TIAGO HENRIQUE NOGUEIRA SIMÕES

TRIAGEM DE ENZIMAS ASSOCIADAS À BIOTRANSFORMAÇÃO
DE HIDROCARBONETOS A PARTIR DE METAGENOMA DE
SEDIMENTOS CONTAMINADOS COM PETRÓLEO E METAIS
PESADOS

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

**São Paulo
2009**

TIAGO HENRIQUE NOGUEIRA SIMÕES

TRIAGEM DE ENZIMAS ASSOCIADAS À BIOTRANSFORMAÇÃO
DE HIDROCARBONETOS A PARTIR DE METAGENOMA DE
SEDIMENTOS CONTAMINADOS COM PETRÓLEO E METAIS
PESADOS

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

Área de Concentração: Biotecnologia

Orientador: Gilson Paulo Manfio

Co- Orientadora: Fabiana Fantinatti
Garboggini

**São Paulo
2009**

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

© reprodução total

Simões, Tiago Henrique Nogueira.

Triagem de enzimas associadas à biotransformação de hidrocarbonetos a partir de metagenoma de sedimentos contaminados com petróleo e metais pesados / Tiago Henrique Nogueira Simões. - São Paulo, 2009.

Orientador: Gilson Paulo Manfio.

Tese (Doutorado) – Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação Interunidades em Biotecnologia USP/IPT/Instituto Butantan. Área de concentração: Biotecnologia. Linha de pesquisa: Microbiologia Aplicada.

Versão do título para o inglês: Screening of enzymes related to biotransformation of hydrocarbons from metagenome of contaminated sediments with oil and heavy metals.

Descritores: 1. Bibliotecas metagenômicas 2. 16S Rdna 3. Triagem de alto desempenho 4. Esterases 5. Epoxido-hidrolases 6. Monoxigenases I. Manfio, Gilson Paulo II. Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós Graduação em Biotecnologia III. Título.

ICB/SBIB34/2009

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

Candidato(a): Tiago Henrique Nogueira Simões.

Título da Tese: Triagem de enzimas associadas à biotransformação de hidrocarbonetos a partir de metagenoma de sedimentos contaminados com petróleo e metais pesados .

Orientador(a): Gilson Paulo Manfio.

A Comissão Julgadora dos trabalhos de Defesa da Tese de Doutorado, em sessão pública realizada a/...../....., considerou

() Aprovado(a)

() Reprovado(a)

Examinador(a): Assinatura:
 Nome:
 Instituição:

Examinador(a): Assinatura:
 Nome:
 Instituição:

Examinador(a): Assinatura:
 Nome:
 Instituição:

Examinador(a): Assinatura:
 Nome:
 Instituição:

Presidente: Assinatura:
 Nome:
 Instituição:

CERTIFICADO DE ISENÇÃO DO COMITÊ DE ÉTICA

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS

Cidade Universitária "Armando de Salles Oliveira"
Av. Prof. Lineu Prestes, 2415 - CEP. 05508-000 São Paulo, SP - Brasil
Telefone : (55) (11) 3091-7733 - telefax : (55) (11) 3091-7438
e-mail: cep@icb.usp.br

Comissão de Ética em Pesquisa

CERTIFICADO DE ISENÇÃO

Certificamos que o Protocolo CEP-ICB N° 280, referente ao projeto intitulado: "*Triagem de enzimas associadas à biotransformação de hidrocarbonetos a partir de metagenoma de sedimentos contaminados com petróleo e metais pesados*" sob a responsabilidade de Tiago Henrique Nogueira Simões, foi analisado na presente data pela CEEA - COMISSÃO DE ÉTICA EM EXPERIMENTAÇÃO ANIMAL e pela CEPsh - COMISSÃO DE ÉTICA EM PESQUISA COM SERES HUMANOS, tendo sido deliberado que o referido projeto não envolve manipulação animal ou humana que justifique uma aprovação quanto aos princípios éticos exigidos por ambas as Comissões.

São Paulo, 09 de dezembro de 2008.

PROF. DR. WOTHAN TAVARES DE LIMA
Coordenador da CEEA - ICB/USP

PROF. DR. LUIZ VICENTE RIZZO
Coordenador da CEPsh - ICB/USP

Dedico,
ao Senhor *Jesus Cristo* e à minha família.

AGRADECIMENTOS

Sem dúvida, quem me conhece sabe que esse tópico é e foi de suma importância para a realização deste trabalho. A amizade e o companheirismo me fizeram ser cada vez mais otimista em relação aos resultados que passo a passo foram conquistados.

A Deus, que permitiu a realização deste trabalho, sabedoria e fiel consolo;

Ao meu pai Lauro, minha mãe Espedita e irmãos André, Márcia e Fernando por manterem sempre seguras as minhas metas;

Ao Dr. Gilson Manfio pela formação profissional e orientação para realização deste trabalho;

Aos meus amigos do laboratório, alegria, apoio, colaboração e companheirismo. Karen, Milena, Rebeca, Cláudia e Suski Maluki às minhas fiéis escudeiras, o meu muito obrigado;

A Dra. Valéria Maia de Oliveira e Dra. Lara Durães Sette, apoio e confiança;

A Dra. Fabiana Fantinatti, colaboração, instrução e amizade;

Ao colaborador e amigo Maurício Drimitov (UNESP - Jaboticabal), parceria, caráter e competência;

Aos colaboradores Anita Marsaioli e Simone Mantovani (IQ-UNICAMP), Dra. Lucia Durrant e sua equipe (FEA), Dra. Ângela Metha dos Reis (Cenagen/Embrapa), Dr. Edmilson Gonçalves (PUCC), Dra. Yoko Bomura Rosato,

Aos membros da Banca de Qualificação Dra. Elizabete Vicente (ICB-USP), Dr. Wellington Luiz de Araújo (Esalq), Dra. Suzan Pantarotto Vasconcellos (CPQBA);

Aos amigos do CPQBA, orgulho e satisfação;

Aos meus amigos de longa data, a graduação não é o limite.

A grandiosa Dani, paciência, carinho, partilha, apoio, atenção, paz..."Presente sempre presente."

A quem que porventura não foi citado. Se você fez parte de algum momento desta construção, sinta-se incluso. Certamente em meus insights diários me recordo positivamente da sua pessoa.

*“Aquele que recebe um benefício não deve jamais esquecê-lo;
aquele que o concede, não deve jamais lembrá-lo.”*

Pierre Charron

A Arte de ser Feliz

“Houve um tempo em que minha janela se abria sobre uma cidade que parecia ser feita de giz. Perto da janela havia um pequeno jardim quase seco. Era uma época de estiagem, de terra esfarelada, e o jardim parecia morto. Mas todas as manhãs vinha um pobre com um balde, e, em silêncio, ia atirando com a mão umas gotas de água sobre as plantas. Não era uma rega: era uma espécie de aspersão ritual, para que o jardim não morresse. E eu olhava para as plantas, para o homem, para as gotas de água que caíam de seus dedos magros e meu coração ficava completamente feliz. Às vezes abro a janela e encontro o jasmineiro em flor. Outras vezes encontro nuvens espessas. Avisto crianças que vão para a escola. Pardais que pulam pelo muro. Gatos que abrem e fecham os olhos, sonhando com pardais. Borboletas brancas, duas a duas, como refletidas no espelho do ar. Marimbondos que sempre me parecem personagens de Lope de Vega. Às vezes, um galo canta. Às vezes, um avião passa. Tudo está certo, no seu lugar, cumprindo o seu destino. E eu me sinto completamente feliz. Mas, quando falo dessas pequenas felicidades certas, que estão diante de cada janela, uns dizem que essas coisas não existem, outros que só existem diante das minhas janelas, e outros, finalmente, que é preciso aprender a olhar, para poder vê-las assim.”

Cecília Meireles

RESUMO

SIMÕES, T.H.N., **Triagem de Enzimas Associadas à Biotransformação de Hidrocarbonetos a Partir de Metagenoma de Sedimentos Contaminados com Petróleo e Metais Pesados**, 2009. 172 f. Tese (Doutorado) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2009.

A metagenômica trouxe novas perspectivas ao estudo de comunidades microbianas no ambiente, permitindo explorar tanto a diversidade taxonômica de microrganismos ainda não-cultivados, como o acesso direto a genes e vias metabólicas. Neste trabalho, foram construídas bibliotecas metagenômicas a partir de amostras de sedimentos de mangue da Baía de Guanabara (RJ), impactadas com hidrocarbonetos de petróleo e metais pesados. Proteobacteria (33,3%), bactérias afiliadas a redutoras-de-sulfato (29,7%) e Firmicutes (20%) representaram os grupos principais nas amostras ambientais, baseado em análises filogenéticas de rDNA 16S, ao passo que isolamentos seletivos utilizando diesel e naftaleno permitiram a recuperação preferencial de delta-Proteobacteria e actinomicetos. Bibliotecas metagenômicas dos sedimentos enriquecidos com óleo diesel, com insertos entre 25 e 35 Kb clonados em foscídeos, foram triadas para detecção de genes catabólicos de monoxigenases (*alkB1*) e expressão de epóxido-hidrolases, esterases, lipases e monoxigenases em ensaios de alto desempenho (*HTS, high throughput screening*). Clones reativos a *alkB1* foram detectados, porém não foram funcionais nas condições de HTS testadas. Nas bibliotecas de foscídeos triadas, vários clones apresentaram atividade enzimática, sendo que dois apresentaram atividade de lipase-esterase com alta seletividade, elevada taxa de conversão de substratos e excesso enantiomérico (*ee* >99%). Os resultados de HTS comprovaram a eficiência do uso da clonagem direta de DNA ambiental na expressão de vias metabólicas de interesse com potencial de aplicação biotecnológica.

Palavras-chave: Bibliotecas metagenômicas; *alkB1*; rDNA 16S; HTS; Epóxido-hidrolase; Esterase; Lipase; Monoxigenase.

