ambiente devido a sua origem natural e antropogênica. Nitroalcanos são produzidos por determinados microrganismos e algumas leguminosas, mas também são encontrados em combustíveis de foguetes, solventes e de materiais expelidos da combustão interna de máquinas e desta forma, causando diversos danos ao ambiente. A propriedade de NAO em catalizar a oxidação de nitroalcanos inertes quimicamente é única entre as flavoproteínas, o que tem despertado o interesse na indústria de processos de biocatálise na utilização da enzima NAO na biossíntese de fármacos.

A hidroxilação de compostos de esteróides é de grande interesse para a indústria farmacêutica pela sua aplicação na síntese de corticóides como, por exemplo, o composto 9- α -fluorohidrocortisona (PETRUSMA, DIJKHUIZEN & GEIZE, 2009). Diferentes microrganismos são amplamente empregados em reações de hidroxilações de esteróides, principalmente os actinomicetos dos gêneros *Rhodococcus* (PETRUSMA, DIJKHUIZEN & GEIZE, 2009; van der GEIZE *et al.*, 2008), *Mycobacterium* (ANDOR *et al.*, 2009; ARNELL *et al.*, 2007) e *Arthrobacter* (STRIJEWSKI, 1982). Aparentemente, *D. cinnamea* P4 possui duas prováveis vias de degradação de compostos de colesterol. Uma via cujas enzimas são codificadas pelos genes chave *kshAB* e *kstD* hipotéticos, e uma segunda via composta por duas colesterol oxidases diferentes. A primeira via de degradação de colesterol descrita no genoma da estirpe P4 utiliza a reação de inserção de um átomo de oxigênio no anel B da molécula de 3-cetoesteróide e a abertura do anel A pelas enzimas KshAB. Esta via é descrita

no genoma de outros actinomicetos como *R. jostii* RHA1, *Rhodococcus erythropolis* SQ1 e *Mycobacterium tuberculosis* H37Rv (McLEOD *et al.*, 2006; KNOL *et al.*, 2008). Outra via catabólica de colesterol possível envolve a ação de duas enzimas colesterol oxidase distintas entre si. A classe I da enzima colesterol oxidase possui como principais representantes bactérias Gram-positivas dos gêneros *Rhodococcus*, *Mycobacterium* e *Streptomyces*. A literatura sugere após experimentos de filogenia e da determinação da composição dos resíduos de aminoácidos que as enzimas colesterol oxidase da subclasse I-1 possuem uma sequência sinal presente que as direcionam para uma localização ligada à face citoplasmática da membrana celular, enquanto que nas sequências I-2 esta sequência é ausente, sugerindo uma provável localização citoplasmática (NAVAS *et al.*, 2001).

A anotação do genoma de *D. cinnamea* P4 identificou 64 prováveis oxigenases presentes no seu genoma. As estirpes pertencentes às espécies *R. jostii* RHA1 e *B. xenovorans* LB400, descritas como detentoras de um grande número de oxigenases e como degradadoras de diversos compostos recalcitrantes, apresentam 203 e 134 oxigenases, respectivamente (CHAIN *et al.*, 2006; McLEOD *et al.*, 2006). Determinando-se a relação entre o número de oxigenases por tamanho do genoma, chegamos a um resultado de cerca de 21 oxigenases por Mb no genoma de *R. jostii* RHA1 e 13,8 oxigenases no genoma de *B. xenovorans* LB400, enquanto o genoma da estirpe P4 apresenta cerca de 19,1 oxigenases por Mb. Desta forma, este resultado revela um elevado número de oxigenases por Mb na estirpe P4. Entre as 67 oxigenases, nove pertencem à classe citocromo P450, o que resulta em um média de aproximadamente 2,5 P450 por Mb, valor semelhante ao descrito no genoma de *R. jostii* RHA1 (McLEOD *et al.*, 2006). Oxigenases são enzimas que catalisam a inserção de um ou dois átomos de oxigênio em um substrato orgânico, utilizando o oxigênio molecular como doador do átomo de oxigênio. Existem duas classes de oxigenases: enzimas que catalisam a inserção de um átomo de oxigênio, chamadas de monoxigenases e as que catalisam a inserção de dois átomos de oxigênio, denominadas de dioxigenases. Em *D. cinnamea* P4, além das prováveis oxigenases envolvidas na degradação de compostos de hidrocarbonetos, compostos nitrogenados e colesterol, também foram descritas enzimas hipotéticas envolvidas na síntese de diversos produtos de interesse industrial como, por exemplo, antibióticos, compostos de terpenóides e vitamina B6. Do mesmo modo que em *D. cinnamea* P4, o genoma de *Nocardia farcinica* IFM 10152 também apresenta uma grande diversidade de oxigenases, como alcanos monoxigenase, nitropropano dioxigenase, citocromo P450 (envolvida no metabolismo de

ácido graxos), dihidrofenolacetato dioxigenase e 14 α -esterol demetilase (ISHIKAWA *et al.*, 2004).

A presença de 25 prováveis enzimas envolvidas na assimilação de carbono a partir de compostos presentes na parede vegetal como fonte de carbono e energia sugere um possível estilo de vida saprófito pela estirpe *D. cinnamea* P4. Projetos de sequenciamento de bactérias isoladas de ambientes de solo têm revelado um elevado número de enzimas hidrolíticas de carboidratos. O genoma de *Thermobifida fusca* YX revelou um total de 45 prováveis enzimas que atuam na degradação de compostos oligo- e polissacarídeos de plantas (LYKIDIS *et al.*, 2007), enquanto que o sequenciamento da bactéria *Bacillus licheniformis* ATCC 15580 revelou cerca de 82 enzimas extracelulares, muitas delas com possíveis funções de degradação de compostos de origem vegetal (REY *et al.*, 2004). As duas estirpes são descritas como bactérias com aplicações industriais, na produção de enzimas com atividades celulolíticas. Além disso, de modo semelhante a *T. fusca* YX, a estirpe P4 apresenta as prováveis proteínas que compõem o sistema de secreção Sec e Tat, que possuem papel na exportação de enzimas com capacidade celulolítica. Apesar de *D. cinnamea* P4 apresentar um número menor de enzimas hidrolíticas de carboidratos, quando comparada com outras bactérias descritas como celulolíticas, sua capacidade de crescimento em temperaturas próximas de 40°C e em uma ampla faixa de pH, entre 7 e 9, indica uma possível aplicação neste campo da biotecnologia.

A maioria do conhecimento existente atualmente sobre metabolismo secundário vem de estudos realizados com espécies do gênero *Streptomyces*, enquanto que em outros actinomicetos os dados são principalmente oriundos de anotações de genomas, como por exemplo, em *Mycobacterium marinum* (STINEAR *et al.*, 2008) e *R. jostii* RHA1 (McLEOD *et al.*, 2006). O sequenciamento do genoma de *D. cinnamea* P4 revelou diversos prováveis genes envolvidos em metabolismo secundário. Além das vias de síntese de compostos terpenóides, β -caroteno e ramanose foram identificados genes policetídeos sintases hipotéticos (PKS), os quais estão envolvidos em vias biossintéticas de produtos naturais. Produtos naturais pertencem a uma classe de policetídeos com diversas aplicações na indústria farmacêutica, como antibióticos e antifúngicos, imunossuppressores e anti-tumores. As enzimas responsáveis pela síntese destes produtos, as proteínas PKS, catalisam reações de condensações descarboxilativas a partir de um metabólito simples como acetato, malonato, propionato e seus derivados em cadeias crescentes de policetídeos (STHAPIT *et al.*, 2004). As

proteínas PKS são divididas em três classes de enzimas, PKS tipo I, II e III (CHAN *et al.*, 2009). A classe PKS tipo I são geralmente subdivididas em dois grupos, modular e iterativo. O grupo modular, o mais comum em bactérias, é caracterizado por reações de incorporação de diferentes precursores a uma cadeia crescente de policetídeo. O grupo iterativo, ao invés de adicionar vários módulos a uma cadeia crescente, é utilizado um único policetídeo repetitivamente em reações de condensação para gerar uma molécula final. Um exemplo em *D. cinnamea* P4 de uma provável enzima do grupo iterativo é o putativo gene *aviGT4*, responsável pela síntese do antibiótico Avilamicina A (WEITNAUER *et al.*, 2001). Outro exemplo de metabolismo catalisado por enzimas do grupo iterativo que foi observado na anotação do genoma da estirpe P4 é a síntese do potente anti-tumor Azinomicina B, que envolve a enzima AziB, uma metiltransferase e o citocromo P450 (ZHAO *et al.*, 2008).

No genoma de *D. cinnamea* P4, o valor provável de 9% das CDSs direcionadas para o transporte transmembrânico é semelhante ao descrito para outros genomas bacterianos. Em *G. thermodenitrificans* NG80-2, cerca de 10% do seu genoma estão envolvidos no transporte de solutos através de membranas, das quais 368 prováveis CDSs, 193 utilizam o transporte ativo, com gasto de ATP para o transporte do soluto e uso do sistema de transporte ABC. No genoma da estirpe P4, 151 CDSs apresentaram homologia a proteínas transportadoras transmembrânicas que necessitam de ATP. A maioria delas está envolvida na destoxificação celular através do efluxo de drogas e metais pesados. O actinomiceto de solo *K. rhizophila* também envolve cerca de 9% de seus genes no transporte transmembrânico, sendo que destes 3,5% são transportadores que gastam ATP (TAKARADA *et al.*, 2008). Em *R. jostii* RHA1, dos 890 transportadores pelo menos 78 são dependentes de ATP para seu funcionamento (McLEOD *et al.*, 2006). Já no actinomiceto *Thermobifida fusca* YX, apenas 6% das CDSs totais apresentou homologia com transportadores transmembrânico, sendo a maioria envolvida no transporte de carboidratos e condizendo com seu estilo de vida saprófito (LYKIDIS *et al.*, 2007). A anotação do genoma da estirpe P4 confirmou a presença de 14 prováveis CDSs com funções de aquisição de compostos de carboidratos do meio extracelular, tais como xilose, ribose, maltose e outros carboidratos. Na estirpe P4, a classe MFS representa 12% das proteínas transportadores, das quais cerca de 30% são envolvidas no efluxo de drogas. No genoma do actinomiceto *R. jostii* RHA1, 37% dos transportadores MFS atuam no efluxo de drogas (McLEOD *et al.*, 2006), enquanto no genoma de *K. rhizophila* 31% o fazem (TAKARADA *et al.*, 2008).

Uma importante classe de transportadores identificada no genoma de *D. cinnamea* P4, o sistema Sec, é similar ao descrito em *M. tuberculosis*. O sistema Sec se apresenta com duas cópias do gene *secA* com funções distintas, uma delas envolvida no processo infeccioso de *M. tuberculosis* em células de macrófagos (BRAUNSTEIN *et al.*, 2003). Como a bactéria *D. cinnamea* P4 foi isolada em um ambiente de solo, ainda não foi possível determinar o papel fisiológico da provável proteína SecA2 baseado no seu contexto ambiental. Outro sistema de secreção de proteínas descrito no genoma da estirpe P4 é o sistema de secreção tipo *Tat independente de Sec*. Esse sistema está envolvido no transporte de proteínas em sua conformação final, como enzimas com ações celulolíticas para o meio extracelular. Este sistema também foi descrito em outros actinomicetos de solo, como *K. rhizophila* (TAKARADA *et al.*, 2008) e *T. fusca* YX (LYKIDIS *et al.*, 2007), os quais possuem um elevado número de enzimas celulolíticas.

O nicho onde *D. cinnamea* P4 foi isolada provavelmente apresenta características de aerobiose e anaerobiose/microaerobiose, presença de ácidos húmicos de origem vegetal, metais pesados, exposição a radiação UV e estresses osmótico, os quais exercem grande impacto sobre sua fisiologia, além de provocarem possíveis danos à parede celular e ao DNA. Pelo que sugere a anotação do genoma, a estirpe P4 desenvolveu um elaborado sistema de tolerância e reparo a diversos estresses ambientais. Na estirpe P4, as respostas a estresses oxidativos empregam reações de dismutação do superóxido em O₂ e H₂O₂ e decomposição de H₂O₂ a H₂ e reações redox, empregando as enzimas glutathione e tioredoxinas. O genoma do actinomiceto *Kineococcus radiotolerans* SRS30216, isolado de um sítio contaminado por dejetos nucleares, apresenta diversas enzimas de destoxificação de espécies reativas de oxigênio semelhantes as descritas no genoma da estirpe P4, como catalases, peroxidases e superóxidos dismutase (BAGWELL *et al.*, 2008). A análise do genoma de *N. farcinica* IFM 10152 identificou quatro diferentes catalases, duas superóxidos dismutase (SodC e SodF) e uma alqui peroxidase, além dos genes que compõem o complexo nitrato redutase Nar. Estes dados sugerem um papel na sobrevivência de *N. farcinica* IFM 10152 no interior de células de macrófagos durante uma infecção (ISHIKAWA *et al.*, 2004). De modo semelhante, em *D. cinnamea* P4, além das enzimas envolvidas na destoxificação de espécies reativas de oxigênio, também foram descritas as prováveis enzimas que compõem o complexo nitrato redutase. Os principais sistemas de reparo de DNA induzidos por radiação UV, Rec e Ruv, identificados no genoma da estirpe P4 são homólogos aos descritos no genoma de *K.*

radiotolerans SRS30216 e de *T. fusca* YX, (LYKIDIS *et al.*, 2007; BAGWELL *et al.*, 2008).

Os resultados de crescimento e manutenção celular de *D. cinnamea* P4 em diferentes condições de biorreator abrem novas possibilidades de aplicações desta estirpe em processos fermentativos. Isso porque as indústrias de bioprocessos buscam por organismos capazes de realizar reações na interface solvente/aquosa e em condições mais extremas, como altas temperaturas e valor de pH mais elevado. Baseado nestes resultados, esperamos padronizar os processos fermentativos em diversas condições de crescimento utilizando a estirpe P4 como veículo de biotransformação de produtos de interesse econômico.

8- Conclusões

A anotação do sequenciamento do genoma de *D. cinnamea* P4 confirmou a presença de todas as vias centrais comumente presentes em bactérias.

A identificação dos putativos genes responsáveis pela oxidação de monóxido de carbono e da via dissimilatória de nitrato sugere uma via alternativa na produção de energia.

A provável presença dos putativos genes que permitem a assimilação de fontes de carbono através de compostos de um único C, como CO₂ e formaldeído indica uma provável capacidade de autotrofismo.

A identificação das vias hipotéticas de degradação de compostos de hidrocarbonetos alifáticos, cíclicos e aromáticos aliado aos resultados de expressão diferencial dos genes da via Alk, além dos genes responsáveis pela degradação de poluentes nitrogenados confirmam seu potencial para aplicações em processos de biorremediação.

As vias metabólicas de colesterol presentes no genoma indicam uma possível aplicação biotecnológica, na síntese de produtos de interesse farmacêutico.

O elevado número de enzimas celulolíticas e oxigenases diferentes presentes no genoma de P4, algumas delas sem função ainda descrita na literatura científica, possibilitam a exploração destas enzimas em aplicações biotecnológicas.

O metabolismo secundário identificado pela anotação do genoma identificou uma possível síntese de antibióticos e do potente anti-tumor Azinomicina B.

A viabilidade de crescimento e a manutenção celular da estirpe P4 por longo período de incubação em biorreatores sugerem sua aplicação em processos fermentativos.