ABSTRACT

SIMÕES, T.H.N., **Screening of Enzymes Related to Biotransformation of Hydrocarbons from Metagenome of Contaminated Sediments with Oil and Heavy Metals**, 2009. 172 p. Thesis (Ph. D.) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2009.

Metagenomics brought a new perspective to the study of microbial communities in the environment, enabling access to the taxonomic diversity of uncultured microorganisms, as well as direct access to genes and metabolic pathways. In the current study, metagenomic libraries were constructed from mangrove sediment samples of the Guanabara Bay (RJ, Brazil), impacted with oil hydrocarbons and heavy metals. Proteobacteria (33.3%), sulfate-reducing affiliated bacteria (29.7%) and Firmicutes (20%) represented the main groups in the environmental samples based upon 16S rDNA phylogenetic analysis, whereas selective isolation using diesel and naphtalene yielded delta-Proteobacteria and actinomycetes. Metagenomic libraries of diesel-enriched sediment samples, with 25 to 35 Kb fosmid inserts, were screened for detecting monooxygenase genes (*alkB1*) and expression of epoxide hydrolases, esterases, lipases and monooxygenases in *high throughput screening (HTS)* assays. Clones reactive to the *alkB1* probe were detected, but were not functional under the HTS conditions used. Several functional clones were detected in the clone library, and two showed lipase-esterase activity with high rates of substrate conversion and enantiomeric ratio (ee >99%). The results obtained on HTS showed the efficiency of the direct cloning of environmental DNA for the expression of metabolic pathways with potential biotechnological application.

Key words: Metagenomic libraries; *alkB1*, 16S rDNA; High throughput screening; Epoxide-hydrolase; Esterase; Lipase; Monooxygenase.

LISTA DE FIGURAS

Figura 1. Esquema de degradação de alcanos em bactérias Gram-negativas, mostrando a localização e a função de produtos dos genes alk.....	32
Figura 2. Esquema de catabolismo aeróbio de compostos aromáticos.....	34
Figura 3. Organização dos genes tod em <i>Pseudomonas putida</i> DOT-T1.....	34
Figura 4. Construção de uma biblioteca metagenômica - fragmentos grandes de DNA obtidos de microrganismos não-cultivados de solo são ligados em vetores BAC.....	36
Figura 5. Esquema ilustrando estratégia de clonagem e expressão em plasmídios BAC - fragmentos de DNA abrangem operons inteiros, cuja expressão se dá em níveis baixos na célula hospedeira.....	37
Figura 6. Fenóis conjugados utilizados em substratos cromogênicos e fluorogênicos.....	41
Figura 7. Microplaca ilustrando a fluorescência da umbeliferona em reações positivas quanto à atividade enzimática (A); Ensaio em microplaca com fluorescência aparente (B).....	41
Figura 8. Ensaio fluorogênico para detecção de hidrolases.....	43
Figura 9. Modelo de conversão substrato produto para monoxigenase flavina-dependentes.....	46
Figura 10. Ensaio fluorogênico para detecção de monoxigenases (sondas 6a e 6b) e álcool desidrogenase (7).....	47
Figura 11. Mapa da área de estudo. Localização geográfica do estado do Rio de Janeiro (A); Localização da Baía de Guanabara (B); Localização geográfica da Ilha do Fundão (C); Representação do ponto de coleta (D) e ponto de coleta detalhe CENPS (Petrobrás) (E).....	50
Figura 12. Canal do Fundão (background) (A). Artigos para coleta: termômetro (B); trado (C); resíduos urbanos na região de coleta (D); sacos de coleta estéreis (E); alíquota da amostra de sedimento em placa de Petri (F).....	53

Figura 13. Imagem do experimento em microcosmo. Amostras controle, 3 e 10% de petróleo (A, B, C, respectivamente). Vista geral das jarras de ensaio (D), detalhe de jarro montado (E,F,G,H).....	56
Figura 14. Ensaios de pré-enriquecimento de amostras de sedimento de mangue provenientes de microcosmos impactado com petróleo.....	60
Figura 15. Observação com DAPI. Bactérias de enriquecimento aeróbio com 3% de óleo diesel. Material controle (A) e cultivados 24h (B); 72h (C) e 7dias (D). Aumento de 60X.....	62
Figura 16. Produção de Carbono (CO ₂) acumulado nos períodos de 15 dias (A); 30 dias (B); 45 dias (C) e 60 dias (D).Eixo X: Tempo em dias. Eixo Y: µg de C-CO ₂ 150 g ⁻¹ de amostra.....	66
Figura 17. Visualização de colônias isoladas em meio mineral salino (2% NaCl) e naftaleno (250 mg mL ⁻¹)......	69
Figura 18. Árvore filogenética mostrando a relação entre os isolados obtidos em meio mínimo + HPA (naftaleno) e linhagens de microrganismos relacionados com base na sequência de rDNA 16S.....	70
Figura 19. A distribuição dos clones e suas proporções entre os diversos grupos taxonômicos do Domínio Bactéria e os clones não afiliados (A). Proporção de clones no Filo Proteobacteria das amostras de sedimento de mangue (B).....	72
Figura 20. Análise filogenética do gene rDNA16S a partir de sedimento de mangue impactado com petróleo.....	74
Figura 21. Esquema resumo de endereçamento de clones em membrana, utilizada para a construção dos macroarranjos (4X4).....	81
Figura 22. Procedimento de detecção imunológica para certificar a qualidade de marcação e a diluição do DNA sonda adequada para os procedimentos posteriores de hibridização.....	85
Figura 23. Esquema de gel preparativo para seleção de fragmentos de alto peso molecular.	87

Figura 24 – Digestão enzimática de produtos de miniprep. Agarose 1%. (1) DNA λ – 100 ng; (2) DNA λ – 200ng; (3) DNA λ digerido EcoRI ; (4) DNA λ digerido Pst I; (5) Ladder 1 Kb; (6) DNA inserto controle; (7) vetor aberto pcc1FOS,1 μ g; clone C3 (SED3) digeridos BamHI, BglI e XhoI, respectivamente; (8,9,10).....	88
Figura 25. Membrana de macroarray da biblioteca metagenômica de enriquecimentos a partir de sedimentos impactados com petróleo. Círculos em vermelho destacam clones selecionados como hits positivos (portadores do gene alkB1).....	89
Figura 26. A - microplaca de 96 poços com 200 μ L de capacidade, B - leitor de fluorescência Flashscan 530 Analitic Jena com filtro de excitação ex=365 nm e leitura em=460 nm.	93
Figura 27. Resolução cinética dos substratos 17, 18 e 19.....	98

LISTAS DE TABELAS

Tabela 1. Mercado global da indústria de enzimas.....	39
Tabela 2. Teor total de metais (método: SW 3051 EPA-US).....	54
Tabela 3. Concentração de hidrocarbonetos poliaromáticos (HPA) das amostras coletadas Canal do Fundão (RJ).....	55
Tabela 4. Controles de pH no início do tratamento (t=0).....	56
Tabela 5. Meio Mineral 2% salino (Zinder - modificado).....	59
Tabela 6. Soluções-traço para Meio Mineral 2% salino (Zinder).....	59
Tabela 7. Reação de seqüenciamento para 16S rDNA.....	64
Tabela 8. Valores para determinação de atividade de enzimas em sedimentos de mangue.....	67
Tabela 9. Conjunto de primers utilizados para amplificação de genes de monoxigenases de n-alcanos.....	83
Tabela 10. Reação de PCR para preparação de sondas de genes catabólicos (<i>alk B1</i>).....	83
Tabela 11. Clones identificados preservados em microplacas 96 poços portadores de genes <i>alkB1</i>	89
Tabela 12. Otimização da concentração de células + sobrenadante nos ensaios HTS.....	95
Tabela 13. Resumo dos resultados obtidos dos ensaios de atividade enzimática em bactérias isoladas de enriquecimentos com óleo diesel.....	96
Tabela 14. Resultados obtidos pela triagem enzimática para hidrolases da biblioteca metagenômica construída a partir de sedimento utilizando as sondas fluorogênicas 1a, 1b, 1c, 2a e 2b.....	97
Tabela 15. Reação de hidrólise dos substratos 17 e 18 utilizando como biocatalisadores os hits positivos testados por HTS.....	98

SUMÁRIO

1 INTRODUÇÃO GERAL.....	18
2 OBJETIVOS.....	20
3 REVISÃO BIBLIOGRÁFICA.....	22
3.1 Comunidades microbianas em ambientes impactados com petróleo	22
3.2 Monitoramento do comportamento metabólico da comunidade microbiana de solos e sedimentos.....	26
3.2.1 Enriquecimento direcionado de amostras ambientais.....	28
3.2.2 Biodegradação e biotransformação microbiana de componentes do petróleo.....	29
3.2.2.1 Metabolismo de n-alcanos.....	31
3.2.2.2 Metabolismo de HPA's.....	32
3.3 Metagenômica e o potencial de exploração biotecnológica de comunidades microbianas não-cultivadas.....	35
3.4 Atividade biocatalítica.....	38
3.5 Triagens enzimática de alto desempenho (HTS).....	40
3.5.1 Hidrolases.....	42
3.5.2 Esterases e lipases.....	43
3.5.3 Epóxido Hidrolases.....	44
3.6 Monoxigenases.....	46
4. Caracterização da Região de Estudo - Canal do Fundão, Baía de Guanabara (RJ): Amostras de Sedimento de Mangue.....	48
CAPÍTULO 1- ESTUDO DA DIVERSIDADE MICROBIANA EM AMOSTRAS DE SEDIMENTOS IMPACTADAS POR HIDROCARBONETOS DE PETRÓLEO.....	51

1.1 Introdução.....	52
1.2 Material e Métodos.....	53
1.2.1 Coleta das amostras de sedimento de mangue.....	53
1.2.2 Montagem dos microcosmos e enriquecimento direcionado para degradadores de HTP e HPA's.....	55
1.2.3 Determinação da taxa de respiração microbiana.....	56
1.2.4 Determinação da atividade de enzimas microbianas em amostras de sedimentos.....	57
1.2.5 Enriquecimento direcionado para organismos degradadores de hidrocarbonetos derivados de petróleo.....	58
1.2.6 Isolamento e identificação de colônias a partir de enriquecimentos.....	60
1.2.7 Visualização por DAPI de amostras de enriquecimento aeróbio com óleo diesel - microscopia de epifluorescência.....	61
1.2.8 Caracterização da diversidade taxonômica nas amostras de sedimento de mangue da Baía de Guanabara.....	63
1.3 Resultados e Discussão.....	65
1.3.1 Quantificação C-CO₂ produzido em microcosmos.....	65
1.3.2 Determinação de enzimas de metabolismo microbiano em amostras de sedimento.....	66
1.3.3 Isolamento e identificação de bactérias a partir dos enriquecimentos.....	68
1.3.4 Caracterização da diversidade bacteriana de amostras de sedimento contaminado da Baía de Guanabara.....	71
CAPÍTULO 2- CONSTRUÇÃO E TRIAGEM GENÉTICA DE BIBLIOTECAS METAGENÔMICAS DE ALTO PESO CONSTRUÍDAS A PARTIR DE AMOSTRAS DE SEDIMENTOS IMPACTADAS POR HIDROCARBONETOS DE PETRÓLEO.....	75
2.1 Introdução.....	76

2.2 Material e Métodos.....	77
2.2.1 Extração de DNA de alto peso molecular.....	77
2.2.2 Procedimentos de clonagem de DNA metagenômico.....	77
2.2.3 Macroarrays - fixação das bibliotecas em membranas de náilon.....	80
2.2.3.1 Processamento das membranas para fixação de DNA.....	82
2.2.4 Construção das sondas de genes catabólicos.....	82
2.2.5 Screening genético para detecção de genes catabólicos em macroarranjos de bibliotecas metagenômicas de grandes vetores.....	84
2.3. Resultados e Discussão.....	87
2.3.1 Clonagem direta de DNA de Alto Peso Molecular.....	87
2.3.2 Validação dos clones da bibliotecas de alto peso molecular.....	88
CAPÍTULO 3- TRIAGEM FUNCIONAL DE BIBLIOTECAS METAGENÔMICAS DE ALTO PESO CONSTRUÍDAS A PARTIR DE AMOSTRAS DE SEDIMENTOS IMPACTADAS POR HIDROCARBONETOS DE PETRÓLEO	91
3.1 Introdução.....	92
3.2 Material e Métodos.....	92
3.2.1 Método HTS para bactérias isoladas a partir dos enriquecimentos direcionados.....	92
3.2.2 Triagem enzimática de bibliotecas metagenômicas.....	94
3.3 Resultados e Discussão.....	96
3.3.1 Triagem enzimática de biblioteca metagenômicas.....	97
5 DISCUSSÃO GERAL.....	100
RERERÊNCIAS BIBLIOGRAFICAS.....	103
ANEXOS.....	118

“A vida é o produto cumulativo de interações entre muitas espécies de substâncias químicas que constituem as células de um organismo”.

Neil A. Campbell

INTRODUÇÃO GERAL

Durante muitas décadas, o petróleo foi o grande propulsor da economia internacional, chegando a representar cerca de 50% do consumo mundial de energia primária no início dos anos 1970. Embora com uma redução gradativa ao longo dos anos, sua participação na matriz energética mundial é ainda da ordem de 40% (2004). Segundo o *World Energy Assessment*, a importância do petróleo na matriz energética mundial deverá manter-se expressiva por várias décadas.

As indústrias de extração e petroquímica desenvolvem atividades altamente impactantes para o meio ambiente, atingindo diversos ecossistemas no planeta, incluindo oceanos, regiões de estuários (geralmente portuárias), rios, águas subterrâneas e áreas continentais (solos). O aumento crescente de regiões impactadas por petróleo tem mobilizado a comunidade científica e iniciativa privada para o desenvolvimento de tecnologias para remoção de resíduos em áreas contaminadas. Com essa finalidade, diversos estudos na literatura se baseiam na aplicação de microrganismos nativos ou consórcios microbianos selecionados capazes de metabolizar (transformar ou mineralizar) petróleo e derivados presentes no ambiente.

Microrganismos adaptados ao metabolismo de hidrocarbonetos expressam enzimas capazes de degradar moléculas complexas e apresentam potencial para aplicações biotecnológicas em processos industriais, incluindo biotransformação, biodegradação de resíduos e biorremediação.

Estudos envolvendo a análise de comunidades microbianas geralmente empregam metodologias tradicionais dependentes-de-cultivo, baseadas no uso de condições seletivas, que permitem a recuperação de grupos microbianos específicos com baixa diversidade taxonômica e metabólica. Assim, esses

métodos são limitados para a avaliação da diversidade microbiana e de genes catabólicos em comunidades adaptadas em ambientes impactados.

O acesso à diversidade microbiana através do estudo de bibliotecas de rDNA 16S construídas a partir de DNA extraído diretamente de amostras ambientais se tornou um procedimento valioso para conhecer a real composição e distribuição de microrganismos nos ecossistemas. Dados assim obtidos são importantes para monitoramento ambiental, uma vez que variações ambientais relevantes são refletidas em alterações na composição e estrutura das comunidades. Porém, o método é limitado quanto à geração de informação a respeito da fisiologia de organismos não-cultivados.

A metagenômica é uma estratégia que permite a construção, triagem e expressão de fragmentos de DNA alto peso molecular isolado diretamente de comunidades microbianas em amostras ambientais. Estes fragmentos podem conter genes, operons e até mesmo vias metabólicas completas, permitindo acesso a características de microrganismos ainda não-cultivados ou de difícil isolamento por métodos tradicionais. Esta abordagem complementa a estratégia anterior de caracterização taxonômica da microbiota ambiental, agregando informações sobre a diversidade metabólica das comunidades. Além disso, a triagem de bibliotecas metagenômicas surge como uma nova alternativa para atender à demanda da indústria por novas enzimas com propriedades e aplicações diferenciadas.

A triagem funcional de bibliotecas metagenômicas de alto peso molecular de amostras de sedimento impactadas por hidrocarbonetos de petróleo, apresentada neste trabalho, representa uma abordagem inédita no Brasil. Este estudo consolida no país a viabilidade técnica do emprego de bibliotecas metagenômicas e triagem enzimática de alta eficiência (HTS) no estudo da diversidade microbiana.

OBJETIVOS

O estudo tem como objetivos principais:

- Caracterizar a diversidade de comunidades bacterianas presentes em sedimentos de mangue contaminados com hidrocarbonetos derivados de petróleo e metais pesados;
- Obter sequências gênicas de monoxigenases do tipo *alkB1* a partir de amostras de DNA metagenômico;
- Obter clones funcionais para atividades de esterases, lipases e monoxigenases a partir da biblioteca metagenômica.

Assim, o estudo “Triagem de Enzimas Associadas à Biotransformação de Hidrocarbonetos a Partir de Metagenoma de Sedimentos Contaminados com Petróleo e Metais Pesados” será apresentado em três partes, de modo a refletir os objetivos principais do trabalho, detalhadas a seguir.

- Capítulo 1. Estudo da diversidade microbiana em amostras de sedimento impactadas por hidrocarbonetos de petróleo;
- Capítulo 2. Construção e triagem genética de bibliotecas metagenômicas de alto peso construídas a partir de amostras de sedimento impactadas por hidrocarbonetos de petróleo;
- Capítulo 3: Triagem funcional de bibliotecas metagenômicas de alto peso construídas a partir de amostras de sedimento impactadas por hidrocarbonetos de petróleo para esterases, lipases e monoxigenases.

Capítulo 1

Capítulo 2

Capítulo 3

REVISÃO BIBLIOGRÁFICA

Comunidades microbianas em ambientes impactados com petróleo

No início da década de 50, Claude Zobell inicia seus estudos sobre a ação de microrganismos em hidrocarbonetos de petróleo. Em 1969, este autor já havia descrito mais de 100 espécies de microrganismos, incluindo 30 gêneros bacterianos distintos capazes de utilizar hidrocarbonetos durante o ciclo de vida (ATLAS, 1981). Esses estudos estruturaram as bases metodológicas para estudos posteriores em microbiologia do petróleo, biofilmes e amostras impactadas com óleo.

Embora os hidrocarbonetos estejam amplamente distribuídos no meio ambiente, a capacidade de degradação dessas moléculas por animais e plantas é limitada. No entanto, bactéria, fungos filamentosos e leveduras possuem recursos enzimáticos adaptados, com alta capacidade para degradar e metabolizar esses compostos (SPORMANN e WIDDEL, 2000; BACCHOON *et al.*, 2001). A capacidade adaptativa destes grupos de bactérias possibilita a sua sobrevivência em ambientes de derramamento, ou até mesmo em depósitos de óleo a grandes profundidades, sob condições ambientais extremas, altas temperaturas, pressão elevada e baixa tensão de oxigênio (FISHER, 1987; BARTH, 1991; BERNARD *et al.*, 1992).

Estudos recentes catalogam mais de 50 gêneros de bactérias capazes de biotransformar petróleo e derivados, tanto de habitats marinhos, como terrestres. Exemplos de bactérias degradadoras de alcanos e compostos aromáticos, tal como: tolueno, naftaleno e fenantreno, incluem Actinobacteria, Proteobacteria, Chloroflexi, Acidobacteria, dos gêneros *Achromobacter*, *Acinetobacter*, *Alcaligenes*, *Arthrobacter*, *Bacillus*, *Flavobacterium*, *Nocardia* e *Pseudomonas* spp. (HARAYAMA *et al.*, 2004; MELO *et al.*, 2001).