9- Referências bibliográficas

1. AL-SALEH E. & OBUEKWE C. 2005. Inhibition of hydrocarbon bioremediation by lead in a crude oil-contaminated soil. *Inter biodeterioration and biodegradation*. 56:1-7.
2. AL-TAHHAN R.A., SANDRIN T.R., BODOUR A.A., MAIER R.M. 2000. Rhamnolipid-induced removal of lipopolysaccharide from *Pseudomonas aeruginosa*: effect on cell surface properties and interaction with hydrophobic substrates. *Appl Environ Microbiol*. 66(8):3262-68.
3. ALTSCHUL S.F., GISH W., MILLER W., MYERS E.W., LIPMAN D.J. 1990. Basic local alignment search tool. *J Mol Biol*. 215(3):403-10.
4. ANDOR A., JEKKELE A., HOPWOOD D.A., JEANPLONG F., ILKOY E., KÓNYA A., KURUCZ I., AMBRUS G. 2009. Generation of useful insertionally blocked sterol degradation pathway mutants of fast-growing mycobacteria and cloning, characterization, and expression of the terminal oxygenase of the 3-ketosteroid 9 α -hydroxylase in *Mycobacterium smegmatis* mc(2)155. *Appl Environ Microbiol*. 72(10):6554-9.
5. ANSORGE W.J. 2009. Next-generation DNA sequencing techniques. *N Biotechnol*. 25(4):195-203.
6. ARNELL R., JOHANNISSON R., LINDHOLM J., FORNSTEDT T., ERSSON B., BALLAGI A., CALDWELL K. 2007. Biotechnological approach to the synthesis of 9 α -hydroxylated steroids. *Prep Biochem Biotechnol*. 37(4):309-21.
7. ATLAS R.M. & PHILP J. 2005. Bioremediation: Applied Microbial Solutions for Real-World Environmental Cleanup. Washington, DC.
8. ATLAS R.M. 1981. Microbial degradation of petroleum hydrocarbons: an environmental perspective. *Microbiol Rev*. 45(1):180-209.
9. BAGWELL C.E., BHAT S., HAWKINS G.M., SMITH B.W., BISWAS T., HOOVER T.R., SAUNDERS E., HAN C.S., TSODIKOV O.V., SHIMKETS L.J. 2008. Survival in nuclear waste, extreme resistance, and potential applications gleaned from the genome sequence of *Kineococcus radiotolerans* SRS30216. *PLoS One*. 3(12):e3878.
10. BALBA M.T., AL-AWANDHI N. & AL-DAHER R. 1998. Bioremediation of oil-

contaminated soil: microbiological methods for feasibility assessment and field evaluation. *J. Microbiol. Methods*. 32:155-64.

11. BAXEVANIS A.D. & QUELLETTE B.F. 2001. Bioinformatics: a practical guide to the analysis of genes and proteins. (2nd edition), Wiley-Liss, Inc.
12. BEARD T.M. & PAGE M.I. 1998. Enantioselective biotransformations using rhodococci. *Antonie Van Leeuwenhoek*. 74(1-3):99-106.
13. BOELS I.C., BEERTHUYZEN M.M., KOSTERS M.H., VAN KAAUWEN M.P., KLEEREBEZEM M., DE VOS W.M. 2004. Identification and functional characterization of the *Lactococcus lactis* rfb operon, required for dTDP-rhamnose Biosynthesis. *J Bacteriol*. 186(5):1239-48.
14. BRASLAVSK Y.I., HEBERT B., KARTALOV E., QUAKE S.R. 2003. Sequence information can be obtained from single DNA molecules. *Proc Natl Acad Sci U S A*. 100(7):3960-64.
15. BRAUNSTEIN M., ESPINOSA B.J., CHAN J., BELISLE J.T., JACOBS J.R. 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. *Mol Microbiol*. 48(2):453-64.
16. BRENNER C. 2002. Catalysis in the nitrilase superfamily. *Curr Opin Struct Biol*. 12(6):775-82.
17. BRITO E.M., GUYONEAUD R., GONI-URRIZA M., RANCHOU-PEYRUSE A., VERBAERE A., CRAPEZ M.A., WASSERMAN J.C., DURAN R. 2006. Characterization of hydrocarbonoclastic bacterial communities from mangrove sediments in Guanabara Bay, Brazil. *Res Microbiol*. 157(8):752-62.
18. BRÜNKE P., STERNER O., BAILEY J.E., MINAS W. 2001. Heterologous expression of the naphthocyclinone hydroxylase gene from *Streptomyces arenae* for production of novel hybrid polyketides. *Antonie Van Leeuwenhoek*. 79(3-4):235-45.
19. BUCHELI-WITSCHERL M. & EGLI T. 2001. Environmental fate and microbial degradation of aminopolycarboxylic acids. *FEMS Microbiol Rev*. 25(1):69-106.
20. CHAILLAN F., LE FLECHE A., BURY E., PHANTAVONG Y.H., GRIMONT P., SALIOT A., OUDOT J. 2004. Identification and biodegradation potential of tropical aerobic hydrocarbon-degrading microorganisms. *Res Microbiol*. 155(7):587-95.

21. CHAIN P.S., DENEV V.J., KONSTANTINIDIS K.T., VERGEZ L.M., AGULLÓ L., REYES V.L., HAUSER L., CÓRDOVA M., GÓMEZ L., GONZÁLEZ M., LAND M., LAO V., LARIMER F., LIPUMA J.J., MAHENTHIRALINGAM E., MALFATTI S.A., MARX C.J., PARNELL J.J., RAMETTE A., RICHARDSON P., SEEGER M., SMITH D., SPILKER T., SUL W..J, TSOI T.V., ULRICH L.E., ZHULIN I.B., TIEDJE J.M. 2006. *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. *Proc Natl Acad Sci U S A*. 103(42):15280-87.
22. CHENG Q., THOMAS S.M., KOSTICHKA K., VALENTINE J.R., NAGARAJAN V. 2000. Genetic analysis of a gene cluster for cyclohexanol oxidation in *Acinetobacter* sp. Strain SE19 by in vitro transposition. *J Bacteriol*. 182(17):4744-51.
23. CHISTOSERDOVA L., LAUKEL M., PORTAIS J.C., VORHOLT J.A., LIDSTROM M.E. 2004. Multiple formate dehydrogenase enzymes in the facultative methylotroph *Methylobacterium extorquens* AM1 are dispensable for growth on methanol. *J Bacteriol*. 186(1):222-8.
24. COLEMAN N.V., BUI N.B. & HOLMES A.J. 2006. Soluble di-iron monooxygenase gene diversity in soils, sediments and ethane enrichments. *Environ Microbiol*. 8(7):1228-39.
25. DA SILVA M.L. & ALVAREZ P.J. 2004. Enhanced anaerobic biodegradation of benzene-toluene-ethylbenzene-xylene-ethanol mixtures in bioaugmented aquifer columns. *Appl Environ Microbiol*. 70(8):4720-26.
26. DARWIN A.J. & STEWART V. 1995. Expression of the *narX*, *narL*, *narP*, and *narQ* genes of *Escherichia coli* K-12: regulation of the regulators. *J Bacteriol*. 177(13):3865-69.
27. DELCHER A.L., BRATKE K.A., POWERS E.C., SALZBERG S.L. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics*. 23(6):673-79.
28. DEPREZ E., GILL E., HELMS V., WADE R.C., HUI BON HOA G. 2002. Specific and non-specific effects of potassium cations on substrate-protein interactions in cytochromes P450cam and P450lin. *J Inorg Biochem*. 91(4):597-606.
29. DESAI J.D. & BANAT I.M. 1997. Microbial production of surfactants and their

- commercial potential. *Microbiol Mol Biol Rev.* 61(1):47-64.
30. DÍAZ E. 2004. Bacterial degradation of aromatic pollutants: a paradigm of metabolic versatility. *Int Microbiol.* 7(3):173-80.
 31. DINAMARCA M.A., ARANDA-OLMEDO I., PUYET A., ROJO F. 2003. Expression of the *Pseudomonas putida* OCT plasmid alkane degradation pathway is modulated by two different global control signals: evidence from continuous cultures. *J Bacteriol.* 185(16):4772-78.
 32. DOUGHTY D.M., SAYAVEDRA-SOTO L.A., ARP D.J., BOTTOMLEY P.J. 2006. Product repression of alkane monooxygenase expression in *Pseudomonas butanovora*. *J Bacteriol.* 188(7):2586-92.
 33. DOUKYU N. 2009. Characteristics and biotechnological applications of microbial cholesterol oxidases. *Appl Microbiol Biotechnol.* 83(5):825-37.
 34. DUBOURDIEU M. & DEMOSS J.A. 1992. The *narJ* gene product is required for biogenesis of respiratory nitrate reductase in *Escherichia coli*. *J Bacteriol.* 174(3):867-72.
 35. EGLI T. 2001. Biodegradation of metal-complexing aminopolycarboxylic acids. *J Biosci Bioeng.* 92(2):89-97.
 36. EVANS F.F., ROSADO A.S., CASELLA R., SEBASTIAN G.V., MACHADO P.L.O.A., HOLMSTRÖM C., KJELLEBERG S., VAN ELSAS J.D., SELDIN L. 2004. Impact of oil contamination and biostimulation on the diversity of indigenous bacterial communities in soil microcosms. *FEMS Microbiol. Ecol.* 49:295-305.
 37. EWING B., HILLIER L., WENDL M.C., GREEN P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* 8(3):175-85.
 38. FENG L., WANG W., CHENG J., REN Y., ZHAO G., GAO C., TANG Y., LIU X., HAN W., PENG X., LIU R., WANG L. 2007. Genome and proteome of long-chain alkane degrading *Geobacillus thermodenitrificans* NG80-2 isolated from a deep-subsurface oil reservoir. *Proc Natl Acad Sci U S A.* 104(13):5602-7.
 39. FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F., KERLAVAGE A.R., BULT C.J., TOMB J.F., DOUGHERTY B.A., MERRICK J.M. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science.* 269(5223):496-512.
 40. FUHRMANN S., FERNER M., JEFFKE T., HENNE A., GOTTSCHALK G., MEYER

- O. 2003. Complete nucleotide sequence of the circular megaplasmid pHCG3 of *Oligotropha carboxidovorans*: function in the chemolithoautotrophic utilization of CO, H₂ and CO₂. *Gene*. 322:67-75.
41. FUNHOFF E.G., BAUER U., GARCÍA-RUBIO I., WITHOLT B., VAN BEILEN J.B. 2006. CYP153A6, a soluble P450 oxygenase catalyzing terminal-alkane hydroxylation. *J Bacteriol*. 188(14):5220-27.
 42. GARDNER P.R. 2008. Assay and characterization of the NO dioxygenase activity of flavohemoglobins. *Methods Enzymol*. 436:217-37.
 43. GIBAS C. & JAMBECK P. 2001. Developing Bioinformatics Computer Skills (3rd Ed). O'Reilly Media.
 44. GINZINGER D.G. 2002. Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. *Exp Hematol*. 30(6):503-12.
 45. GORKE B. & STÜLKE J. 2008. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat Rev Microbiol*. 6(8):613-24.
 46. GUNSCH C.K., KINNEY K.A., SZANISZLO P.J., WHITMAN C.P. 2006. Quantification of homogentisate-1,2-dioxygenase expression in a fungus degrading ethylbenzene. *J Microbiol Methods*. 67(2):257-65.
 47. HAMAMURA N., YEAGER C.M. & ARP D.J. 2001. Two distinct monooxygenases for alkane oxidation in *Nocardioides* sp. strain CF8. *Appl Environ Microbiol*. 67(11):4992-98.
 48. HAN L. & REYNOLDS KA. 1997. A novel alternate anaplerotic pathway to the glyoxylate cycle in streptomycetes. *J Bacteriol*. 179(16):5157-64.
 49. HÉCTOR-OSORIO H., MARTÍNEZ V., NIETO P.A., HOLMES D.S., QUATRINI R. 2008. Microbial iron management mechanisms in extremely acidic environments: comparative genomics evidence for diversity and versatility. *BMC Microbiol*. 8:203.
 50. HENDRICKSON E.L., KAUL R., ZHOU Y., BOVEE D., CHAPMAN P., CHUNG J., CONWAY DE MACARIO E., DODSWORTH J.A., GILLET W., GRAHAM D.E., HACKETT M., HAYDOCK A.K., KANG A., LAND M.L., LEVY R., LIE T.J., MAJOR T.A., MOORE B.C., PORAT I., PALMEIRI A., ROUSE G., SAENPHIMMACHAK C., SÖLL D., VAN DIEN S., WANG T., WHITMAN

- W.B., XIA Q., ZHANG Y., LARIMER F.W., OLSON M.V., LEIGH J.A. 2004. Complete genome sequence of the genetically tractable hydrogenotrophic methanogen *Methanococcus maripaludis*. *J Bacteriol.* 186(20):6956-69.
51. HERTER S., FARFSING J., GAD'ON N., RIEDER C., EISENREICH W., BACHER A., FUCHS G. 2001. Autotrophic CO₂ fixation by *Chloroflexus aurantiacus*: study of glyoxylate formation and assimilation via the 3-hydroxypropionate cycle. *J Bacteriol.* 183(14):4305-16.
 52. HUANG X. & MADAN A. 1999. CAP3: A DNA sequence assembly program. *Genome Res.* 9(9):868-77. *J Bacteriol.* 182(4):874-81.
 53. HÜGLER M., KRIEGER R.S., JAHN M., FUCHS G. (2003) Characterization of acetyl-CoA/propionyl-CoA carboxylase in *Metallosphaera sedula*. Carboxylating enzyme in the 3-hydroxypropionate cycle for autotrophic carbon fixation. *Eur J Biochem.* 270(4):736-44.
 54. ISHIKAWA J., YAMASHITA A., MIKAMI Y., HOSHINO Y., KURITA H., HOTTA K., SHIBA T., HATTORI M. 2004. The complete genomic sequence of *Nocardia farcinica* IFM 10152. *Proc Natl Acad Sci U S A.* 101(41):14925-30.
 55. IWAMOTO T. & NASU M. 2001. Current bioremediation practice and perspective. *J Biosci Bioeng.* 92(1):1-8.
 56. KAMERBEEK N.M., MOONEN M.J., VAN DER VEN J.G., VAN BERKEL W.J., FRAAIJE M.W., JANSSEN D.B. 2001. 4-Hydroxyacetophenone monooxygenase from *Pseudomonas fluorescens* ACB. A novel flavoprotein catalyzing Baeyer-Villiger oxidation of aromatic compounds. *Eur J Biochem.* 268(9):2547-57.
 57. KAMINSKA K.H., PURTA E., HANSEN L.H., BUJNICKI J.M., VESTER B., LONG K.S. 2010. Insights into the structure, function and evolution of the radical-SAM 23S rRNA methyltransferase Cfr that confers antibiotic resistance in bacteria. *Nucleic Acids Res.* 38(5):1652-63.
 58. KANEHISA.M. & BORK P. 2003. Bioinformatics in the post-sequence era. *Nat Genet.* 33 Suppl:305-10.
 59. KIM K., LEE S., LEE K., LIM D. 1998. Isolation and characterization of toluene-sensitive mutants from the toluene-resistant bacterium *Pseudomonas putida* GM73. *J Bacteriol.* 180(14):3692-96.
 60. KIM K.J., KIM S., LEE S., KANG B.S., LEE H.S., OH T.K., KIM M.H. 2006.

Crystallization and initial crystallographic characterization of the *Corynebacterium glutamicum* nitrilotriacetate monooxygenase component A. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 62(Pt 11):1141-43.