Diversos estudos na literatura apontam o uso viável de microrganismos autóctones como agentes despoluidores de regiões impactadas (RADWAN *et al.*, 1995; SO e YUNG, 1999; NAIR *et al.*, 2008). Em isolamentos microbianos de sedimentos impactados, Rapp e Gabriel-Jürgens (2003) verificaram que uma linhagem de *Rhodococcus* sp. degradava di-, tri- e tetraclorobenzenos e produzia biossurfactantes durante o crescimento em uma grande variedade de alcanos. Observaram que esta linhagem degradava *n*-alcanos com cadeia de C₁₀ – C₁₇ tão prontamente, ou até melhor que aqueles de cadeias menores, mostrando potencial para a biorremediação de solos e sedimentos contaminados por compostos recalcitrantes, uma vez que cadeias longas de hidrocarbonetos são mais dificilmente degradadas na natureza (CURY, 2006).

Microrganismos anaeróbicos também estão presentes em áreas contaminadas e em reservatórios de óleo. Representantes importantes deste grupo incluem as arqueas metanogênicas, bactérias fermentativas e bactérias redutoras de ferro e manganês (MAGOT *et al.*, 2000).

Segundo Magot e colaboradores (2000), na maioria dos ambientes anaeróbicos é necessária uma cooperação metabólica entre os microrganismos para que ocorra uma degradação completa da matéria orgânica. Através de técnicas moleculares baseadas em clonagem de genes RNAr 16S e hibridização *in situ*, foram detectadas populações de microrganismos acetogênicos afiliados ao grupo das Delta-Proteobacteria, associadas a populações de organismos metanogênicos relacionados a *Methanocalculus* e *Methanosaeta*. Neste sentido, a presença de ácidos graxos, como estearato e palmitato em algumas amostras contendo óleo bruto pode ser correlacionada com o processo de biodegradação, uma vez que os ácidos graxos são intermediários importantes para a degradação de matéria orgânica complexa (CRESPIN, 2008).

A Ecologia Molecular Microbiana, baseada em análises independentes-de-cultivo, tem colaborado para estudos que envolvem a análise da estrutura de comunidades no ambiente. Esta abordagem já tem sido empregada na caracterização da distribuição, riqueza e abundância de grupos microbianos em ambientes impactados por petróleo e derivados no Brasil (PIZA, 2004).

Em outros estudos recentes de caracterização de comunidades de bactérias associadas ao petróleo utilizando métodos moleculares independentes-de-cultivo, foram encontrados dez diferentes grupos taxonômicos em amostras de óleo provenientes de poços de produção com diferentes graus de biodegradação, desde não-biodegradado, até altamente biodegradado (OLIVEIRA *et al.*, 2008). Estes microrganismos mostraram-se relacionados a *Acinetobacter*, *Alicyclobacillus acidoterrestris*, *Arcobacter*, *Bacillus*, *Halanaerobium*, *Leuconostoc*, *Marinobacter*, *Streptomyces*, *Propionibacterium* e *Streptococcus*, pertencendo a quatro classes dentro do Domínio Bacteria: Firmicutes (Gram-positivas de baixo conteúdo de GC); Actinobacteria (Gram-positivas de alto conteúdo de GC); Gama-Proteobacteria e Epsilon-Proteobacteria. Foi verificado que alguns destes grupos de microrganismos eram exclusivos da amostra de óleo degradada, como os gêneros *Acinetobacter*, *Bacillus* e *Streptococcus* (OLIVEIRA *et al.*, 2008). Destes, membros dos gêneros *Acinetobacter* e *Bacillus* têm sido associados com petróleo e degradação de hidrocarbonetos em muitos outros estudos (NAZINA *et al.*, 2002; SAADOUN, 2002; BACH *et al.*, 2003; LU *et al.*, 2003).

Outros gêneros bacterianos relativamente comuns foram também descritos em estudos de biodegradação, incluindo *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus subtilis*, *Dietzia maris*, *Halomonas*, *Marinobacter*, *Microbacterium testaceum* e *Micrococcus* (CRESPIN, 2008; HUQ *et al.*, 1990, MOREIRA, SIQUEIRA e BRUSSARD, 2008; VAN HAMME, SINGH e WARD, 2003).

Neste sentido, podemos especular que as vias metabólicas importantes para a degradação de hidrocarbonetos de petróleo encontram-se amplamente dispersas em diferentes espécies microbianas, com características de metabolismo distintas. Assim, frente aos relatos da literatura, espera-se encontrar em amostras de sedimentos de mangue impactados por petróleo a presença de grupos especializados na biodegradação de hidrocarbonetos relacionados a diferentes táxons, ocupando, possivelmente, diferentes nichos nas amostras.

Em ambientes marinhos, os manguezais constituem um ecossistema de papel fundamental na manutenção da vida marinha em regiões tropicais e subtropicais (CURY, 2002). As suas características de solo, incluindo elevada concentração de matéria orgânica e baixa tensão de oxigênio, são favoráveis à acumulação de poluentes derivados de atividades antropogênicas, como, por exemplo, derivados do processo de refino de petróleo e pesticidas agrícolas. Esse acúmulo, em longo prazo, e em determinadas concentrações de poluentes, comprometem processos essenciais para a manutenção dos ecossistemas (CURY, 2002).

A contaminação de manguezais por derramamentos acidentais de petróleo e trânsito de embarcações cargueiras é frequente em todo o mundo (JACKSON *et al.*, 1989). Naturalmente, comunidades microbianas presentes no solo são diretamente impactadas pela presença desses poluentes, e se respondem de modo a se adaptarem ao tipo e à concentração do poluente presente (LYNCH e HOBBIE, 1990; ZANARDI, 1996).

A caracterização das comunidades em solos e sedimentos de mangue se ampliou com o uso de ferramentas de biologia molecular, tais como PCR em tempo real, DGGE, hibridização DNA:DNA e bibliotecas de rDNA16S, que permitem avaliar a diversidade e estrutura das comunidades resistentes a óleo (SEBAT *et al.*, 2003.). Pela aplicação dessas novas técnicas moleculares, além do acesso aos grupos ainda não-cultivados, é facilitado também o entendimento das relações entre eles nas comunidades (HUQ *et al.*, 1990; MOREIRA, SIQUEIRA e BRUSSARD, 2008). Estudos recentes tem focado também a caracterização de genes envolvidos no metabolismo de hidrocarbonetos derivados de petróleo em isolados de amostras impactadas (TIEDJE *et al.*, 1994, PELLIZARI *et al.*, 1996).

Segundo Van Hamme e colaboradores (2003), dados obtidos na caracterização de comunidades associadas ao metabolismo de hidrocarbonetos de petróleo podem:

- gerar conhecimento sobre as comunidades degradadoras de petróleo em amostras ambientais;

- fornecer parâmetros para o monitoramento de processos de biorremediação em derramamento acidentais de petróleo;
- auxiliar na definição de estratégias de recuperação de ambientes impactados, permitindo estimular o crescimento de comunidades degradadoras e intensificar taxas de degradação *in situ*;
- permitir a comparação com comunidades desenvolvidas em biorreatores e sistemas de tratamento de efluentes da indústria petroquímica.

Neste sentido, o estudo de amostras de sedimentos de mangue impactados com petróleo, utilizando uma abordagem metagenômica, pode nos fornecer subsídios importantes tanto para o conhecimento da composição, como da estrutura da comunidade degradadora, integrando informações importantes para a definição de estratégias de seleção e triagem de isolados com potencial para aplicação industrial.

Monitoramento do comportamento metabólico da comunidade microbiana de solos e sedimentos

Como anteriormente expresso, comunidades microbianas em solos e sedimentos geralmente apresentam alta diversidade e abundância de microrganismos (BOHANNAN e HUGHES, 2003; LAMBAIS *et al.*, 2005; MOREIRA, SIQUEIRA e BRUSSARD, 2008). As diferentes comunidades presentes no ambiente desenvolvem papéis fundamentais para o balanço material dos ecossistemas, através da sua participação nos ciclos biogeoquímicos (MELO, 1997; STENBERG, 1999).

Diversos métodos podem ser utilizados para acessar informações relativas ao metabolismo de microrganismos presentes em amostras ambientais. O isolamento de microrganismos em meios microbiológicos seletivos, apesar de limitado, ainda é o método mais utilizado para determinação do potencial metabólico de comunidades ambientais (VAN HAMME, SINGH E WARD, 2003). Além do isolamento, métodos indiretos de avaliação da atividade microbiana podem ser utilizados em paralelo. A respiração microbiana é um dos mais antigos parâmetros para quantificar a atividade microbiana em solos (MOREIRA e SIQUEIRA, 2002). A respiração microbiana em solo consiste na produção de CO₂, ou consumo de O₂, como resultado dos processos metabólicos de todos os organismos vivos presentes no solo. Ela representa a oxidação da matéria orgânica por microrganismos

aeróbios do solo, que, portanto, utilizam O_2 comoceptor final de elétrons. Parkinson e Coleman (1991) consideram que a respiração microbiana basal (RMB) é capaz de fornecer resultados válidos sobre a atividade microbiana do solo, tal qual, que em alguns casos, é utilizada como um índice de fertilidade para estudos agronômicos.

Assim como a respiração microbiana, a determinação de enzimas de metabolismo microbiano em solos evidenciam transformações bioquímicas da matéria orgânica para o aproveitamento energético dessas moléculas como recursos metabólicos (MAEGESIN *et al.*, 2002).

De acordo com Dick e colaboradores (1999), as avaliações da atividade enzimática nos solos pode ser útil para indicar em que medida um solo está desempenhando seu potencial de ciclagem de nutrientes, nitrificação, oxidação e outros processos.