61. KING G.M. & WEBER C.F. 2007. Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nat Rev Microbiol.* 5(2):107-18.
62. KLEINSTEUBER S., RIIS V., FETZER I., HARMS H., MULLER S. 2006. Population dynamics within a microbial consortium during growth on diesel fuel in saline environments. *Appl Environ Microbiol.* 72(5):3531-42.
63. KNOL J., BODEWITS K., HESSELS G.I., DIJKHUIZEN L., van der GEIZE R. 2008. 3-Keto-5alpha-steroid Delta(1)-dehydrogenase from *Rhodococcus erythropolis* SQ1 and its orthologue in *Mycobacterium tuberculosis* H37Rv are highly specific enzymes that function in cholesterol catabolism. *Biochem J.* 410(2):339-46.
64. KOBAYASHI M., KOMEDA H., NAGASAWA T., NISHIYAMA M., HORINOUCHE S., BEPPU T., YAMADA H., SHIMIZU S. 1993. Amidase coupled with low-molecular-mass nitrile hydratase from *Rhodococcus rhodochrous* J1. Sequencing and expression of the gene and purification and characterization of the gene product. *Eur J Biochem.* 217(1):327-36.
65. KOROTKOVA N. & LIDSTROM M.E. 2004. MeaB is a component of the methylmalonyl-CoA mutase complex required for protection of the enzyme from inactivation. *J Biol Chem.* 279(14):13652-58.
66. KÜHNER S., WÖHLBRAND L., FRITZ I., WRUCK W., HULTSCHIG C., HUFNAGEL P., KUBE M., REINHARDT R., RABUS R. 2005. Substrate-dependent regulation of anaerobic degradation pathways for toluene and ethylbenzene in a denitrifying bacterium, strain EbN1. *J Bacteriol.* 187(4):1493-503.
67. LABINGER J.A. & BERCAW J.E. 2002. Understanding and exploiting C-H bond activation. *Nature.* 417(6888):507-14.
68. LARKIN M.J., KULAKOV L.A. & ALLEN C.C. 2005. Biodegradation and *Rhodococcus*--masters of catabolic versatility. *Curr Opin Biotechnol.* 16(3):282-90.
69. LEE D.S., YAMADA A., SUGIMOTO H., MATSUNAGA I., OGURA H., ICHIHARA K., ADACHI S., PARK S.Y., SHIRO Y. 2003. Substrate recognition and

molecular mechanism of fatty acid hydroxylation by cytochrome P450 from *Bacillus subtilis*. Crystallographic, spectroscopic, and mutational studies. *J Biol Chem.* 278(11):9761-67.

70. LEHNINGER A. 2007. Princípio de Bioquímica. (4nd Edição) Ed. Savier.
71. LEPESHEVA G.I. & WATERMAN M.R. 2007. Sterol 14alpha-demethylase cytochrome P450 (CYP51), a P450 in all biological kingdoms. *Biochim Biophys Acta.* 1770(3):467-77.
72. LEYTHAEUSER D. & RÜCKHEIM J. 1989. Heterogeneity of oil composition within a reservoir as a reflection of accumulation history. *Geochimica et Cosmochimica Acta.* 53(8):2119-23.
73. LI Q., KHOSLA C., PUGLISI J.D. & LIU C.W. 2003. Solution structure and backbone dynamics of the holo form of the frenolicin acyl carrier protein. *Biochemistry.* 42(16):4648-57.
74. LI R., LI Y., KRISTIANSEN K., WANG J. 2008. SOAP: short oligonucleotide alignment program. *Bioinformatics.* 24(5):713-14.
75. LORENZ M.C. & FINK G.R. 2002. Life and death in a macrophage: role of the glyoxylate cycle in virulence. *Eukaryot Cell.* 1(5):657-62.
76. LOVLEY D.R. & LLOYD J.R. 2000. Microbes with a mettle for bioremediation. *Nat Biotechnol.* 18(6):600-01.
77. LUZ A.P., PELLIZARI V.H., WHYTE L.G., GREER C.W. 2004. A survey of indigenous microbial hydrocarbon degradation genes in soils from Antarctica and Brazil. *Can J Microbiol.* 50(5):323-33.
78. LYKIDIS A., MAVROMATIS K., IVANOVA N., ANDERSON I., LAND M., DIBARTOLO G., MARTINEZ M., LAPIDUS A., LUCAS S., COPELAND A., RICHARDSON P., WILSON D.B., KYRPIDES N. 2007. Genome sequence and analysis of the soil cellulolytic actinomycete *Thermobifida fusca* YX. *J Bacteriol.* 189(6):2477-86.
79. MAIER T., FÖRSTER H.H., ASPERGER O., HAHN U. 2001. Molecular characterization of the 56-kDa CYP153 from *Acinetobacter* sp. EB104. *Biochem Biophys Res Commun.* 286(3):652-58.
80. MARIN M.M., YUSTE L. & ROJO F. 2003. Differential expression of the components of the two alkane hydroxylases from *Pseudomonas aeruginosa*. *J Bacteriol.*

185(10):3232-37.

81. MARSHALL A.G. & RODGERS R.P. 2004. Petroleomics: the next grand challenge for chemical analysis. *Acc Chem Res.* 37(1):53-9.
82. MARTINEZ-FLEITES C., PROCTOR M., ROBERTS S., BOLAM D.N., GILBERT H.J., DAVIES G.J. 2006. Insights into the synthesis of lipopolysaccharide and antibiotics through the structures of two retaining glycosyltransferases from family GT4. *Chem Biol.* 13(11):1143-52.
83. MARTÍNKOVÁ L., UHNÁKOVÁ B., PÁTEK M., NESVERA J., KREN V. 2009. Biodegradation potential of the genus *Rhodococcus*. *Environ Int.* 35(1):162-77.
84. MATHIES R.A., PECK K. & STRYER L. 1990. Optimization of high-sensitivity fluorescence detection. *Anal Chem.* 62(17):1786-91.
85. MCGINNIS S. & MADDEN T.L. 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res.* 32:20-25.
86. MCLEOD M.P., WARREN R.L., HSIAO W.W., ARAKI N., MYHRE M., FERNANDES C., MIYAZAWA D., WONG W., LILLQUIST A.L., WANG D., DOSANJH M., HARA H., PETRESCU A., MORIN R.D., YANG G., STOTT J.M., SCHEIN J.E., SHIN H., SMAILUS D., SIDDIQUI A.S., MARRA M.A., JONES S.J., HOLT R., BRINKMAN F.S., MIYAUCHI K., FUKUDA M., DAVIES J.E., MOHN W.W., ELTIS L.D. 2006. The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. *Proc Natl Acad Sci U S A.* 103(42):15582-87.
87. MEDINI D., SERRUTO D., PARKHILL J., RELMAN D.A., DONATI C., MOXON R., FALKOW S., RAPPUOLI R. 2008. Microbiology in the post-genomic era. *Nat Rev Microbiol.* 6(6):419-30.
88. MONGODIN E.F., SHAPIR N., DAUGHERTY S.C., DEBOY R.T., EMERSON J.B., SHVARTZBEYN A., RADUNE D., VAMATHEVAN J., RIGGS F., GRINBERG V., KHOURI H., WACKETT L.P., NELSON K.E., SADOWSKY M.J. 2006. Secrets of soil survival revealed by the genome sequence of *Arthrobacter aurescens* TC1. *PLoS Genet.* 2(12):e214.
89. MORENO C.M., GONZÁLEZ B. & BLANCO-SANTOS M.J. 2004. Tratamientos biológicos de suelos contaminados: contaminación por hidrocarburos. Aplicaciones de hongos en tratamientos de biorrecuperación. Ver. Iberoam Micol.

21: 103-120.

90. MORENO-VIVIÁN C., CABELLO P., MARTÍNEZ-LUQUE M., BLASCO R., CASTILLO F. 1999. Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases. *J Bacteriol.* 181(21):6573-84.
91. NAVAS J., GONZÁLEZ-ZORN B., LADRÓN N., GARRIDO P., VÁZQUEZ-BOLAND J.A. 2001. Identification and mutagenesis by allelic exchange of choE, encoding a cholesterol oxidase from the intracellular pathogen *Rhodococcus equi*. *J Bacteriol.* 183(16):4796-805.
92. NOORDMAN W.H. & JANSSEN D.B. 2002. Rhamnolipid stimulates uptake of hydrophobic compounds by *Pseudomonas aeruginosa*. *Appl Environ Microbiol.* 68(9):4502-08.
93. OCHMAN H., LAWRENCE J.G. & GROISMAN E.A. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature.* 405(6784):299-304.
94. OKOH A.I., AJISEBUTU S., BABALOLA G.O., TREJO-HERNANDEZ M.R. 2002. Biodegradation of Mexican heavy crude oil (*Maya*) by *Pseudomonas aeruginosa*. *J. Trop. Biosci.* 2(1): 12-24.
95. OUTTEN F.W., WOOD M.J., MUNOZ F.M., STORZ G. 2003. The SufE protein and the SufBCD complex enhance SufS cysteine desulfurase activity as part of a sulfur transfer pathway for Fe-S cluster assembly in *Escherichia coli*. *J Biol Chem.* 278(46):45713-9.
96. PACE H.C. & BRENNER C. 2001. The nitrilase superfamily: classification, structure and function. *Genome Biol.*;2(1).
97. PACE N.R. 1997. A molecular view of microbial diversity and the biosphere. *Science.* 276(5313):734-40.
98. PACE N.R., STAHL D.A., LANE D., OLSEN G.J. 1986. The analysis of natural population by ribosomal rna sequence. *Adv. Gen. Microbiol. Eco.* 9:1-55.
99. PEARSON W.R. 1998. Empirical statistical estimates for sequence similarity searches. *J Mol Biol.* 276(1):71-84.
100. PELZMANN A., FERNER M., GNIDA M., MEYER-KLAUCKE W., MAISEL T., MEYER O. 2009. The CoxD protein of *Oligotropha carboxidovorans* is a predicted AAA+ ATPase chaperone involved in the biogenesis of the CO

dehydrogenase [CuSMoO₂] cluster. *J Biol Chem.* 284(14):9578-86.

101. PERNA N.T., PLUNKETT G., BURLAND V., MAU B., GLASNER J.D., ROSE D.J., MAYHEW G.F., EVANS P.S., GREGOR J., KIRKPATRICK H.A., PÓSFAI G., HACKETT J., KLINK S., BOUTIN A., SHAO Y., MILLER L., GROTBECK E.J., DAVIS N.W., LIM A., DIMALANTA E.T., POTAMOUSIS K.D., APODACA J., ANANTHARAMAN T.S., LIN J., YEN G., SCHWARTZ D.C., WELCH R.A., BLATTNER F.R. 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature.* 409(6819):529-33.
102. PETRUSMA M., DIJKHUIZEN L. & van der GEIZE R. 2009. *Rhodococcus rhodochrous* DSM 43269 3-ketosteroid 9 α -hydroxylase, a two-component iron-sulfur-containing monooxygenase with subtle steroid substrate specificity. *Appl Environ Microbiol.* 75(16):5300-7.
103. POOLE L.B. 2005. Bacterial defenses against oxidants: mechanistic features of cysteine-based peroxidases and their flavoprotein reductases. *Arch Biochem Biophys.* 433(1):240-54.
104. POWELL S.M., FERGUSON S.H., BOWMAN J.P., SNAPE I. 2006. Using real-time PCR to assess changes in the hydrocarbon-degrading microbial community in Antarctic soil during bioremediation. *Microb Ecol.* 52(3):523-32.
105. RAINEY F.A., KLATTE S., KROPPENSTEDT R.M., STACKEBRANDT E. 1995. *Dietzia*, a new genus including *Dietzia maris* comb. nov., formerly *Rhodococcus maris*. *Int J Syst Bacteriol.* 45(1):32-36.
106. REICHENBECHER W. & MURRELL J.C. 1999. Linear alkanesulfonates as carbon and energy sources for gram-positive and gram-negative bacteria. *Arch Microbiol.* 171(6):430-38.
107. REY M.W., RAMAIYA P., NELSON B.A., BRODY-KARPIN S.D., ZARETSKY E.J., TANG M., LOPEZ DE LEON A., XIANG H., GUSTI V., CLAUSEN I.G., OLSEN P.B., RASMUSSEN M.D., ANDERSEN J.T., JØRGENSEN P.L., LARSEN T.S., SOROKIN A., BOLOTIN A., LAPIDUS A., GALLERON N., EHRLICH S.D., BERKA R.M. 2004. Complete genome sequence of the industrial bacterium *Bacillus licheniformis* and comparisons with closely related *Bacillus* species. *Genome Biol.* 5(10):R77.
108. RIEGER P.G., MEIER H.M., GERLE M., VOGT U., GROTH T., KNACKMUSS H.J.

2002. Xenobiotics in the environment: present and future strategies to obviate the problem of biological persistence. *J Biotechnol.* 94(1):101-23.
109. ROGERS S.L. & MCCLURE N. 2003. in Bioremediation: a critical review (eds. Head, I. M. Singleton, I & Miller, M. G.) 27-59 (Horin Scientific Press. Wymondham, UK).
110. ROJO F. 2005. Specificity at the end of the tunnel: understanding substrate length discrimination by the AlkB alkane hydroxylase. *J Bacteriol.* 187(1):19-22.
111. ROJO F. 2009. Degradation of alkanes by bacteria. *Environ Microbiol.* 11(10):2477-90.
112. ROTHSCILD L.J. 2008. The evolution of photosynthesis...again? *Philos Trans R Soc Lond B Biol Sci.* 363(1504):2787-801.
113. RUIZ-MANZANO A., YUSTE L. & ROJO F. 2005. Levels and activity of the *Pseudomonas putida* global regulatory protein Crc vary according to growth conditions. *J Bacteriol.* 187(11):3678-86.
114. SALMINEN J.M., TUOMI P.M. & JØRGENSEN K.S. 2008. Functional gene abundances (*nahAc*, *alkB*, *xylE*) in the assessment of the efficacy of bioremediation. *Appl Biochem Biotechnol.* 151(2-3):638-52.
115. SAMANTA S.K., SINGH O.V. & JAIN R.K. 2002. Polycyclic aromatic hydrocarbons: environmental pollution and bioremediation. *Trends Biotechnol.* 20(6):243-48.
116. SAMBROOK J. & RUSELL D.W. 2001. Molecular cloning: a laboratory manual (3rd Ed.). Cold Spring Harbor, NT: Cold Spring Harbor Laboratory.
117. SAMPSON N.S. & VRIELINK A. 2003. Cholesterol oxidases: a study of nature's approach to protein design. *Acc Chem Res.* 36(9):713-22.
118. SANTIAGO B., SCHÜBEL U., EGELSEER C., MEYER O. 1999. Sequence analysis, characterization and CO-specific transcription of the *cox* gene cluster on the megaplasmid pHCG3 of *Oligotropha carboxidovorans*. *Gene.* 236(1):115-24.
119. SCHNEIKER S., MARTINS DOS SANTOS V.A., BARTELS D., BEKEL T., BRECHT M., BUHRMESTER J., CHERNIKOVA T.N., DENARO R., FERRER M., GERTLER C., GOESMANN A., GOLYSHINA O.V., KAMINSKI F., KHACHANE A.N., LANG S., LINKE B., MCHARDY A.C., MEYER F., NECHITAYLO T., PÜHLER A., REGENHARDT D., RUPP O., SABIROVA J.S., SELBITSCHKA W., YAKIMOV M.M., TIMMIS K.N., VORHÖLTER F.J., WEIDNER S., KAISER O., GOLYSHIN P.N. 2006. Genome sequence of the

ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis*. *Nat Biotechnol.* 24(8):997-1004.