As enzimas presentes nos solos desempenham, portanto, papel fundamental neste meio, pois, através da coordenação dos ciclos metabólicos nele presentes, asseguram a transformação e o movimento de materiais entre suas porções bióticas e abióticas (PIZA *et al.*, 2004). Como a extração de uma enzima do sedimento é extremamente difícil, sua presença no mesmo costuma ser caracterizada pela avaliação de sua atividade sob um conjunto de condições previamente estipuladas (pH, temperatura, umidade, concentração de substrato). Desta forma, as atividades enzimáticas de solo costumam medir sua atividade potencial e não sua atividade *in situ* (DICK *et al.*, 1999). Conforme Moreira e Siqueira (2002), várias atividades enzimáticas de solo podem ser medidas, sendo que os produtos de suas reações podem ser analisados por diversos métodos dependendo da sua natureza química.

Outro modo de monitorar e avaliar a biomassa ativa presente em amostras ambientais é através do uso de métodos visuais de microscopia como, por exemplo, a microscopia de contraste, de epifluorescência e eletrônica (KEPNER e PRATT, 1994). A microscopia de contraste é recomendada para acompanhar o desenvolvimento da biomassa em tempo real, numa superfície transparente. A microscopia de epifluorescência (EPF) é

uma alternativa excelente na quantificação de células aderidas às superfícies. Para visualizar a adesão bacteriana, usam-se substâncias fluorescentes como o alaranjado de acridina para coloração direta das células, ou anticorpos fluorescentes que se ligam às células, permitindo sua observação (COSTA, 1999).

O 4'-6'-diamidino-2-fenilidol (DAPI) é um fluorocromo que se liga ao DNA de células vivas é utilizado na determinação do conteúdo nuclear de células eucarióticas e procarióticas, principalmente em ensaios que visam análise de ciclo celular (KAPUSCINSKI, 1995). Desde a década de 80 o DAPI vem sendo utilizado para visualização e contagem de bactérias também em amostras ambientais, sendo inicialmente considerado o método mais eficiente para estimar o número mais provável (NMP) de bactérias contidas na amostra (KEPNER e PRATT, 1994). Atualmente, em conjunto com outras técnicas como FISH (*fluorescence in situ hybridization*), por exemplo, se torna uma ferramenta imprescindível nas determinações quantitativas e qualitativas de comunidades microbianas em amostras ambientais (LOY *et al.*, 2005).

Apesar do grande número de publicações com protocolos bem definidos para o uso de DAPI como forma de visualização e contagem, cada amostra ambiental estudada representa um desafio e demanda ajustes e otimização de metodologia para que seja possível a aplicação da técnica na geração de dados para estudos ambientais.

Enriquecimento direcionado de amostras ambientais

Existem diversas referências na literatura relativas à biomassa de amostras ambientais. Alguns trabalhos apresentam uma média entre $1,4 \times 10^7$ a $1,4 \times 10^9$ bactérias por grama de sedimento (WHEELER e STAHL, 1996). Porém, quando se trata de sedimentos impactados, sabe-se que esses números se reduzem de acordo com a complexidade e degradabilidade dos compostos presentes, podendo chegar até o máximo de 10^4 células por grama (MALTSEV *et al.*, 2005). Além disso, amostras impactadas exigem um tratamento diferenciado para que se acesse toda a biomassa, dadas as características físico-químicas dos poluentes presentes, como a viscosidade do

óleo, que dificultam a suspensão celular e até impossibilitam a extração de ácidos nucleicos.

O enriquecimento de amostras ambientais vem sendo muito empregado por diversos grupos de pesquisa, à mediada que a matriz de extração de DNA seja de difícil acesso. Diversos estudos na literatura apresentam resultados a partir do *screening* de bibliotecas metagenômicas com fragmentos de alto peso molecular. Porém, na maioria dos trabalhos apresentam resultados a partir de matrizes de fácil acesso à biomassa (LORENZ e ECK, 2005), ou ainda abordam o uso de pré-enriquecimento de solos e sedimentos.

Para este estudo a estratégia de montagem de microcosmos foi utilizada como ferramenta para estimular o crescimento de organismos degradadores de óleo.

Biodegradação e biotransformação microbiana de componentes do petróleo

O petróleo é uma mistura complexa de compostos orgânicos e com alto conteúdo de energia (MELO, 1997). Além de atuar como matriz energética em diversos países, o petróleo é um recurso natural que participa ativamente em diversos setores industriais na geração de subprodutos como: insumos agrícolas, alimentos, fármacos, entre outros produtos com amplo interesse em processos biotecnológicos.

As características físico-químicas do petróleo variam de acordo com a mistura de compostos orgânicos que o compõe. Os hidrocarbonetos que caracterizam o óleo podem ser agrupados em parafinas voláteis (*n*-alcanos ramificados ou não, até C₁₀), parafinas não voláteis (*n*-alcanos lineares e ramificados de C₁₀ a C₄₀), naftenos (cicloalcanos), alcenos (oleofinas) e aromáticos (monoaromáticos e poliaromáticos). Em um grupo à parte estão os compostos das resinas e asfaltenos.

Hidrocarbonetos derivados de petróleo são fontes de carbono consumíveis por determinados microorganismos, principalmente bactérias e fungos (VAN HAMME, SINGH e WARD, 2003).

O interesse em estudar comunidades microbianas resistentes a condições extremas de vida vem se intensificado de modo prospectivo para aplicação de recursos naturais, como enzimas termo-estáveis, agentes antimicrobianos competidores no desenvolvimento de simbiose microambiental em processos industriais, *etc.*

A indústria petroleira desenvolve duas rotas principais nos estudos de agentes microbianos. A primeira rota é focada em microbiota de potencial biodegradador de óleos, o que acarreta uma diminuição no potencial de refino pelo decaimento energético do litro produzido e biocorrosão de equipamentos por formação de biofilmes resistentes e duradouros. A segunda rota está envolvida com a capacidade de biotransformação dos hidrocarbonetos de petróleo para aplicações em processos biotecnológicos

A ação de biodegradação e biotransformação microbiana deriva de uma habilidade de conversão para quase todos os substratos orgânicos aquo-solúveis, incluindo compostos de solubilidade muito baixa, como hidrocarbonetos e esteróides, através de complexas reações catalisadas por enzimas, geralmente por eles produzidas (LIMA *et al.*, 2001; KLIBANOV, 2001; REYMOND, 2006). A biodegradação e a biotransformação microbiana é um fato observado pela humanidade há milhares de anos com procedimentos que variavam, desde cuidados com a preservação de alimentos durante longos períodos, a processos fermentativos para fabricação de bebidas e alimentos (WACKETT e HERSHBERGER, 2001).

O metabolismo dos hidrocarbonetos derivados de petróleo pode ser dividido em: metabolismo de fração saturada (*n*-alcanos, alcanos ramificados e cicloalcanos) e metabolismo da fração aromática (monoaromáticos e poliaromáticos (MELO *et al.*, 2001). A degradação efetiva de compostos químicos na natureza só é possível pela presença de consórcios microbianos adaptados às condições ambientais.

Metabolismo de *n*-alcanos

Vias degradativas de alcanos são bastante estudadas (VAN BEILEN *et al.*, 2002). Os alcanos são hidrocarbonetos saturados lineares ou ramificados

que perfazem grande parte da composição do petróleo (MELO *et al.*, 2001), de onde é retirada a fração correspondente à gasolina. Nesta classe encontram-se, ainda o metano (CH₄), etano (C₂H₆), propano (C₃H₈) e butano (C₄H₁₀) (CHAPMAN, 1983). Existem vários sistemas enzimáticos que geram o processo degradativo desses compostos.

Em ambiente aeróbio as oxigenases iniciam o processo de degradação adicionando um ou dois átomos de oxigênio aos seus substratos, tornando-os mais suscetíveis ao metabolismo por um grande número de microrganismos. As vias de degradação aeróbia dos *n*-alcanos geralmente iniciam-se com a oxidação do alcano à álcool primário, seguido por um aldeído e um ácido monocarboxílico. O ácido carboxílico é degradado via beta-oxidação onde há a formação de ácidos graxos com dois carbonos a menos e a formação de acetil-coenzima A, com liberação de CO₂ (MELO e AZEVEDO, 1997). Em leveduras, as enzimas citocromo P-450 realizam o passo inicial de monoxigenação dos *n*-alcanos, um fato já descrito para alguns isolados bacterianos como *Acinetobacter calcoaceticus* EB-104 (LAMB *et al.*, 2000).

Uma grande variedade genética de monoxigenases bacterianas é descrita na literatura mostrando-se, muitas delas, conservadas em diferentes gêneros bacterianos como é o caso da monoxigenase (AlkB) de *Pseudomonas putida* GPo1 (VAN BAILEN, 2001).

A via de degradação dos alcanos consiste em um sistema de hidroxilases formado basicamente por três componentes enzimáticos (Figura 1): hidroxilase de alcano (AlkB), rubredoxina (AlkG) e rubredoxina redutase (AlkT). A AlkB é uma proteína de membrana que potencializa reação de hidroxilação. Essa proteína é codificada pelo gene *alkB*, com aproximadamente 1,2 Kb, o qual tem sido detectado em uma larga fração de microrganismos de ambientes contaminados por óleo, sendo marcante no gênero *Pseudomonas* (SMITS *et al.*, 2002). Por esse motivo, podemos inferir que o gene *alkB* é um potencial marcador molecular de contaminação ou processos biodegradativos de petróleo, podendo ser utilizado em estudos de prospecção em amostras de petróleo bruto.

Com isso, os HPA's induzem uma repressão co-metabólica, fazendo com que ocorra interrupção no processo de degradação dos próprios HPA's e de outros compostos tóxicos ou não integrados no cometabolismo (VAN HAMME, 2003).