120. SCHULZ S., PERÉZ-DE-MORA A., ENGEL M., MUNCH J.C., SCHLOTTER M. 2010. A comparative study of most probable number (MPN)-PCR vs. real-time-PCR for the measurement of abundance and assessment of diversity of *alkB* homologous genes in soil. *J Microbiol Methods.* 80(3):295-8.
121. SCULLION J. 2006. Remediating polluted soils. *Naturwissenschaften.* 93(2):51-65.
122. SEKINE M., TANIKAWA S., OMATA S., SAITO M., FUJISAWA T., TSUKATANI N., TAJIMA T., SEKIGAWA T., KOSUGI H., MATSUO Y., NISHIKO R., IMAMURA K., ITO M., NARITA H., TAGO S., FUJITA N., HARAYAMA S. 2008. Sequence analysis of three plasmids harboured in *Rhodococcus erythropolis* strain PR4. *Environ Microbiol.* 8(2):334-46.
123. SELDIN L. & DUBNAU L. 1985. DNA homology among *Bacillus polymyxa*, *Bacillus macerans*, *Bacillus azotofixans* and other nitrogen fixing *Bacillus* strains. *Int. J. Syst. Bacteriol.* 35, 151-54.
124. SHANKLIN J. & WHITTLE E. 2003. Evidence linking the *Pseudomonas oleovorans* alkane omega-hydroxylase, an integral membrane diiron enzyme, and the fatty acid desaturase family. *FEBS Lett*; 545(2-3):188-92.
125. SHENDURE J. & JI H. 2008. Next-generation DNA sequencing. *Nat Biotechnol.* 26(10):1135-45.
126. SHENDURE J., PORRECA G.J., REPPAS N.B., LIN X., MCCUTCHEON J.P., ROSENBAUM A.M., WANG M.D., ZHANG K., MITRA R.D., CHURCH G.M. 2005. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science.* 309(5741):1728-32.
127. SIDDIQUI S. & ADAMS W.A. 2002. The fate of diesel hydrocarbons in soils and their effect on the germination of perennial ryegrass. *Environ Toxicol.* 17(1):49-62.
128. SMITH L.M., SANDERS J.Z., KAISER R.J., HUGHES P., DODD C., CONNELL C.R., HEINER C., KENT S.B., HOOD L.E. 1986. Fluorescence detection in automated DNA sequence analysis. *Nature.* 321(6071):674-9.
129. SMITS T.H., BALADA S.B., WITHOLT B., VAN BEILEN J.B. 2002. Functional analysis of alkane hydroxylases from gram-negative and gram-positive bacteria. *J Bacteriol.* 184(6):1733-42.

130. STARKENBURG S.R., CHAIN P.S., SAYAVEDRA-SOTO L.A., HAUSER L., LAND M.L., LARIMER F.W., MALFATTI S.A., KLOTZ M.G., BOTTOMLEY P.J., ARP D.J., HICKEY W.J. 2006. Genome sequence of the chemolithoautotrophic nitrite-oxidizing bacterium *Nitrobacter winogradskyi* Nb-255. *Appl Environ Microbiol.* 72(3):2050-63.
131. STHAPIT B., OH T.J., LAMICHHANE R., LIOU K., LEE H.C., KIM C.G., SOHNG J.K. 2004. Neocarzinostatin naphthoate synthase: an unique iterative type I PKS from neocarzinostatin producer *Streptomyces carzinostaticus*. *FEBS Lett.* 566(1-3):201-6.
132. STINEAR T.P., SEEMANN T., HARRISON P.F., JENKIN G.A., DAVIES J.K., JOHNSON P.D., ABDELLAH Z., ARROWSMITH C., CHILLINGWORTH T., CHURCHER C., CLARKE K., CRONIN A., DAVIS P., GOODHEAD I., HOLROYD N., JAGELS K., LORD A., MOULE S., MUNGALL K., NORBERTCZAK H., QUAIL M.A., RABBINOWITSCH E., WALKER D., WHITE B., WHITEHEAD S., SMALL P.L., BROSCHE R., RAMAKRISHNAN L., FISCHBACH M.A., PARKHILL J., COLE S.T. 2008. Insights from the complete genome sequence of *Mycobacterium marinum* on the evolution of *Mycobacterium tuberculosis*. *Genome Res.* 18(5):729-41.
133. STRIJEWSKI A. 1982. The steroid-9 alpha-hydroxylation system from *Nocardia* species. *Eur J Biochem.* 128(1):125-35.
134. TAKARADA H., SEKINE M., KOSUGI H., MATSUO Y., FUJISAWA T., OMATA S., KISHI E., SHIMIZU A., TSUKATANI N., TANIKAWA S., FUJITA N., HARAYAMA S. 2008. Complete genome sequence of the soil actinomycete *Kocuria rhizophila*. *J Bacteriol.* 190(12):4139-46.
135. TETTELIN H., MASIGNANI V., CIESLEWICZ M.J., DONATI C., MEDINI D., WARD N.L., ANGIUOLI S.V., CRABTREE J., JONES A.L., DURKIN A.S., DEBOY R.T., DAVIDSEN T.M., MORA M., SCARSELLI M., MARGARIT Y ROS I., PETERSON J.D., HAUSER C.R., SUNDARAM J.P., NELSON W.C., MADUPU R., BRINKAC L.M., DODSON R.J., ROISOVITZ M.J., SULLIVAN S.A., DAUGHERTY S.C., HAFT D.H., SELENGUT J., GWINN M.L., ZHOU L., ZAFAR N., KHOURI H., RADUNE D., DIMITROV G., WATKINS K.,

- O'CONNOR K.J., SMITH S., UTTERBACK T.R., WHITE O., RUBENS C.E., GRANDI G., MADOFF L.C., KASPER D.L., TELFORD J.L., WESSELS M.R., RAPPUOLI .R, FRASER C.M. 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". *Proc Natl Acad Sci U S A*. 102(39):13950-55.
136. THOMAS J.M. & RAJA R. (2005) Design of a "green" one-step catalytic production of epsilon-caprolactam (precursor of nylon-6). *Proc Natl Acad Sci U S A*. 102(39):13732-6.
137. THUY T.T., LEE H.C., KIM C.G., HEIDE L., SOHNG J.K. 2005. Functional characterizations of novWUS involved in novobiocin biosynthesis from *Streptomyces spheroides*. *Arch Biochem Biophys*. 436(1):161-67.
138. van BEILEN J.B. & WITHOLT B. 2004. Alkane degradation by pseudomonas, p. 397-423. In: J. L. Ramos (ed.), *The Pseudomonads cells*. vol. 3. Kluwer Academic Publishers, Dordrecht, The Netherlands.
139. van BEILEN J.B., DUETZ W.A., SCHMID A., WITHOLT B. 2003a. Practical issues in the application of oxygenases. *Trends Biotechnol*. 21(4):170-77.
140. van BEILEN J.B., EGGINK G., ENEQUIST H., BOS R, WITHOLT B. 1992 . DNA sequence determination and functional characterization of the OCT-plasmid-encoded *alkJKL* genes of *Pseudomonas oleovorans*. *Mol Microbiol*. 6(21):3121-36.
141. van BEILEN J.B., LI Z., DUETZ W.A., SMITS T.H.M., WITHOLT B. 2003b. Diversity of alkane hydroxylase systems in the environment. *Oil Gas Sci. Technol*. 58:427-440.
142. van BEILEN J.B., MARIN M.M., SMITS T.H., ROTH LISBERGER M., FRANCHINI A.G., WITHOLT B., ROJO F. 2004. Characterization of two alkane hydroxylase genes from the marine hydrocarbonoclastic bacterium *Alcanivorax borkumensis*. *Environ Microbiol*. 6(3):264-73.
143. van BEILEN J.B., NEUENSCHWANDER M., SMITS T.H., ROTH C., BALADA S.B., WITHOLT B. 2002a. Rubredoxins involved in alkane oxidation. *J Bacteriol*. 184(6):1722-32.
144. van BEILEN J.B., SMITS T.H., ROOS F.F., BRUNNER T., BALADA S.B., ROTH LISBERGER M., WITHOLT B. 2005. Identification of an amino acid

position that determines the substrate range of integral membrane alkane hydroxylases. *J Bacteriol.* 187(1):85-91.

145. van BEILEN J.B., SMITS T.H., WHYTE L.G., SCHORCHT S., ROTH LISBERGER M., PLAGGEMEIER T., ENGESSER K.H., WITHOLT B. 2002b. Alkane hydroxylase homologues in Gram-positive strains. *Environ Microbiol.* 4(11):676-82.
146. van BEILEN J.B., WUBBOLTS M.G., WITHOLT B. 1994. Genetics of alkane oxidation by *Pseudomonas oleovorans*. *Biodegradation.* 5(3-4):161-74.
147. van der GEIZE R. & DIJKHUIZEN L. 2004. Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. *Curr Opin Microbiol.* 7(3):255-61.
148. van der GEIZE R., HESSELS G.I., NIENHUIS-KUIPER M., DIJKHUIZEN L. 2008. Characterization of a second *Rhodococcus erythropolis* SQ1 3-ketosteroid 9 α -hydroxylase activity comprising a terminal oxygenase homologue, KshA2, active with oxygenase-reductase component KshB. *Appl Environ Microbiol.* 74(23):7197-203.
149. van der GEIZE R., YAM K., HEUSER T., WILBRINK M.H., HARA H., ANDERTON M.C., SIM E., DIJKHUIZEN L., DAVIES J.E., MOHN W.W., ELTIS L.D. 2007. A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into *Mycobacterium tuberculosis* survival in macrophages. *Proc Natl Acad Sci U S A.* 104(6):1947-52.
150. van der PLOEG J.R., EICHHORN E. & LEISINGER T. 2001. Sulfonate-sulfur metabolism and its regulation in *Escherichia coli*. *Arch Microbiol.* 176(1-2):1-8.
151. van HAMME J.D., SINGH A. & WARD OP. 2003. Recent advances in petroleum microbiology. *Microbiol Mol Biol Rev.* 67(4):503-49.
152. VARGAS C., ARGANDOÑA M., REINA-BUENO M., RODRÍGUEZ-MOYA J., FERNÁNDEZ-AUNIÓN C., NIETO J.J. 2008. Unravelling the adaptation responses to osmotic and temperature stress in *Chromohalobacter salexigens*, a bacterium with broad salinity tolerance. *Saline Systems.* 4:14.
153. VAYSSE P.J., PRAT L., MANGENOT S., CRUVEILLER S., GOULAS P., GRIMAUD R. 2009. Proteomic analysis of *Marinobacter hydrocarbonoclasticus* SP17 biofilm formation at the alkane-water interface reveals novel proteins and cellular

- processes involved in hexadecane assimilation. *Res Microbiol.* 160(10):829-37.
154. VELICER G.J., RADDATZ G., KELLER H., DEISS S., LANZ C., DINKELACKER I., SCHUSTER S.C. 2006. Comprehensive mutation identification in an evolved bacterial cooperator and its cheating ancestor. *Proc Natl Acad Sci U S A.* 23;103(21):8107-12.
155. VOLKERING F., BREURE A.M. & RULKENS W.H. 1998. Microbiological aspects of surfactant use for biological soil remediation. *Biodegradation.* 8(6):401-17.
156. von der WEID I., MARQUES J.M., CUNHA C.D., LIPPI R.K., DOS SANTOS S.C., ROSADO A.S., LINS U., SELDIN L. 2007. Identification and biodegradation potential of a novel strain of *Dietzia cinnamea* isolated from a petroleum-contaminated tropical soil. *Syst Appl Microbiol. Syst Appl Microbiol.* 30(4):331-9.
157. WANKE M., SKORUPINSKA-TUDEK K. & SWIEZEWSKA E. 2001. Isoprenoid biosynthesis via 1-deoxy-D-xylulose 5-phosphate/2-C-methyl-D-erythritol 4-phosphate (DOXP/MEP) pathway. *Acta Biochim Pol.* 48(3):663-72.
158. WASMUND K., BURNS K.A., KURTBÖKE D.I., BOURNE D.G. 2009. Novel alkane hydroxylase gene (alkB) diversity in sediments associated with hydrocarbon seeps in the Timor Sea, Australia. *Appl Environ Microbiol.* 75(23):7391-8.
159. WATANABE M.E. 2001. Can bioremediation bounce back? *Nat Biotechnol.* 19(12):1111-15.
160. WEITNAUER G., MÜHLENWEG A., TREFZER A., HOFFMEISTER D., SÜSSMUTH R.D., JUNG G., WELZEL K., VENTE A., GIRRESER U., BECHTHOLD A. 2001. Biosynthesis of the orthosomycin antibiotic avilamycin A: deductions from the molecular analysis of the avi biosynthetic gene cluster of *Streptomyces viridochromogenes* Tü57 and production of new antibiotics. *Chem Biol.* 8(6):569-81.
161. WELCH R.A., BURLAND V., PLUNKETT G., REDFORD P., ROESCH P., RASKO D., BUCKLES E.L., LIOU S.R., BOUTIN A., HACKETT J., STROUD D., MAYHEW G.F., ROSE D.J., ZHOU S., SCHWARTZ D.C., PERNA N.T., MOBLEY H.L., DONNENBERG M.S., BLATTNER F.R. 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc Natl Acad Sci U S A.* 99(26):17020-24.
162. WENTZEL A., ELLINGSEN T.E., KOTLAR H.K., ZOTCHEV S.B., THRONE-

- HOLST M. 2007. Bacterial metabolism of long-chain *n*-alkanes. *Appl Microbiol Biotechnol.* 76(6):1209-21.
163. WHYTE L.G., SLAGMAN S.J., PIETRANTONIO F., BOURBONNIERE L., KOVAL S.F., LAWRENCE J.R., INNIS W.E., GREER C.W. 1999. Physiological adaptations involved in alkane assimilation at a low temperature by *Rhodococcus* sp. strain Q15. *Appl Environ Microbiol.* 65(7):2961-68.
164. WHYTE L.G., SMITS T.H., LABBE D., WITHOLT B., GREER C.W., VAN BEILEN J.B. 2002. Gene cloning and characterization of multiple alkane hydroxylase systems in *Rhodococcus* strains Q15 and NRRL B-16531. *Appl Environ Microbiol.* 68(12):5933-42.
165. WILLIAMS W.A., LOBOS J.H. & CHEETHAM W.E. A phylogenetic analysis of aerobic polychlorinated biphenyl-degrading bacteria. *Int J Syst Bacteriol.* 47(1):207-10.
166. WU D., RAYMOND J., WU M., CHATTERJI S., REN Q., GRAHAM J.E., BRYANT D.A., ROBB F., COLMAN A., TALLON L.J., BADGER J.H., MADUPU R., WARD N.L., EISEN J.A. 2009. Complete genome sequence of the aerobic CO-oxidizing thermophile *Thermomicrobium roseum*. *PLoS One.* (1):e4207.
167. WU M., REN Q., DURKIN A.S., DAUGHERTY S.C., BRINKAC L.M., DODSON R.J., MADUPU R., SULLIVAN S.A., KOLONAY J.F., HAFT D.H., NELSON W.C., TALLON L.J., JONES K.M., ULRICH L.E., GONZALEZ J.M., ZHULIN I.B., ROBB F.T., EISEN J.A. 2005. Life in hot carbon monoxide: the complete genome sequence of *Carboxydotherrmus hydrogeniformans* Z-2901. *PLoS Genet.* 1(5):e65.
168. WYATT J.M. & KNOWLES C.J. 1995. The development of a novel strategy for the microbial treatment of acrylonitrile effluents. *Biodegradation* 6:93-107.
169. YAGI J.M., SIMS D., BRETTIN T., BRUCE D., MADSEN E.L. 2009. The genome of *Polaromonas naphthalenivorans* strain CJ2, isolated from coal tar-contaminated sediment, reveals physiological and metabolic versatility and evolution through extensive horizontal gene transfer. *Environ Microbiol.* 11(9):2253-70.
170. YAMADA H. & KOBAYASHI M. 1996 . Nitrile hydratase and its application to industrial production of acrylamide. *Biosci. Biotechnol. Biochem.* 60:1391-1400.

171. YASSIN A.F., HUPFER H. & SCHAAL K.P. 2006. *Dietzia cinnamea* sp. nov., a novel species isolated from a perianal swab of a patient with a bone marrow transplant. *Int J Syst Evol Microbiol.* 56(3):641-45.
172. YUMOTO I., NAKAMURA A., IWATA H., KOJIMA K., KUSUMOTO K., NODASAKA Y., MATSUYAMA H. 2002. *Dietzia psychralcaliphila* sp. nov., a novel, facultatively psychrophilic alkaliphile that grows on hydrocarbons. *Int J Syst Evol Microbiol.* 52(1):85-90.
173. ZHAO Q., HE Q., DING W., TANG M., KANG Q., YU Y., DENG W., ZHANG Q., FANG J., TANG G., LIU W. 2008. Characterization of the azinomycin B biosynthetic gene cluster revealing a different iterative type I polyketide synthase for naphthoate biosynthesis. *Chem Biol.* 15(7):693-705.

10 - Anexo

**INSIGHTS FROM THE DRAFT GENOME OF *Dietzia cinnamea* P4 REVEALS
MECHANISMS OF SURVIVAL IN COMPLEX TROPICAL SOIL HABITATS**

**Luciano Procópio, Vanessa M. Alvarez, Diogo A. Jurelevicius, Marcia .R.R.
Coelho, Lars Hansen, Søren J. Sørensen, Janine S. Cardoso, Marcelo Pádula, Alvaro C.
Leitão, Lucy Seldin, Jan Dirk van Elsas***

Goningen University, Microbial Ecology, Groningen, The Netherlands.

Department of Microbial Ecology, CEES, University of Groningen, Kerklaan 30, 9751 NN Haren, The
Netherlands

* j.d.van.elsas@rug.nl

Abstract

The draft genome of strain P4 was determined using pyrosequencing. In total, 435 supercontigs were obtained and analyzed. We here describe and interpret the main features of the draft genome. The genome contained a total of 3,557,916 bp, arranged in a single replicon with an average G+C percentage of 70.9 %. It revealed the presence of complete pathways for basically all central metabolic routes. Also present were complete sets of genes for the glyoxalate and reductive carboxylate cycles. Autotrophic growth was suggested to occur by the presence of genes for aerobic CO oxidation, formate/formaldehyde oxidation, the reverse tricarboxylic acid cycle and the 3-hydropropionate cycle for CO₂ fixation. Secondary metabolism was evidenced by the presence of genes for the biosynthesis of terpene compounds, frenolicin, nanaomycin and avilamycin A antibiotics. Furthermore, a probable role in azinomycin B synthesis, an important product with antitumor activity, was indicated. The complete Alk operon for the degradation of n-alkanes was found to be present, as were clusters of genes for biphenyl ring dihydroxylation. This study brings new insights in the genetics and physiology of *Dietzia cinnamea* P4, which is useful in biotechnology and

bioremediation.

Keywords

Dietzia cinnamea P4, genome sequence, biotechnology, bioremediation.

Introduction

Dietzia species are Gram-positive bacteria with high G+C content, which belong to the order actinomycetales. They are widely distributed in nature, as they have been isolated from soil, deepsea sediment, soda lakes, reed rhizomes, skin and the intestinal tracts of marine fish (Yassin et al., 2006). *Dietzia cinnamea* strain P4 was isolated in a study on the diversity of hydrocarbon degraders in a tropical rainforest soil in Brazil, with no history of oil contamination (Evans et al., 2004). In subsequent work, von der Weid and colleagues (von der Weid et al., 2007) demonstrated that strain P4 is able to degrade a wide range of *n*-alkanes, pristane as well as phytane, whereas it grows in mineral liquid media amended with aromatic hydrocarbons as sole carbon sources. It thus may have applications in the bioremediation of contaminated soils.

The ability to degrade a wide range of organic compounds probably derives from the lifestyle of *Dietzia cinnamea* like bacteria, which may assimilate a variety of different carbon sources and may also respond rapidly to environmental stressors. To the best of our knowledge, so far not a single *Dietzia* genome has been published. In an effort to better understand the relationship between *Dietzia*-like soil actinomycetes and the tropical soil habitat, including the response to oil contamination, we pyrosequenced and annotated the genome of *D. cinnamea* P4. Our results not only confirm the potential for bioremediation of hydrocarbon contamination, but also the ability to degrade other xenobiotic compounds. Furthermore, a complex metabolic network tuned to energy acquisition and biosynthesis of

products of commercial interest was revealed.

Results and Discussion

Ecology of *D. cinnamea* P4. The habitat from which strain P4 was isolated, tropical rainforest soil (sandy loam Cambisol; pH 5.5), revealed the presence of high levels of humic acids and heavy metals (Evans et al., 2004). The forest soil further showed oxygen characteristics that rapidly fluctuate between anaerobic/microaerobic and aerobic conditions, as a result of rapid fluctuations in moisture content. Hence, cells were likely exposed to periodic (fluctuations in) osmotic pressure, impacting their physiology, as well as directing DNA and cell wall damage. Strain P4 was non-spore forming, non-motile and strictly aerobic. It was also able to grow at pH values up to 10 but not at pH below 6.0. Furthermore, it showed growth at NaCl concentrations from zero to 10%. Assays of degradation with different h petroleum ydrocarbons revealed the ability of strain P4 to use several compounds from oil and diesel, in particular benzene and alkane compounds, as the sole carbon sources (von der Weid et al., 2007).

In order to evaluate the biological responses of strain P4 to environmental stresses conferred by physical and oxidative chemical mutagens, cultures in the exponential growth phase (next to those of *E. coli* wild type strains K12A15 and BH980) were exposed to a series of stresses, i.e. UV-A, UV-B and UV-C irradiations, as well as oxidative stress (using H₂O₂ and cumene hydroperoxide - CH). Surviving fractions in the treated populations were determined by using the ratio between the number of viable cells before and after each treatment (see Supporting information). Strain P4 revealed to be more resistant than the *E. coli* wild type strains to all physical agents tested ($p < 0.05$), indicating that it is endowed with efficient DNA repair or tolerance mechanisms, particularly when considering its resistance to

UV-A (Fig. 1). Regarding oxidative stress tolerance (using H₂O₂ and CH), strain P4 showed a resistance similar to that of control strains K12A15 and BW9091 (data not shown). We also evaluated the mutability of strain P4 compared with strains K12A15 and BW9091. Spontaneous mutagenesis to rifampin (Rp) resistance of strain P4 occurred at the same level as that of *E. coli* K12A15. The latter strain is not considered to be a mutator phenotype, especially when compared to strain BH980 (data not shown). Interestingly, strain P4 was extremely refractory to UV-A-induced mutagenesis, which was significantly different from that of wild-type strain K12A15 ($p < 0.05$). Whereas strain K12A15 revealed a 10.8-fold increase of mutation to Rif^R after UV-A treatment, strain P4 displayed no increase in the prevalence of Rif^R mutants. Under UV-B, strain P4 was quite susceptible to mutagenesis, but to a lesser extent than K12A15, both in terms of absolute number of Rif^R mutants as in fold increase ($p < 0.05$). A similar response was observed between strains P4 and K12A15 after challenge with UV-C, and these two strains did not reveal the absolute Rif^R mutant levels of the BH980 mutator strain. Overall, these data indicate that *Dietzia cinnemea* P4 is not a spontaneous mutator strain. It is particularly resistant to UV-induced mutagenesis, especially to the oxidative (UV-A) counterpart of the solar UV spectrum.

General genome features. Using high-coverage pyrosequencing on an FLX machine, we then assembled the genome (coverage 30X). The sequencing yielded 565,157 shotgun reads for a total of 3,556,793 nucleotides (Table 1). The resulting assembly finally consisted of 435 supercontigs, with a length range of 5 to 61 kb, giving a mean length of about 8,000 nucleotides. Comparison with the PFGE results showed that the 435 supercontigs covered roughly 99% of the single chromosome seen on gel (data not shown). Given the difficulty to completely close the genome (due to repeats and extremely-high G+C% regions), we focused on the roughly 3.2 Mb of sequence for annotation. The average G+C content of the genome

was 70.9%, similar to that of genomes of related actinomycetes, such as *Rhodococcus jostii* RHA1 and *Nocardia farcinica* IFM 10152 (Table 2).

Similarity and domain searches showed that the *D. cinnamea* strain P4 genome contains 3,328 predicted protein-encoding genes (CDS). The putative CDSs encode 2,455 proteins of known function, next to 389 conserved hypothetical proteins without function (source: NCBI, NR database and Conserved Domains Database (CDD)). A total of 484 proteins were unique to *D. cinnamea* P4, with no significant similarity to sequences that represented known functions in the database (Table1). Two rRNA operons were found, next to 49 tRNA genes for all 20 natural amino acids.

A major part of the strain P4 genome revealed suites of genes encoding putative proteins devoted to energy acquisition and compound transport (Fig. 2). We did not find any gene encoding flagellum synthesis and motility, which is consistent with lack of motility as described by von der Weid and colleagues (von der Weid et al., 2007). The high number of genes devoted to different carbon acquisition pathways, coupled to the versatile pathway degradation of hydrocarbons and plant compounds, and elaborate stress response systems, showed potential strategies that allow survival and competitiveness of strain P4 in soil settings. The elevated number of uncharacterized genes encoding proteins of unknown function (14%) and hypothetical proteins (10%) is in accordance with findings for other soil actinomycetes (Takarada et al., 2008; Mongodin et al., 2006; McLeod et al., 2006).

Among the 3328 CDSs of the *D. cinnamea* P4 genome, about 42% showed similarity to genes of *Rhodococcus*, 16% to those of *Mycobacterium* sp. and 13% to those of *Corynebacterium* spp., while 14% were similar to genes of *Gordonia bronchiales* DSM43247, *Nocardia farcinica* IFM10152 and *Tsukamurella paurometabola* DSM20162. The high degree of similarity to genes of *Rhodococcus* is understandable given its taxonomic proximity. In

fact, strain P4 was initially classified as a member of the genus *Rhodococcus* (Evans et al., 2004).

Alternative energy generation pathways. In spite of the fact that gene annotations do not necessarily imply functionality, genome sequences of other bacteria have provided insights into key potential functions. Thus, aerobic CO oxidizers have been uncovered, including several actinomycetes from soils (King and Weber, 2007). We here indicate the presence of a CO oxidation pathway in the strain P4 genome by the identification of genes that likely encode the aerobic oxidation of CO to CO₂, which is catalyzed by the so-called carbon monoxide dehydrogenase complex (CODH). The genome of strain P4 further revealed the presence of all genes for the large, medium and small subunits of this complex, i.e. the *coxSLM* genes (DCP0504-6) (Table S1). The distribution of these genes may differ between species (Santiago et al., 1999; Fuhrmann et al., 2003; Pelzmann et al., 2009), however the typically clustered *coxSLM* organization was indeed found in the P4 genome. The two described forms of CODH, i.e. form I (which oxidizes CO), and form II (function unknown) were both found (King and Weber, 2007). Our analyses also revealed that the large subunit of the Cox complex of strain P4 matched form I (Fig 3). Three copies of the accessory gene *coxE* (DCP40500-2) as well as a single xanthine and CODH maturation factor (*xdhC/coxF* - DCP40503) were also identified. The presence of accessory genes in addition to the CODH complex support the proposition of Santiago and colleagues (Santiago et al., 1999), that the presence of *cox* accessory genes is essential, with a CuMoO₂ metal center, composed by the Cox complex, for complete CO oxidation.

The genome of strain P4 also showed the possible use of nitrogenous organic compounds as terminal electron acceptors in oxygen-limiting conditions by dissimilatory nitrate reduction. The capacity to reduce NO₃⁻ to NO₂⁻ by strain P4 was suggested by the

finding of large and small subunit membrane-bound NarGH complexes. The accessory soluble proteins NarIJ clustered in the *narGHIJ* operon (DCP40604-07), whereas a single gene for dissimilatory membrane-bound nitrate reductase (DCP40603) was located at another chromosomal position. In addition, we found three copies of *moaA* (DCP40600-01, DCP40608), two copies of *moeA* (DCP40609, DCP40611) and a single copy of *moaC* (DCP40610). These genes clustered together with a gene for multicopper oxidase, CopA (DCP40613), which is involved in the synthesis of molybdenum cofactor, and is essential for nitrate reductase activities (Moreno-Vivián et al., 1999). The supply of nitrate for use by the NarGH complex in the reduction reaction is probably performed by the ABC-type transporter NarK (DCP40602, DCP40612), which clustered along the *narGHIJ* operon, and a second copy which was localized in another site in the P4 chromosome.

Autotrophic carbon fixation. The energy generated and conserved via aerobic CO oxidation can be directed to fix CO₂ into biomolecules. In strain P4, this may proceed by two different mechanisms, i.e. via the reductive TCA cycle (rTCA) and/or the 3-hydroxypropionate cycle. Here, we found evidence for the ability of strain P4 to employ both metabolic pathways by the presence of the main genes that are required for them (Table S1).

Although these cycles are prevalent among anaerobic bacteria, e.g. inhabitants of hydrothermal vents, in actinomycete species it was shown that the key enzymes for autotrophic CO₂ fixation can operate under microaerophilic conditions, sometimes using sulfide or hydrogen as the preferred electron donors (Rothschild, 2008). Our genome analysis revealed the presence of the multisubunit pyruvate:ferredoxin oxidoreductase (DCP40511), phosphoenolpyruvate synthase (DCP40512) and phosphoenolpyruvate carboxylase (DCP40514), which are indicative for the presence of this branch of the TCA cycle, in addition to both 2-oxoglutarate:ferredoxin oxidoreductase subunits (DCP40580-81) and

citrate lyase (DCP40513). The latter are both constituents of the normal TCA cycle, but operating in reverse mode. Autotrophic CO₂ fixation by the 3-hydroxypropionate cycle, which was first described for *Chloroflexus aurantiacus* (Herter et al., 2001, Eisenreich et al., 1993), is possible by a reaction in which two molecules of CO₂ are converted into glyoxylate by action of the large biotin-dependent acetyl/propionyl-CoA carboxylase enzyme and malonyl-CoA reductase. In the strain P4 genome, we found one cluster of genes containing malonyl-CoA reductase (DCP40515) next to genes for both subunits of acetyl/propionyl-CoA carboxylase (DCP40516-17), and the complete pathway for the biosynthesis of biotin compounds (see below). Besides these results, we assessed the utilization of CO₂ at different concentrations as the sole carbon source to spur the growth of strain P4 (see Supporting information).

The ability of strain P4 to use C1 compounds as energy and carbon sources is suggested by our finding of genes involved in formate oxidation and formaldehyde transfer (CH₂O). In formate oxidation, there are two possible mechanisms: formate can be taken up from the environment by the ABC transport system and subsequently become oxidized to CO₂ by the formate dehydrogenase complex, which is encoded by the *fdhABD* genes (DCP40575-77) (Lenger et al., 1997). The second mechanism utilizes formaldehyde as the substrate, which can be converted to formate by the bifunctional enzyme 5,10-methylene-tetrahydrofolate dehydrogenase/methenyl tetrahydrofolate cyclohydrolase (DCP40566). The subsequent reaction catalyzes the oxidation of formate to CO₂ by the formate dehydrogenase complex (DCP40575) (16). The formaldehyde transfer mechanism involves the participation of three different metabolic cycles: serine, glyoxylate regeneration and poly-β-hydroxy isobutyryl-CoA (PHB) cycles, with genes overlapping between distinct pathways. This complex mechanism found in the P4 genome is similar to the one described in

Methylobacterium extorquens strain AM1 (Chistoserdova et al., 2003). This strain utilizes tetrahydrofolate (H₄F) as a cofactor and 5,10-methylene-tetrahydrofolate dehydrogenase/methenyl tetrahydrofolate cyclohydrolase catalyzes the formation of methylene-H₄F from formaldehyde and H₄F. And, it is directed to C1 assimilation cycles by the serine, glyoxilate regeneration and PHB cycles (Table S1).