Os HPAs são considerados compostos potencialmente perigosos, devido suas atividades cancerígena e mutagênica, segundo a Agência de Proteção Ambiental (EUA), que classificou 16 HPAs como poluentes prioritários: naftaleno, acenaftaleno, acenaftileno, fenatreno, fluoreno, fluoreteno, antraceno, benzo[a]antraceno, criseno, pireno, benzo[a]preno, benzo[a]fluoranteno, benzo[k]fluoranteno, dibenzeno[a,h]antraceno, benzo[ghi]perileno e indeno[1,2,3-cd]pireno.

A degradação aeróbia de hidrocarbonetos aromáticos (Figura 2) se inicia por uma diidroxilação do substrato por enzimas oxigenases formando dióis, seguida da clivagem do anel e formação de intermediários comuns (catecol, protocatecol e ginsinato), que posteriormente são degradados a intermediários do ciclo de Krebs (ALEXANDER, 1997; BELLINASSO, 2002).

Figura 2. Esquema de catabolismo aeróbio de compostos aromáticos.
(Fonte: BELLINASSO, 2002.)

A degradação aeróbia de compostos aromáticos é bem representada pelo tolueno que, certamente, possui a via mais estudada e aplicada à biodegradação (ZYLSTRA e GIBSON, 1991). Essa via se inicia pela adição de oxigênio na posição 2,3 do anel aromático por uma dioxigenase. Em seguida, ocorre uma desidrogenação, originando catecol e a clivagem do anel, resultando, como produto final, o ácido clorobenzóico (KANALY e HARAYAMA, 2000). Em *Pseudomonas putida* DOT-T1, Mosqueda e colaboradores (1999) descreveram a via metabólica do tolueno com 13 genes divididos em dois operons: *todXFC1C2BADEGIH* e *todST*, com uma sequência de nucleotídeos de aproximadamente 15 Kb (Figura 3).

Figura 3. Organização dos genes *tod* em *Pseudomonas putida* DOT-T1.
(MOSQUEDA *et al.*, 1999).

Outras vias relacionadas à degradação de compostos aromáticos foram relatadas em *Pseudomonas putida* BE81, (genes *bnz*, degradação de benzeno) e *Pseudomonas* sp. mt-2 (operon *xyl*, degradação de xileno), e em outros gêneros bacterianos (PRITCHARD, *et al.*, 1995; VAN HAMME *et al.*, 2003). Grandes partes dessas vias possuem *operons* com aproximadamente 20 Kb, possibilitando serem clonados e expressos em vetores de alta capacidade (RONDON *et al.*, 2000).

Metagenômica e o potencial de exploração biotecnológica de comunidades microbianas não-cultivadas

Os genomas da microbiota total encontrada em uma comunidade são chamados, coletivamente, de metagenoma (HANDELSMAN *et al.*, 1998; BÉJÀ *et al.* 2000, RONDON *et al.*, 2000) e contém maior informação genética do que aquela verificada na diversidade microbiana cultivada.

Os microrganismos não-cultivados representam a maior parte da diversidade biológica do planeta (RIESENFELD *et al.*, 2004). Na maioria dos ambientes, cerca de 99% da diversidade microbiana ainda não é conhecida. As limitações geradas por técnicas tradicionais de isolamento e enriquecimento inviabilizam o desenvolvimento de organismos fastidiosos e dependentes de co-metabolismos, impossibilitando o acesso à fisiologia de determinados grupos (LORENZ *et al.*, 2002). Após o uso de ferramentas moleculares no estudo de caracterização de comunidades microbianas baseada em rDNA 16S abriu-se um novo panorama de acesso à microrganismos novos para ciência, no entanto, muitos microrganismos não-cultivados, são afiliados a filos que contém membros cultivados. Ao passo que a metade dos filos microbianos conhecidos possui membros não-cultivados (RIESENFELD *et al.*, 2004).

Além do sequenciamento do gene rDNA16S para os estudos da diversidade de comunidades microbianas, técnicas baseadas em extração direta de ácidos nucleicos de amostras ambientais associada às técnicas de hibridização com sondas grupo-específicas e/ou de amplificação por PCR com *primers* grupo-específicos, clonagem e seqüenciamento, vêm permitindo uma avaliação mais precisa da diversidade microbiana no ambiente e a descoberta

de novos grupos de organismos, nunca antes cultivados (GIOVANONNI *et al.*, 1990; TORSVIK *et al.*, 1990; AMANN *et al.* 1995; PACE, 1996). Entretanto, estas metodologias dificultam o acesso ao potencial metabólico destes novos organismos, uma vez que as etapas de isolamento e cultivo são suprimidas dos estudos. Assim, uma estratégia alternativa é a clonagem de metagenomas na forma de grandes fragmentos de DNA, entre 40 a 100 Kb, através do uso de vetores de alta capacidade (Fosmídeos ou BAC - *Bacterial Artificial Chromosome*) (HANDELSMAN *et al.*, 1998; RONDON *et al.*, 1999a; RONDON *et al.*, 2000). Estes vetores têm a habilidade de manter de maneira estável fragmentos grandes de DNA exógeno em hospedeiros como *Escherichia coli*, e têm sido usados para a construção de bibliotecas genômicas de eucariotos (SHIZUYA *et al.*, 1992; KIM *et al.*, 1996). Desse modo, é possível clonar fragmentos que contenham operons ou vias metabólicas inteiras de organismos presentes no ambiente, cultiváveis ou não. Vetores tipo fosmídeos e semelhantes podem ainda expressar genes e/ou operons inteiros quando clonados e transformados em *E. coli* (Figura 4).

Figura 4. Construção de uma biblioteca metagenômica - fragmentos grandes de DNA obtidos de microrganismos não-cultivados de solo são ligados em vetores BAC.
(Fonte: Osburne *et al.*, 2000).

O uso de fosmídeos para clonagem de genomas bacterianos aumenta a probabilidade de se obter a expressão gênica do inserto a partir do promotor presente no vetor, uma vez que o hospedeiro é também um procarioto. Utilizando-se estes novos vetores é possível ainda obter um aumento de até 100 vezes no nível de expressão, facilitando significativamente a detecção dos clones com as atividades biológicas procuradas, como também as etapas de recuperação e purificação do DNA dos clones de interesse.

O uso dessa estratégia também permite combinar o escrutínio de genes funcionais nas bibliotecas metagenômicas com a análise de sequências de rDNA 16S, o que pode representar os primeiros passos no sentido de estabelecer relações entre taxonomia e função e ajudar a elucidar o papel ecológico dos componentes da microbiota no ambiente (OSBURN *et al.*, 2000). Bibliotecas de alto peso molecular oferecem, assim, uma ferramenta valiosa para a avaliação e exploração, de maneira mais abrangente, da diversidade total em um dado ambiente (Figura 5).

Figura 5. Esquema ilustrando estratégia de clonagem e expressão em plasmídios BAC - fragmentos de DNA abrangem operons inteiros, cuja expressão se dá em níveis baixos na célula hospedeira. (Fonte: RONDON *et al.*, 1999a).

Estudos envolvendo a aplicação de técnicas moleculares para detecção e caracterização de genes de microrganismos não-cultivados de comunidades em amostras de petróleo podem contribuir para o desenvolvimento de aplicações biotecnológicas diversas.

Através desses estudos é possível acessar vias metabólicas microbianas, contribuindo para a compreensão da fisiologia de microrganismos não-cultivados presentes na amostra de estudo, permitindo inferir a atividade de consórcios em ambientes de exploração, tal como rocha-reservatório, melhorando o entendimento do processo de biodeteriorização de óleo e aplicando as rotas acessadas em ensaios de biocatálise.

Atividade biocatalítica

Atividades biocatalíticas referem-se aos processos tecnológicos que usam catalisadores biológicos em etapas de conversão química. Estes catalisadores são em geral enzimas imobilizadas ou células íntegras de microrganismos (SANTOS, 2005).

A área multidisciplinar de biocatálise encontra-se atualmente em amplo desenvolvimento. Pesquisas realizadas em vários ramos da química e da biologia têm como principal objetivo o desenvolvimento de novos catalisadores para aplicação industrial (CONTI *et al.*, 2001). O interesse industrial por tecnologias enzimáticas vem aumentando continuamente, principalmente, nas áreas de engenharia de proteínas e enzimologia em meios não convencionais, as quais ampliaram consideravelmente a aplicação das enzimas como catalisadores em processos industriais (Tabela 1).

Entre os processos de maior interesse industrial e ambiental estão as reações de hidrólise, síntese e interesterificação de lipídeos por meio das lipases (Tabela 1).

Entretanto, avaliando-se a aplicabilidade dos estudos acadêmicos em biocatálise, depara-se com um cenário não incentivador no sentido da otimização dos processos desvendados, pois poucas rotas enzimáticas chegam à aplicação em escala industrial (SCHMID *et al.*, 2001). O motivo

deve-se a alguns problemas que não foram ainda inteiramente solucionados, como catálise lenta de substratos não-naturais, baixa estabilidade, pouca atividade em meios não aquosos e/ou requerimento de co-fatores de custo elevado (SCHMID *et al.*, 2001; SCHOEMAKER; MINK; WUBBOLTS, 2003) e (FARINAS; BULTER; ARNOLD, 2001).

Tabela 1. Mercado global da indústria de enzimas.

Setor	Enzimas	Milhões US\$
alimentos fermentados, cervejaria, edulcorantes, panificação, etanol de cereais	alfa-amilase, glucoamilase, glucose-isomerase	500
detergentes	protease, lipase, amilase	450
têxteis	amilases	150
tratamento de couro	enzimas diversas	25
texteis, produtos alimentícios	celulases	25
produtos lácteos	lactase	150
alimentício, cervejaria, panificação, terapêutico	papaína	25
suco de frutas, vinícola	pectinases, pectina-esterases	30
Total		1.355

(Fonte: CASTRO *et al.*, 2004).