Hydrocarbon metabolism. Alkane compounds are the most abundant components of crude oil, making up to over 50% of total oil mass. Alkane molecules are chemically rather inert, but in the presence of O₂ they can be activated by oxygenases and directed to complete oxidation (van Beilen et al., 2001). In previous work, strain P4 showed to be capable of degrading compounds of crude oil and petroleum, such as differently-sized alkanes (C-11 up to C-36) and aromatic hydrocarbons (von der Weid et al., 2007). In the current genome analysis, we identified one gene cluster containing genes for enzymes that are likely involved in *n*-alkane degradation. These encompassed one TetR family transcriptional regulator (DCP41331) and one alkane 1-monooxygenase-rubredoxin fusion protein (DCP41332). We also noted the presence of the gene for a lipid transporter (DCP41332), which clusters with the previous genes. In agreement with the phylogenetic relatedness of the putative transcriptional regulator TetR with AlkS transcriptional regulators of several actinomycetes, this suggested a likely regulatory function of DCP41331 in the alkane biodegradation pathway (Canosa et al., 2000). The putative monooxygenase region revealed 96% amino acid identity with the alkane hydroxylase-rubredoxin fusion protein of *Dietzia* sp. E1. A comparison with the THMMH database indicated that the protein encoded by this gene contains six transmembrane loops, which is similar to other AlkB proteins. The rubredoxin region protein, which is involved in electron transfer, showed similarity with the rubredoxin A (RubA) protein of other actinomycetes. In addition to putative operons involved in *n*-alkane

degradation, we identified four genes that likely encode monooxygenases involved in aliphatic hydrocarbon oxidation, although we could not determine the specific substrates (DCP41334-37).

Analysis of the P4 genome further confirmed the presence of genes encoding enzymes involved in the biphenyl and benzoate degradation pathways. One chromosomal region was identified, which was found to contain the 2,3-dihydrobiphenyl 1,2-dioxygenase, *bphB* (DCP41306), 2,3-dihydroxybiphenyl 1,2-dioxygenase, *bphC* (DCP41311, DCP41327) and 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase, *bphD* (DCP41326) genes. These three genes make part of the upper biphenyl pathway. In addition, we found a LysR transcriptional regulator (DCP41307) clustering with these genes, indicating involvement in the regulation of expression of the *bphBCD* operon. The metabolites resulting from the upper pathway are 2-hydroxy-penta-2,4-dienoate and benzoate. The first compound can be converted to TCA intermediates by the lower biphenyl pathway, which is encoded by *bphEFG* genes that were also found (DCP41308-10). Moreover, benzoate is converted to catechol by the benzoate dioxygenase complex, *benABC* (DCP41315-17). Another potential metabolite, 4-hydroxybenzoate, is converted to protocatechuate by protocatechuate 4,5-dioxygenase (DCP41312). Both products can be directed to the protocatechuate and catechol branches of the β -keto adipate central pathway, respectively (Table S1) (McLeod et al., 2006; Gonçalves et al., 2006). In addition, two genes probably encoding MFS family benzoate transporters were found (DCP41300), which supports the importance of this pathway for strain P4. Another branch of the β -keto adipate pathway present in the strain P4 genome is the phenol oxidation pathway. Phenolic compounds, much like biphenyl compounds, are often toxic and may reduce the efficiency of waste treatment (Saa et al., 2010). Phenol oxidation is catalyzed by phenol monooxygenase (DCP41328), yielding catechol, which is subsequently directed to the

TCA cycle.

Other oxygenases. Oxygenases are enzymes that catalyze the insertion of oxygen atoms from O₂ into an organic substrate. They consist of two types: enzymes that catalyze the insertion of a single oxygen atom, monooxygenases, and enzymes that catalyze the insertion of two oxygen atoms, dioxygenases. Reactions in which organic compounds are oxygenated have great value for organic synthesis, thus there is great interest of biotechnology industries. Genomic analysis of P4 revealed the presence of 48 monooxygenases and 19 dioxygenases (Table S2). The number of oxygenases present in the P4 genome is relatively high, as judged by comparing the number of oxygenases per Mb across the genomes of different species (Table1). *Rhodococcus josti* RHA1 and *Burkholderia xenovorans* LB400, strains known to have large numbers of oxygenases in their genomes, have about 20.9 and 13.8 oxygenases per Mb respectively, while strain P4 has 19.1. Closer examination revealed the presence of some genes for monooxygenases involved in the degradation of aliphatic and cyclic hydrocarbons, such as cyclohexanone, phenylacetate and 4-hydroxyacetophenone (Table S2). The number of cytochrome P450 genes in the P4 genome was nine, which is in line with that in other actinomycetes (Mongodin et al., 2006). Among the cytochrome P450's of strain P4, genes encoding linaol monooxygenase and CypX stood out. Both are involved in secondary metabolism (Table S2).

We identified coding genes for 3-keto 9- β -hydroxylase, which is responsible for opening the ring of cholesterol compounds. In addition, we noted two coding genes for cholesterol oxidase. There are two different classes of cholesterol oxidase, each with two subclasses. Class I contains the FAD cofactor non-covalently bound to the enzyme; this class occurs mainly in actinomycetes. Class II contains the cofactor covalently linked to the enzyme, with Gram-negative bacteria as main representatives. The P4 genome was found to

contain two different cholesterol oxidases of class I (DCP41295-96). Phylogenetic and similarity analyses suggested that CDSs DCP41296 and DCP41295 were distant from each other, being the first similar to class I-1, and the second to class I-2. The latter lacks a signal sequence, suggesting cytoplasmic localization (Navas et al., 2001). In addition, strain P4 also possessed genes described for catabolic pathways: specifically, genes encoding lactate monooxygenase, dimethylalanine monooxygenase, glycolate oxidase and lysine/ornithine monooxygenase (Table S2).

Secondary metabolism. The genome sequence of strain P4 revealed genes coding for the biosynthesis of terpene compounds, non-ribosomal peptide and polyketide synthases, which were distributed non-uniformly over the P4 genome. We found genes of the so-called DOXP/MEP pathways of terpenoid backbone biosynthesis, responsible for the synthesis of geranyl biphosphate and undecaprenyl biphosphate compounds (DCP41417-28). Subsequently, the geranyl can be converted into (S)-carvone by limonene 6-monooxygenase and carveol dehydrogenase genes (DCP41429-30). The full pathway for L-rhamnose biosynthesis was also detected in the P4 genome (DCP414211-14). L-rhamnose is a component of cell walls of pathogens which spurs virulence, and so this characteristic might confer virulence characteristics on this bacterium. We also found the complete pathway for the secondary carotenoid metabolite that is probably connected to UV-resistance (see below).

Polyketide synthases (PKS) catalyze a decarboxylative condensation of simple metabolites like acetate, malonate, propionate, butyrate and their derivatives into growing polyketide chains (van der Ploeg et al., 2001). Microbial PKSs are divided into three major types. The PKS type I system consists of one or more multifunctional proteins that contain a different active site for each modification. PKS type II comprises individual proteins for the synthesis of aromatic products. PKS type III is composed of homodimers grouped in the

synthase superfamily. In addition, there is an atypical PKS type, which catalyzes iterative condensation of extenders in a polyketide chain, called iterative type I PKS. We identified 9 putative polyketide synthases of PKS type II, and a single regulator of polyketide synthase expression. We deduced the presence of three coding genes that likely catalyze the synthesis of new antibiotics: a gene homologous to *aviGT4*, the product of which is responsible for the synthesis of avilamycin A (DCP41439), one akin to *fnrE*, which is involved in the biosynthesis of the aromatic antibiotics frenocilin and nanaomycin (DCP41432) and one similar to *ncnH*, possibly involved in biosynthesis of the aromatic antibiotic bnapthocyclinone (DCP41441). The latter antibiotic exhibits activity against Gram-positive bacteria (Sthapit et al., 2004). The genome annotation further identified a single copy of a gene for naphthoate synthase, a component of the interactive type I PKS family that participates in the biosynthetic pathway of Azinomycin B, a complex natural polyketide with strong antitumor activity (Brünke et al., 2001).

Stress responses and DNA repair mechanisms. The strain P4 genome showed elaborate systems that allow tolerance to environmental stresses. Also, multiple DNA repair mechanisms were found. Strain P4 may resist reactive oxygen radicals [produced by redox reactions in aerobic metabolism] using three superoxide dismutases, namely two copies of SodA (DCP45063-65) and a single Cu-Zn SodC (DCP45041). Furthermore, four heme-catalases (DCP45044-45, DCP45055, DCP45073), two Mn-catalases (DCP45046-47), two peroxisomal catalases (DCP45043, DCP45056) and two peroxidases (DCP45042, DCP45057) may be involved. We also noted the transcriptional regulator OxyR (DCP45058), which - at high levels of hydrogen peroxide (H_2O_2) - activates the detoxifying enzyme alkyl peroxidase (DCP45048).

The strain P4 genome was further found to contain genes involved in thiol redox

control, e.g. thioredoxin (DCP45067) and thioredoxin reductase (DCP45069, DCP45062), glutathione peroxidase (DCP45060) and glutathione reductase (DCP45049). Also, we found the organic hydroperoxide resistance gene (DCP45059), which is involved in detoxifying organic oxidative agents, methionine sulfoxide reductase and peroxyredoxin. In addition, we detected a gene cluster composed of genes homologous to *sufBCDSE* (DCP45074-79). This cluster is likely expressed under conditions of iron starvation and oxidative stress, working in protection or repair of [Fe-S] proteins (Zhao et al., 2008).

Strain P4 probably repairs damage to its DNA through nucleotide excision, as evidenced by the finding of *uvrA*, *uvrBC* and *uvrD*-like helicases, genes for DNA alkylation repair and DNA glycosylases. Furthermore, pyrimidine dimer resolving enzymes, e.g. photo lyase, and recombination repair proteins - *ruvAB* and *recAB* – were found. Also, other components of the recombination apparatus, *recNOQ*, and other enzymes of Holliday junction resolution were found. Four copies of the Nudix family of genes (which has as a role in cleaning the cell of toxic nucleotide metabolites), were found. In addition, we noted the presence of phytoene synthetase (DCP41407), phytoene dehydrogenase (DCP41404, DCP41408) and lycopene cyclase (DCP41400) genes. These genes compose the β -carotene biosynthesis pathway, which is suggested to confer radiation resistance by scavenging electrons from reactive oxygen species (Bessman et al., 1996).

The ability of *D. cinnamea* P4 to protect itself from toxic effects of heavy metals was evidenced by the 32 putative coding genes identified that are potentially involved in conferring resistance against these compounds (Table S1). Genes for arsenic resistance (through extrusion) and inactivation by reduction were localized at two different sites in the genome. The first was a region composed of genes for an arsenite export protein (DCP45011) and a single arsenate reductase (DCP45012), the second contained two transporter systems in

tandem, probably involved as an efflux pump of arsenic compounds (DCP45016). Also, genes for mercury resistance were detected: one putative mercury reductase protein (DCP45016) and one transporter system involved in heavy metal translocation (DCP45015). Interestingly, two transcriptional regulators were found to be localized at different sites in the genome (DCP45005, DCP45018), beside other transporter systems involved in heavy metal translocation, such as copper, zinc and cadmium. Although we found three copies of *terC* genes (DCP45026-27, DCP45032), responsible for resistance to tellurium ions; other genes involved in this, eg. *terBDE*, were missing. Resistance to copper was associated with the presence of the sensor genes *copR* (DCP45014) and *cusS* (DCP45014), three copies of copper-translocating P-type ATPase *copA* (DCP45008, DCP45021, DCP45023) and one of *copC* (DCP45004). This suggested one mechanism of copper trafficking, triggered by redox-dependent coordination properties of the metal (Phillips et al., 2002). In spite of some metals being necessary for cell homeostasis, a single gene coding for a protein involved in the efflux of iron was detected, which probably has a role in the resistance to iron at high concentrations. Another seven ATPase efflux transporter systems for cadmium, cobalt and zinc were found in the P4 genome, showing the importance of metal resistance for strain P4 ecology (Table S1).

The genome information further showed that strain P4 may sense osmotic stress by virtue of an osmolarity sensor protein (DCP45113), after which a response is triggered. The internal osmotic equilibrium is probably maintained by accumulation of organic osmolytes (compatible solute strategy), which are synthesized by the enzymes choline dehydrogenase (DCP45111) and betaine aldehyde dehydrogenase (DCP45110, DCP45112). We also noted the presence of genes encoding transport mechanisms for these compounds (DCP45114-20). Also, we identified the so-called *ehuAB* genes, which are involved in the uptake of

extracellular ectoine compounds (DCP43050-51). Intracellular accumulation of trehalose and glycogen has been reported under water stress conditions. The genes in the P4 genome responsible for trehalose biosynthesis from glycogen were found to be clustered in the *treYXZ* operon (DCP40100-02). TreX is a glycogen-debranching enzyme, yielding 1,4- α -glucan compounds, TreY reduces the end of 1,4- α -glucan to form trehalose residues and TreZ liberates trehalose from the product of the TreY-catalyzed reaction (Arnesano et al., 2003). In addition, the presence of an aquaporine Z gene (DCP45121) suggests the role of control of water movement in efforts to osmotic stress tolerance.

One of the desirable characteristics in a strain with potential for bioremediation/biocatalysis is tolerance to solvents. Strain P4 is able to grow in oil, diesel, benzene and alkanes (von der Weid et al., 2007). This tolerance is probably due to the presence of an ABC transporter system encoded by the *tigABC* operon. These genes may be constitutively expressed, as their control is normally accomplished by the transcriptional repressor *tigR*, which we could not find in the genome annotation. The genome of strain P4 further possesses universal stress-related proteins (USPs) (DCP45136), which are involved in a wide variety of responses to environmental stress, such as carbon, nitrogen and phosphate starvation, exposure to heat shock and UV radiation. Also, we found the gene for the transcriptional regulator of heat shock, HcrA (DCP45127), *dnaJ-dnaK* (DCP45128, DCP45139-40, DCP45145), *clpBCX* (encoding chaperonins) (DCP45131, DCP45135), *groEL* (DCP45133), *hsp90* (DCP45141), and cold shock (DCP45132).

Among the genes that are potentially involved in survival and adaptation of P4 in soil, a wide repertoire of antibiotic resistance genes was found. In total, 12 genes encoding B and C classes beta-lactamases and penicillin biosynthesis were found, next to at least 12 putative ATP-ABC and 10 MFS transport systems (Table S1). The latter are likely involved in the

efflux of drugs like enterobactin (DCP45110) and daunorubicin (DCP43123). We found four copies of genes coding for aminoglycoside 3'-phosphotransferases (DCP45087), which are involved in the modification/efflux of neomycin and kanamycin, next to two genes involved in export of polyketide antibiotics (DCP45092, DCP45126). One gene coding for the radical SAM 23-S rRNA Methyltransferase of the Cfr family (DCP45088) was found. This is thought to perform post-transcriptional modification of ribosomal RNA, thus conferring resistance to ribosome-targeting antibiotics (Kaminska et al., 2010). Also, we identified a single copy of a gene for erythromycin esterase (DCP45096), which protects against an own antibiotic.

Transport of molecules. Analysis of the P4 genome using the THMMH, NR, CDD and TransportDB databases revealed the presence of about 311 CDSs with putative transport roles (DCP43000-43307). This value represents about 8.8% of the whole P4 genome, which is consistent with values from other sequenced actinomycete genomes. The best-represented transporter class was the ABC family (49%), followed by the MFS family (13%). In both families, the presence of genes involved in drug export, solvent and metal efflux stood out. This may reflect a requirement for these systems in an environment like soil, where the presence of antibiotic producers is likely. Hence, the ability to compete in the niche is enhanced by the use of efficient antibiotic efflux systems next to the synthesis of own antimicrobials. In addition to the uptake transporters described above, the P4 genome contains several genes for permeases for diverse sugar classes, glycerol, nucleosides, amino acids, acetate, urea, vitamin B12 and C4-dicarboxylate. The *secADEFY* genes (DCP43190-3204), which encode the Type II secretion system, were also detected in P4 genome. The *tatC* translocase gene (DCP43197) was also found, this encodes a component of the Sec-independent protein translocation system, is involved in translocation of folded proteins across of periplasmic membrane.

Concluding remarks

The tropical soil from which strain P4 was isolated is characterized as a complex environment with frequent chemical and physical inputs. The resulting energy flux is mainly due to the primary production, which is reflected in root exudation as well as in the decomposition of plant parts. The current annotation of the P4 genome assists us in gaining new insights in the mechanisms employed by *D. cinnamea* P4 in its adaptation to soil. The presumably saprophytic lifestyle of *D. cinnamea* P4 is reflected in the complex network of gene products involved in the acquisition of carbon, as well as in the seemingly efficient processes of protection against environmental stress. The diverse array of ABC- and MFS-like transporters, which are predicted to be involved in the acquisition of different materials from soil, next to the complex catabolic mechanisms for the (simultaneous) use of plant compounds like limone, pinene, alkane/alkanesulfonate, cyclic alkanes, nitrogen compounds and complex carbohydrates, highlights the strategies for growth and survival in a complex habitat. The genome sequence also confirmed the existence of pathways for the transformation of aromatic and phenolic compounds which are generated following decomposition of plant materials. Here, the assimilation of aromatic compounds like biphenyl, benzoate and phenol converge to the common β -ketoacid pathway. The ability to use aerobic CO oxidation and a possible anaerobic dissimilative nitrate reduction pathway as energy generation mechanisms, and the fixation of CO₂, support the capacity to grow autotrophically. However, an aerobic heterotrophic lifestyle may predominate in the organism. The harsh climate can be inhospitable to microorganisms dwelling in tropical soils, as alternating periods of drought and high humidity, high sun incidence raising the temperature, and the presence of chemicals constitute typical stressors. Another factor of importance for the fitness of strain P4 is the high

number of genes devoted to a rapid response to stress signals. Strain P4 was shown to possess elaborate systems for osmotic stress control, also revealing the complete pathway for synthesis of carotenoid compounds that confer UV and oxidative stress resistance. Moreover, the multiplicity of genes devoted to antibiotic resistance in the P4 genome reveals the importance of drug tolerance in its habitat.

Finally, the *D. cinnamea* P4 genome was found to encode a high number of enzymes with potential application in biocatalysis. Oxygenases involved in secondary metabolite biosynthesis, e.g. of carveol compounds and genes for the biosynthesis of aromatic antibiotics exemplify some products of commercial interest. Finally, the unexpected metabolic capabilities of *D. cinnamea* P4, combined with the tolerance of this organism to environmental stresses and its features of easy cultivation in bioreactors, provides a stimulus to future applications in biocatalysis and bioremediation.

Material and Methods

Strain and growth. *D. cinnamea* P4 (DSM44994, CIP109208) was previously isolated from samples of an acidic sandy loam (Cambisol) soil collected from an environmentally protected area of tropical Atlantic forest (Biological Reserve of Poço das Antas) located in Rio de Janeiro state, Brazil. It was cultivated and maintained in Luria–Bertani (LB; tryptone 1%, NaCl 0.5%, yeast extract 0.5%) medium at 30°C under shaking up to exponential phase. Genomic DNA was obtained from strain P4 as described by Seldin and Dubnau (1985).

Escherichia coli K12A15 wild-type is proficient in all DNA repair/damage tolerance systems. This strain was used as a positive control in tests for cellular resistance to stressors (4). In contrast, *Escherichia coli* BH980 (*mutY*[−]) mutant strain is deficient in base excision repair mediated by the MutY protein. It was used as the positive control in tests of

spontaneous mutagenesis (Nghiem et al.,1988,).

Hydrocarbon degradation and growth experiments with *D. cinnamea* P4 were performed in minimal Bushnell Hass Broth (BHB), supplemented with the appropriate carbon source. Strain P4 was initially grown on complex Luria Bertani (LB) broth until mid-exponential phase, then cells were harvested and cell pellets washed 3X in 1X phosphate-buffered saline (PBS). About 10^4 cells were used as an inoculum into 30 ml of BHB in 250-ml bottles. *n*-hexadecane and benzene were added at final concentrations of 1% and 0.5% respectively, and cultures were incubated at 28°C with shaking at 200 rpm. Samples from each bottle were taken on a daily basis to assess growth by CFU counting. All experiments were conducted in triplicate and at least two independent growth experiments were conducted for each carbon source. Growth experiments were also performed in 2-L bioreactors in aerobic conditions using 1 L of BHB supplemented with *n*-hexadecane at 1%. Here, we analyzed growth at pH 5 to 9, and temperatures from 25°C to 45°C. At least two independent growth experiments were conducted for each condition.

Survival experiments. Cell cultures in exponential growth phase were exposed to a series of physical and chemical agents. Surviving cell fractions were determined by the ratio between the number of remaining cells after and the number of viable cells before each treatment. UV-A, UV-B and UV-C were used as the stress factors, as follows. Cultures in the exponential growth phase ($\sim 2 \times 10^8$ CFU/ml in 10 ml LB) were centrifuged at $6,000 \times g$ at 4°C, washed twice with phosphate buffer (PB), as described above. For ultraviolet C (UV-C), cell suspensions were irradiated with a General Electric 15-W germicidal lamp (G15T8, 254 nm) with increasing doses. After each treatment, aliquots were taken and plated on solid LB. Plates were incubated for 24–48 h and scored for remaining CFUs. UV strength was determined by means of a Latarjet dosimeter. For ultraviolet B (UV-B), a VL215 lamp (312

nm) was used. The strength was also determined by using a VLX-3W dosimeter with a UV-B photocell. For ultraviolet A (UV-A), irradiation was performed with a VL212 lamp (365 nm), while fluence was determined with a Black Ray Dosimeter with a long wave photocell.

Chemical agents. *Hydrogen peroxide (H₂O₂) and Cumene hydroperoxide (CH)* - Resistance to the oxidative agents H₂O₂ (inorganic) and CH (organic) was determined at three different concentrations: 0.1, 1 and 10 M. Cells were grown in LB medium as described. After, 50 µl were taken and added to 2.5 ml of molten top agar. The mixture was poured onto LB plates and left at room temperature for 5 minutes. Whatman filters (#1) were placed on the plates (four per plate) and soaked with 10 µl of each H₂O₂ or CH solution. Plates were kept in the dark at 4°C for 1 hour, then incubated for 24h at 30°C and finally observed to determine growth inhibition haloes. A relative scale was used, ranging from one to four crosses (+), respectively 2 up to 12 mm haloes.

Mutagenesis. Both spontaneous and induced mutagenesis were evaluated using the rifampicin resistance system. Cells were challenged with UV-A, UV-B and UV-C to evaluate induced mutagenesis and then monitored for the appearance of rifampicin-resistant mutants (Rif^R) over the spontaneous levels. Cells in the exponential growth phase were diluted and plated on solid LB, while undiluted aliquots were plated on LB containing 100 µg/ml rifampicin. After incubation at 30°C for 24–48 h, colonies were scored and mutation frequency was determined. Mutagenesis was expressed as the frequency of Rif^R mutants per 10⁸ cells (Rif^R/10⁸ cells).

Statistical analyses of survival and mutagenesis experiments. The results are presented as means of at least three experiments with respective standard errors. Data were analyzed by ANOVA followed by the Student Newman Keuls multiple comparison test using the statistical program InStat version 3.01 (GraphPad Software, San Diego, CA, USA). A 5% level of

significance was adopted to evaluate the data.

Pyrosequencing, genome analysis and annotation. Pyrosequencing was performed (one run) on a FLX Roche apparatus at the University of Copenhagen, Copenhagen, Denmark. Over 500,000 reads of good quality were obtained. Following contig formation by automated software, the genome was automatically annotated, after which all annotation was manually controlled and analyzed. The total of sequenced bases was 3579857, which corresponds to about 99% of all P4 genome. Predicted protein-encoding genes were identified using GLIMMER v3.02 (48) and EMBOSS for prokaryotes, followed by automatic annotation using search results against the Pfam, Smart, COG and CDD databases (see respective websites). Metabolic pathways found were categorized using the KEGG database. Transmembrane domains were identified using TransportDB or THMMH (see website) and SignalP was used to predict signal peptide regions (<http://www.cbs.dtu.dk/services/SignalP/>) using the hidden Markov model (HMM) algorithm. All annotation was done manually by using NCBI BLASTP against the non-redundant database. Insertion sequences were found by using the IS finder database (<http://www-is.biotoul.fr/is.html>). Ribosomal RNA (rRNA) genes were detected by searching the whole genome sequence against rRNA databases with NCBI BLASTN and transfer RNAs by using tRNAscan-SE (45). The G+C content was determined by using the software EMBOSS (see website). The total nucleotide sequence and annotation of *Dietzia cinnamea* P4 is deposited in GenBank under accession no. AEKG00000000.

Acknowledgements

Thanks are due to CNPq and Coimbra Institute for providing a fellowship to L. Procopio to come over to the Groningen lab, where the genome annotation and physiological experiments were performed. This study was carried out as part of a research project supported by PETROBRAS.

References

- Yassin A.F., Hupfer H. and Schaal K.P. (2006) *Dietzia cinnamea* sp. nov., a novel species isolated from a perianal swab of a patient with a bone marrow transplant. *Int J Syst Evol Microbiol.* 56(3):641-45.
- Evans F.F., Rosado A.S., Casella R., Sebastian G.V., Machado P.L.O.A., Holmström C., Kjelleberg S., van Elsas J.D. and Seldin L. (2004) Impact of oil contamination and biostimulation on the diversity of indigenous bacterial communities in soil microcosms. *FEMS Microbiol. Ecol.* 49:295-305.
- von der Weid I., Marques J.M., Cunha C.D., Lippi R.K., Dos Santos S.C., Rosado A.S., Lins U. and Seldin L. (2007) Identification and biodegradation potential of a novel strain of *Dietzia cinnamea* isolated from a petroleum-contaminated tropical soil. *Syst Appl Microbiol. Syst Appl Microbiol.* 30(4):331-9.
- Takarada H., Sekine M., Kosugi H., Matsuo Y., Fujisawa T., Omata S., Kishi E., Shimizu A., Tsukatani N., Tanikawa S., Fujita N. And Harayama S. (2008) Complete genome sequence of the soil actinomycete *Kocuria rhizophila*. *J Bacteriol.* 190(12):4139-46.
- Mongodin E.F., Shapir N., Daugherty S.C., Deboy R.T., Emerson J.B., Shvartzbeyn A., Radune D., Vamathevan J., Riggs F., Grinberg V., Khouri H., Wackett L.P., Nelson K.E. and Sadowsky M.J. (2006) Secrets of soil survival revealed by the genome sequence of *Arthrobacter aurescens* TC1. *PLoS Genet.* 2(12):e214.
- McLeod M.P., Warren R.L., Hsiao W.W., Araki N., Myhre M., Fernandes C., Miyazawa D., Wong W., Lillquist A.L., Wang D., Dosanjh M., Hara H., Petrescu A., Morin R.D., Yang G., Stott J.M., Schein J.E., Shin H., Smailus D., Siddiqui A.S., Marra M.A., Jones S.J., Holt R., Brinkman F.S., Miyauchi K., Fukuda M., Davies J.E., Mohn W.W. and Eltis L.D. (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. *Proc Natl Acad Sci U S A.* 103(42):15582-87.
- King G.M. and Weber C.F. (2007) Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nat Rev Microbiol.* 5(2):107-18.
- Santiago B., Schübel U., Egelseer C. and Meyer O. (1999.) Sequence analysis, characterization and CO-specific transcription of the *cox* gene cluster on the megaplasmid pHCG3 of *Oligotropha carboxidovorans*. *Gene.* 236(1):115-24.
- Fuhrmann S., Ferner M., Jeffke T., Henne A., Gottschalk G. and Meyer O. (2003) Complete nucleotide sequence of the circular megaplasmid pHCG3 of *Oligotropha*

carboxidovorans: function in the chemolithoautotrophic utilization of CO, H₂ and CO₂. *Gene*. 322:67-75.

- Pelzmann A., Ferner M., Gnida M., Meyer-Klaucke W., Maisel T. and Meyer O. (2009) The CoxD protein of *Oligotropha carboxidovorans* is a predicted AAA+ ATPase chaperone involved in the biogenesis of the CO dehydrogenase [CuSMoO₂] cluster. *J Biol Chem*. 284(14):9578-86.
- Santiago B., Schuebel U., Egelseer C. and Meyer, O. (1999) Sequence analysis, characterization and CO-specific transcription of the cox gene cluster on the megaplasmid pHCG3 of *Oligotropha carboxidovorans*. *Gene* 236, 115–124.
- Moreno-Vivián C., Cabello P., Martínez-Luque M., Blasco R. and Castillo F. (1999) Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases. *J Bacteriol*. 181(21):6573-84.
- Rothschild L.J. (2008) The evolution of photosynthesis...again? *Philos Trans R Soc Lond B Biol Sci*. 363(1504):2787-801.
- Herter S., Farfsing J., Gad'On N., Rieder C., Eisenreich W., Bacher A. and Fuchs G. (2001) Autotrophic CO₂ fixation by *Chloroflexus aurantiacus*: study of glyoxylate formation and assimilation via the 3-hydroxypropionate cycle. *J Bacteriol*. 183(14):4305-16.
- Eisenreich W., Strauss G., Werz U., Fuchs G. and Bacher A. (1993) Retrobiosynthetic analysis of carbon fixation in the phototrophic eubacterium *Chloroflexus aurantiacus*. *Eur J Biochem*. 215(3):619-32.
- Canosa I., Sánchez-Romero J.M., Yuste L. and Rojo F. (2000) A positive feedback mechanism controls expression of AlkS, the transcriptional regulator of the *Pseudomonas oleovorans* alkane degradation pathway. *Mol Microbiol*. 35(4):791-9.
- Lenger R., Herrmann U., Gross R., Simon J. and Kröger A. (1997) Structure and function of a second gene cluster encoding the formate dehydrogenase of *Wolinella succinogenes*. *Eur J Biochem*. 246(3):646-51.
- Chistoserdova L., Laukel M., Portais J.C., Vorholt J.A. and Lidstrom M.E. (2003) Multiple formate dehydrogenase enzymes in the facultative methylotroph *Methylobacterium extorquens* AM1 are dispensable for growth on methanol. *J Bacteriol*. 186(1):222-8.
- van Beilen J.B., Panke S., Lucchini S., Franchini A.G., Röthlisberger M. and Witholt B. (2001) Analysis of *Pseudomonas putida* alkane-degradation gene clusters and flanking insertion sequences: evolution and regulation of the alk genes. *Microbiology*. 147(Pt

6):1621-30.