Muitos dos processos biocatalíticos atuais envolvem resoluções racêmicas para a obtenção de intermediários enantiopuros, os quais são dificilmente obtidos utilizando catalisadores químicos, como é usualmente feito na síntese orgânica. Há diversas etapas que envolvem a geração de um novo produto ou processo em biocatálise (SCHMID *et al.*, 2001). Primeiramente, é preciso conhecer em detalhes sobre os mecanismos de conversão dos reagentes a produtos. A próxima etapa consiste da seleção do biocatalisador/biocatalista, a qual pode ser feita a partir de um grupo de biocatalisadores já disponíveis no comércio ou mediante o isolamento de uma nova enzima ou microrganismo. Esta etapa é denominada bioprospecção e, muitas vezes, representa os maiores investimentos para o estabelecimento de um novo processo (SANTOS, 2005). A terceira e última etapa consiste no isolamento e caracterização do biocatalisador, seguindo-se à produção em larga escala. (SANTOS, 2005).

O desenvolvimento de métodos que permitam o estudo e o acesso à biocatalisadores ainda desconhecidos são de extrema importância dentro do contexto de tecnologia enzimática. A biotecnologia molecular compõe um conjunto de técnicas que fornecem novas alternativas em enzimologia, tanto para o melhoramento de enzimas já disponíveis (DNA *shuffling*, ou evolução dirigida *in vitro*) que se refere a sua estereoespecificidade e termoestabilidade, quanto pela descoberta de enzimas nunca isoladas anteriormente através da clonagem direta de fragmentos de DNA metagenômico (GODDARD e REYMOND, 2004; GALVÃO, MOHN e LORENZO, 2005; HARDEMAN e SJÖLIN, 2007). Sendo assim, a biotecnologia molecular supre as necessidades e gera inovação tecnológica em processos convencionais para obtenção e aplicação de enzimas microbianas distribuídas na natureza (CONTI *et al.*, 2001; CASTRO *et al.*, 2004; PARALES *et al.*, 2002).

Triagens enzimáticas de alto desempenho (*High Throughput Screening* -HTS)

Os ensaios no formato HTS, consistem na utilização de um substrato capaz de liberar um cromóforo ou fluoróforo como consequência da ação enzimática. Dentre esses substratos, têm-se os derivados do tão conhecido ânion nitrofenolato (8), que apresenta cor amarela, e os ânions fluorescentes umbeliferona (9), resofurina (10) e a fluoresceína (11) que fluorescem no azul, vermelho e verde respectivamente (Figura 6). Muitos desses substratos são comercialmente disponíveis e amplamente utilizados para detecção de enzimas hidrolíticas.

Figura 6. Fenóis conjugados utilizados em substratos cromogênicos e fluorogênicos que na forma neutra são pouco coloridos e não-fluorescentes (Fonte: Mantovani, 2007).

Através da aplicação desta metodologia é possível detectar vários tipos de enzimas, com alta seletividade, incluindo álcool desidrogenases (KLEIN e REYMOND, 1998), aldolases (JOURDAIN *et al.*, 1998), lipases e esterases (KLEIN e REYMOND, 1999), acilases e fosfatases (BADALASSI *et al.*, 2000), proteases (BADALASSI *et al.*, 2002), transaldolases (GONZALEZ-GARCIA *et al.*, 2003), BAEYER-VILLIGERASES (Figura 7) (GUTIERREZ *et al.*, 2003) e epóxido hidrolases (BADALASSI *et al.*, 2004).

Figura 7. Microplaca ilustrando a fluorescência da umbeliferona em reações positivas quanto à atividade enzimática (A); Ensaio em microplaca com fluorescência aparente (B). (Fonte: VASCONCELOS, 2006).

A triagem enzimática de microrganismos isolados é realizada por muitos grupos de pesquisas resultando na descoberta de novas linhagens com atividades diferenciadas. Um grande número de microrganismos entre bactérias, fungos filamentosos e leveduras, são conhecidos e utilizados pela indústria para obtenção de compostos com aplicação em diversos setores como: produção de fármacos, alimentos embutidos, química fina, cosméticos, processamentos de resíduos em tratamentos de efluentes, entre outros (CASTRO *et al.*, 2001).

Com o mesmo objetivo, a triagem enzimática de alto desempenho (HTS) de bibliotecas metagenômicas surge como uma alternativa promissora para buscar novas rotas de produção de compostos alvos produzidos no ambiente. A abordagem metagenômica se torna interessante ao passo que possibilita a triagem das atividades enzimáticas de genomas de organismos nunca cultivados e/ou de difícil acesso.

O estudo das biotransformações de compostos orgânicos por microrganismos se acumulam na literatura resultando no conhecimento de um número amplo e crescente de reações para aplicações industriais (REYMOND, 2006). O aumento na demanda de novos biocatalisadores para a indústria química sugere a importância no uso de bibliotecas metagenômicas como fonte de recursos enzimáticos ainda não acessados (LORENZ *et al.*, 2002; REYMOND, 2006).

Hidrolases

As hidrolases são enzimas capazes de hidrolisar ou formar ésteres, amidas, lactonas, lactamas, epóxidos, nitrilas, anidridos, glicosídeos e organolídeos. As reações catalisadas por essa classe de enzimas são rotas utilizadas na resolução de racematos, pois são mais seletivas e baratas que os métodos químicos convencionais. Além disso, são bastante atrativas pelo fato de serem independentes de co-fatores e reagirem com uma grande variedade de substratos. As lipases e esterases são as mais amplamente utilizadas em

resoluções e dessimetrizações de álcoois primários, ácidos carboxílicos, e substratos complexos contendo um ou mais centros estereogênicos.

A técnica para detecção de hidrolases inicia-se pela hidrólise dos substratos fluorogênicos (1 e 2) pelo biocatalisador, gerando um produto (3), cuja oxidação por NaIO_4 leva ao composto ariloxi-carbonílico (4) precursor direto do composto fluorescente (5). Nesta última etapa, a albumina de soro bovino (BSA) promove uma beta-eliminação (Figura 8).

Figura 8. Ensaio fluorogênico para detecção de hidrolases.

Esterases e lipases

Esterases representam um grupo diversificado de hidrolases que catalisam a clivagem e a formação de ligações éster. O substrato padrão para esterases é a tributirina (tributirilglicerol), que, entretanto, podem também ser hidrolisada por lipases (JAEGER; DIJKSTRA; REETZ, 1999; THOMSON, DELAQUIS e MAZZA, 1999). As esterases e lipases estão amplamente distribuídas em animais, plantas e microrganismos (JAEGER e REETZ, 1998). Essas enzimas apresentam tolerância a uma grande variedade de substratos, além de mostrarem alta regio e estereo-especificidade, o que faz desses

biocatalisadores uma ferramenta atrativa para produção de compostos em indústrias de química fina. O interesse nas esterases reside também no fato de serem independentes de co-fatores, e geralmente muito estáveis e ativas em solvente orgânico (JAEGER; DIJKSTRA; REETZ, 1999). As esterases, junto das lipases, representam o grupo mais importante de biocatalisadores para aplicações biotecnológicas. As lipases caracterizam-se por catalisar reações de hidrólise, esterificação e transesterificação de triglicerídeos compostos principalmente por ácidos graxos de cadeia longa (C8 – C10), enquanto as “verdadeiras” esterases hidrolisam preferencialmente ésteres de cadeia curta, como acetatos e propionatos, triglicerídeos como ésteres de ácido graxo de cadeia menor que (C6) (JAEGER; DIJKSTRA; REETZ, 1999). A maior diferença entre lipases e esterases surge das interações físico-químicas com seus substratos. Enquanto as esterases obedecem à cinética clássica de *Michaelis-Menten*, ou seja, sua atividade é dependente da concentração do substrato, as lipases necessitam de uma concentração mínima de substrato antes de serem ativadas. Esse fenômeno é chamado de ativação interfacial, e é observado somente para lipases. O mecanismo de hidrólise de ésteres é essencialmente o mesmo para lipases e esterases.

Organismos bem descritos pela literatura como *Candida antarctica*, *Pseudomonas fluorescens*, *Candida rugosa*, *Chromobacterium viscosum*, *Burkholderia cepacia*, *Rhizopus oryzae*, são utilizados pela indústria farmacêutica, cosmética e química fina na produção dos mais variados compostos. Entre os produtos, estão os flavonídeos, agentes anti-tumorais, precursores quirais de vitaminas, biolubrificantes, biosurfactantes, poliéster, policarbonatos, biodiesel, *etc.* (SANTOS, 2005; THOMSON, DELAQUIS e MAZZA, 1999).

Epóxido Hidrolases

A biodegradação de alcanos e aromáticos envolve sistemas enzimáticos, os quais são responsáveis pela ação primária aos substratos. Monoxigenases e dioxigenases são as enzimas chaves na completa mineralização de moléculas alifáticas e monoaromáticas, ou de um dos anéis de um sistema poliaromático. Entretanto, hidrocarbonetos poliaromáticos são alvos de ataque

de monooxigenases microbianas, onde a atividade destas enzimas resulta na formação de um epóxido (Figura 8), o qual é um intermediário altamente reativo. Epóxidos, por sua vez, podem ser transformados em dióis, através da ação de epóxido hidrolases microbianas.

As epóxido-hidrolases estão envolvidas em várias rotas de degradação e na biossíntese de importantes compostos biológicos. Recentemente, epóxido hidrolases de microrganismos têm recebido considerável atenção nos processos que envolvem a biotransformação de vários compostos, promovendo reações estereosseletivas de grande importância. Dois aspectos importantes aumentaram os interesses por epóxido hidrolases, onde o primeiro é que tais enzimas não necessitam de cofatores e o segundo é que as epóxido hidrolases possuem um amplo leque de substratos, sendo que ambos, epóxidos e dióis, opticamente puros, podem ser utilizados como importantes blocos de construção em sínteses assimétricas (SWAVING e BONT, 1996).