- Gonçalves E.R., Hara H., Miyazawa D., Davies J.E., Eltis L.D. and Mohn W.W. (2006) Transcriptomic assessment of isozymes in the biphenyl pathway of *Rhodococcus* sp. strain RHA1. *Appl Environ Microbiol.* 72(9):6183-93.
- Saa L., Jaureguibeitia A., Largo E., Llama M.J. and Serra J.L. (2010) Cloning, purification and characterization of two components of phenol hydroxylase from *Rhodococcus erythropolis* UPV-1. *Appl Microbiol Biotechnol.* 86(1):201-11.
- Mongodin E.F., Shapir N., Daugherty S.C., DeBoy R.T., Emerson J.B., Shvartzbeyn A., Radune D., Vamathevan J., Riggs F., Grinberg V., Khouri H., Wackett L.P., Nelson K.E. and Sadowsky M.J. (2006) Secrets of soil survival revealed by the genome sequence of *Arthrobacter aurescens* TC1. *PLoS Genet.* 2(12):e214.
- Navas J., González-Zorn B., Ladrón N., Garrido P. and Vázquez-Boland J.A. (2001) Identification and mutagenesis by allelic exchange of *choE*, encoding a cholesterol oxidase from the intracellular pathogen *Rhodococcus equi*. *J Bacteriol.* 183(16):4796-805.
- van der Ploeg J.R., Barone M. and Leisinger T. (2001) Functional analysis of the *Bacillus subtilis* *cysK* and *cysJI* genes. *FEMS Microbiol Lett.* 201(1):29-35.
- Sthapit B., Oh T.J., Lamichhane R., Liou K., Lee H.C., Kim C.G. and Sohng J.K. (2004) Neocarzinostatin naphthoate synthase: an unique iterative type I PKS from neocarzinostatin producer *Streptomyces carzinostaticus*. *FEBS Lett.* 566(1-3):201-6.
- Brünke P., Sterner O., Bailey J.E. and Minas W. (2001) Heterologous expression of the naphthocyclinone hydroxylase gene from *Streptomyces arenae* for production of novel hybrid polyketides. *Antonie Van Leeuwenhoek.* 79(3-4):235-45.
- Zhao Q., He Q., Ding W., Tang M., Kang Q., Yu Y., Deng W., Zhang Q., Fang J., Tang G. and Liu W. (2008) Characterization of the azinomycin B biosynthetic gene cluster revealing a different iterative type I polyketide synthase for naphthoate biosynthesis. *Chem Biol.* 15(7):693-705.
- Bessman M.J., Frick D.N. and O'Handley S.F. (1996) The MutT proteins or "Nudix" hydrolases, a family of versatile, widely distributed, "housecleaning" enzymes. *J Biol Chem.* 271(41):25059-62.
- Phillips R.W., Wiegel J., Berry C.J., Fliermans C., Peacock A.D., White D.C. and Shimkets L.J. (2002) *Kineococcus radiotolerans* sp. nov., a radiation-resistant, gram-

- positive bacterium. *Int J Syst Evol Microbiol.* 52(Pt 3):933-8.
- Arnesano F., Banci L., Bertini I., Mangani S. and Thompson A.R. (2003) A redox switch in CopC: an intriguing copper trafficking protein that binds copper(I) and copper(II) at different sites. *Proc Natl Acad Sci U S A.* 100(7):3814-9.
- Kaminska K.H., Purta E., Hansen L.H., Bujnicki J.M., Vester B. and Long K.S. (2010) Insights into the structure, function and evolution of the radical-SAM 23S rRNA methyltransferase Cfr that confers antibiotic resistance in bacteria. *Nucleic Acids Res.* 38(5):1652-63.
- Seldin L. and Dubnau L. (1985) DNA homology among *Bacillus polymyxa*, *Bacillus macerans*, *Bacillus azotofixans* and other nitrogen fixing *Bacillus* strains. *Int. J. Syst. Bacteriol.* 35, 151-54.
- Nghiem Y., Cabrera M., Cupples C.G. and Miller J.H. The *mutY* gene: a mutator locus in *Escherichia coli* that generates G.C----T.A transversions. *Proc Natl Acad Sci U S A.* 85(8):2709-13.

Legends to figures

Fig. 1. Survival of *D. cinnamea* P4 and *E. coli* AB1157 after UV-A, UV-B and UV-C treatment. Cultures in exponential growth phase were treated with increasing doses of UV radiation (365 nm, 312 nm, 254 nm, respectively) as described in Materials and Methods.

Fig. 2. Functional distribution of class proteins over the 435 contigs based on KEGG categorization and COG classifications. The **Others** column include oxygenases, cell division, multidrug resistance and cell wall biogenesis proteins.

Fig. 3. Phylogenetic relationship between amino acid sequence DCP0505 of *D. cinnamea* P4 and the sequences of form I and II large carbon monoxide dehydrogenase proteins (CoxL) of related organisms.

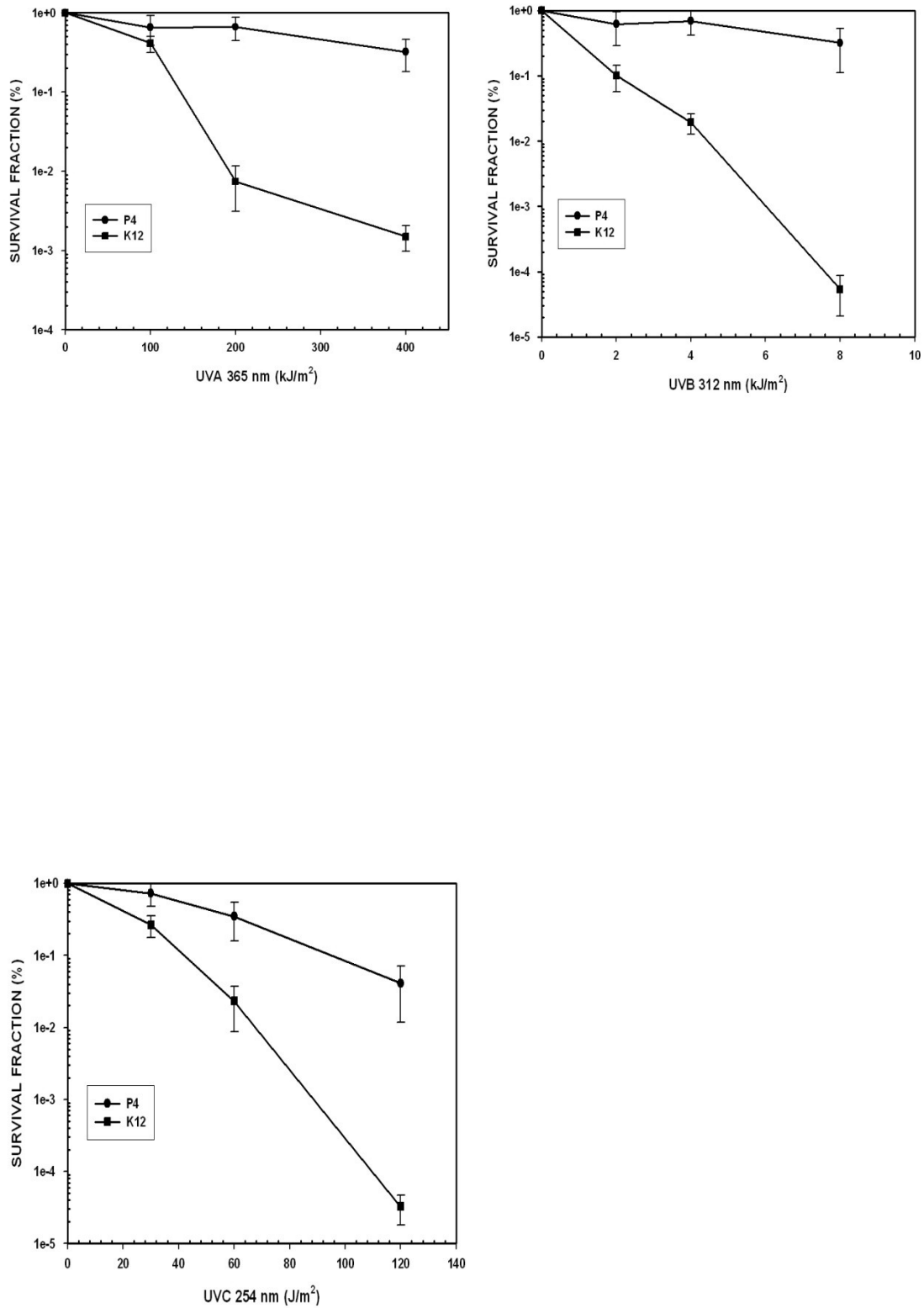


Fig. 1. Survival of *E. coli* AB1157 and P4 strains after UVA, UVB and UVC treatment. Cultures in exponential growth phase were treated with increasing doses of UVA radiation

(365 nm, 312 nm, 254 nm, respectively) as described in Materials and Methods.

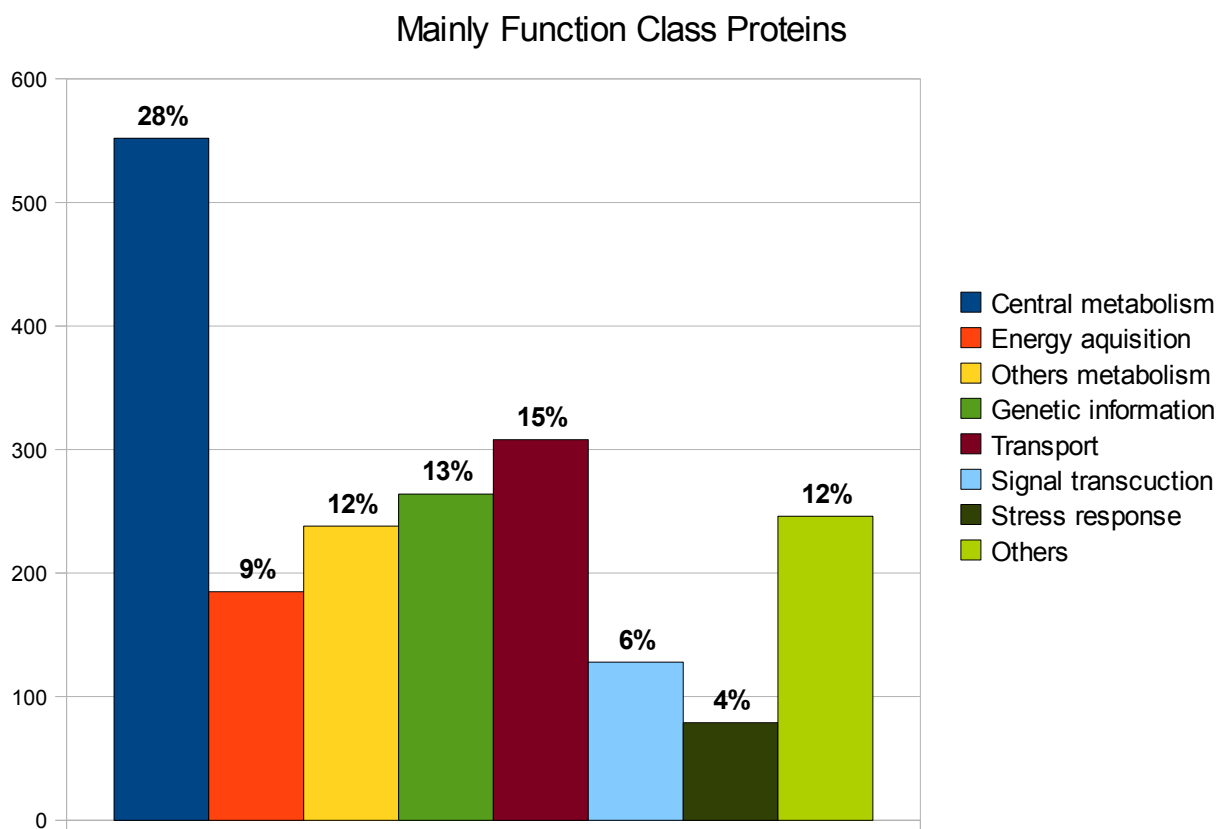


Fig. 2. Functional distribution of class proteins over the 435 contigs based on KEGG categorization and COG classifications. The **Others** column include oxygenases, cell division, multidrug resistance and cell wall biogenesis proteins.

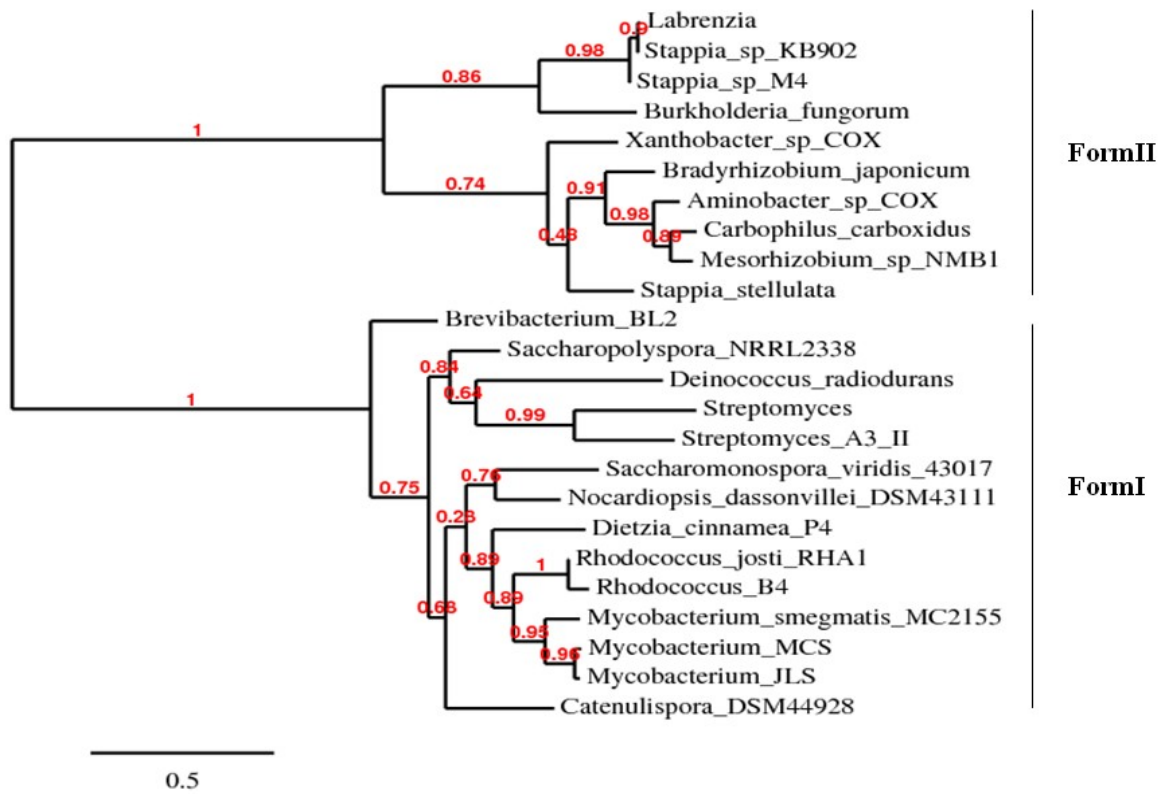


Fig. 3. Phylogenetic relationship between amino acid sequence DCP0505 of *D. cinnamea* P4 and the sequences of form I and II large carbon monoxide dehydrogenase proteins (CoxL) of related organisms.

Livros Grátis

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

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

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