Os primeiros exemplos de epóxido hidrolases microbianas foram descobertos nas bactérias dos gêneros *Flavobacterium* (MARTIN e FOSTER, 1955) e *Pseudomonas* (ALLEN e JAKOBY, 1969), e são responsáveis pela conversão de 2,3-epoxisuccinato em ácido tartárico. É importante ressaltar que succinato também é descrito como produto da biotransformação de compostos aromáticos após a atividade de monooxigenases microbianas.

A abertura do epóxido pela enzima óxidosqualeno ciclase, leva à formação de uma molécula policíclica, semelhante ao colesterol, o qual sabe-se ser o precursor natural do biomarcador colestano, presente no petróleo. Esqualeno epoxidase, por sua vez, catalisa a epoxidação de esqualeno a 2,3-óxidoesqualeno, o qual é substrato para a formação de esteróis e uma grande variedade de triterpenóides (VOLKMAN, 2005). Pode-se assim pressupor sobre a atividade de oxigenases e epóxido hidrolases, não somente na biodegradação de hidrocarbonetos, mas também na biossíntese de biomarcadores do petróleo.

Monoxigenases

As biotransformações oxidativas de hidrocarbonetos envolvem normalmente um complexo enzimático conhecido como oxigenases, as quais catalisam a inserção de um (monooxigenases) ou dois átomos de oxigênio (dioxigenases) em um ponto específico da molécula, freqüentemente com alta regio e estereosseletividade (MASON e COMMARK, 1992).

As monoxigenases são divididas em duas classes principais: as dependentes de flavinas (Figura 9) e as dependentes do citocromo P450. Nos últimos 10 anos, a oxidação enzimática de Baeyer-Villiger tem se mostrado como uma estratégia eficiente, onde estas flavoenzimas, denominadas como Baeyer-Villiger monooxigenases (BVMOs), catalisam a oxidação nucleofílica de cetonas, assim como a oxidação eletrofílica de vários heteroátomos, tais como enxôfre, selênio, nitrogênio e fósforo.

Figura 9. Modelo de conversão substrato produto para monooxigenase flavina-dependentes.

Oxidações de Baeyer-Villiger mediadas por enzimas constituem uma grande ferramenta da Química Verde na produção de lactonas quirais. Muitos organismos têm sido descritos como catalisadores destas *reações*, sendo que as monoxigenases microbianas vêm ganhando destaque na produção de vários intermediários quirais para a indústria (MIHOVILOVIC *et al.*, 2003; MIHOVILOVIC *et al.*, 2006).

Princípio semelhante pode ser aplicado para a detecção de monoxigenases (6) e alcooldesidrogenases (7) por fluorescência (Figura 10). As oxidações enzimáticas das sondas 6a, 6b e 7 levam aos produtos de oxidação (8a, 8b e 9) que na presença de BSA.

Figura 10. Ensaio fluorogênico para detecção de monooxigenases (sondas 6a e 6b) e álcool desidrogenase (7).

O uso de técnicas em triagem enzimática de alto desempenho (HTS) em bibliotecas metagenômicas tem despertado novos interesses dentro da comunidade científica em estudo de amostras ambientais (HARDEMAN e SJÖLIN, 2007). Através da clonagem direta em vetores de alta capacidade, como fosmídeos e BAC's (*Bacterial Artificial Chromosome*) foi possível acessar novas enzimas a partir de solos (HENNE *et al.*, 1999, 2000; RONDON *et al.*, 2000; LEE *et al.*, 2004b; YUN *et al.*, 2004; KIM *et al.*, 2006), água (COTTRELL *et al.*, 1999; RANJAN *et al.*, 2005) e ambientes termais (RHEE *et al.*, 2005). Há um grande interesse no estudo de amostras com elevada diversidade microbiana refletindo maior abundância de genes e vias metabólicas (HARDEMAN e SJÖLIN, 2007). Amostras de sedimentos marinhos são ricas em diversidade e abundância de grupos bacterianos e constituem um importante ecossistema de exploração para biotecnologia (RAVENSCHLAG *et al.*, 1999; LEE *et al.*, 2004a).

Dentro deste contexto as triagens enzimáticas de alto desempenho (HTS) são necessárias para o rápido e eficiente acesso de recursos enzimáticos potenciais em bibliotecas metagenômicas construídas a partir de biomas brasileiros. (VASCONCELLOS, 2006).

CARACTERIZAÇÃO DA REGIÃO DE ESTUDO - CANAL DO FUNDÃO, BAÍA DE GUANABARA (RJ): AMOSTRAS DE SEDIMENTO DE MANGUE

A Baía de Guanabara (RJ), segunda maior baía do litoral brasileiro, possui uma área de cerca de 380km², englobando praticamente toda a Região Metropolitana do Estado do Rio de Janeiro. Os atuais níveis de poluição da Baía de Guanabara são decorrentes de um processo de degradação que se intensificou, principalmente, nas décadas de 1950-1960, com o elevado crescimento urbano verificado, especialmente, na Região Sudeste do Brasil (FEEMA, 2007). Diariamente, são despejadas nas águas da Baía de Guanabara toneladas de esgoto e lixo doméstico, efluentes industriais e resíduos de óleo, entre outros poluentes, levando ao desfalecimento dos diferentes ecossistemas da Baía (FILHO *et al.*, 2003 ; BRITO *et al.*, 2006).

A contaminação por petróleo e seus derivados é uma das formas mais freqüentes de poluição da região da Baía de Guanabara face à diversidade de fontes poluidoras existentes. Há cerca de vinte anos, após detalhado levantamento realizado, estimou-se que são lançados na Baía de Guanabara uma quantidade de 9,5 toneladas dia de óleo e 300 kg de metais pesados (como o Pb, Cr, Hg e Zn) proveniente de duas refinarias de petróleo (REDUC e Manguinhos), dois portos comerciais (Rio de Janeiro e Niterói), 15 terminais de Petróleo, 40 estaleiros, além da contribuição de outras indústrias e de mais de 2000 postos de abastecimento. Esse cenário comprova que a Baía de Guanabara é um exemplo de região altamente afetada por atividades antropogênicas. Além das intensas atividades impactantes, em Janeiro de 2000 a região sofreu com um derramamento accidental de 1,3 milhões de litros de óleo causado por um rompimento de dutos de óleo da região nordeste da Baía (BRITO *et al.*, 2006). A literatura aponta diversos estudos relacionados ao impacto gerado pelo pólo industrial e pelo crescimento urbano desordenado nas margens da Baía de Guanabara. Diversos projetos de despoluição dessa região foram iniciados pelo governo do estado do Rio de Janeiro, porém sem avanços aparentes (CIBG, 2008).

O Canal do Fundão (Figura 10), escolhido como ponto de coleta em nossos estudos, é um dos efluentes que deságuam na Baía de Guanabara. Localizado entre a Ilha do Fundão e o continente, ao longo da Linha Vermelha, o Canal do Fundão é a região que apresenta maior grau de contaminação por diferentes fontes poluidoras na Baía de Guanabara (Resolução no. 07/2007). O local onde foram realizados os procedimentos de coleta ($22^{\circ} 51' 18,4''$ S; $43^{\circ} 14' 11,4''$ W) apresenta acúmulo progressivo de sedimentos finos, coesivos, e altamente contaminados, depositados ao longo do tempo, dentro da seção de escoamento do canal. Esse fato colabora para o intenso assoreamento e intensifica o impedimento da recirculação das marés por toda a extensão do lado sul da Ilha do Fundão (Resolução no. 07/2007). A região do Canal do Fundão, assim como grande parte dos afluentes que deságuam na Baía de Guanabara estão associados a manguezais (SEMADS, 2001). Dos 260 Km² originalmente cobertos por vegetação de mangue, atualmente restam aproximadamente 82 Km² (CIBG, 2008).

Apontamos a importância de estudos relacionados a estrutura da diversidade microbiana resistente ao óleo em áreas impactadas e/ou degradadas como o Canal do Fundão. Além do acesso a informações de cunho ambiental, dados obtidos na caracterização destas comunidades podem ser utilizados como parâmetros para a prospecção de genes e enzimas envolvidas na degradação de componentes do óleo com potencial de aplicações em biotecnologia.

Figura 11. Mapa da área de estudo. Localização geográfica do estado do Rio de Janeiro (A); Localização da Baía de Guanabara (B); Localização geográfica da Ilha do Fundão (C); Representação do ponto de coleta (D) e ponto de coleta detalhe CENPS (Petrobrás) (E).

CAPÍTULO 1 - ESTUDO DA DIVERSIDADE MICROBIANA EM AMOSTRAS DE SEDIMENTOS IMPACTADAS POR HIDROCARBONETOS DE PETRÓLEO

*“A variabilidade genética é definida como sendo o conjunto
de oportunidades no espaço protéico”*

Prof. Dr. Joaquim Machado

Introdução

Os manguezais constituem um ecossistema de papel importante na manutenção da vida marinha em regiões tropicais e subtropicais. Devido suas características de solo, como por exemplo, alta concentração de matéria orgânica e baixa tensão de oxigênio, poluentes derivados de atividades antropogênicas, como refino do petróleo, produção agrícola com adição de pesticidas, indústria pesqueira, entre outras, se tornam altamente evasivas gerando efeito cumulativo no solo. Esse acúmulo em longo prazo e em determinadas concentrações de poluentes, comprometem processos essenciais para a manutenção dos ecossistemas (CURY, 2002). A contaminação de manguezais por derramamentos acidentais de petróleo e trânsito de embarcações cargueiras é freqüente em todo o mundo (JACKSON *et al.*, 1989). A Baía de Guanabara (Rio de Janeiro, Brasil) é um exemplo de região afetada por atividade de indústria do petróleo, tráfego intenso de navios e descarte de esgoto doméstico da região metropolitana. Além das intensas atividades impactantes, em Janeiro de 2000 a região sofreu com um derramamento acidental de 1,3 milhões de litros de óleo causado por um rompimento de ductos de óleo da região nordeste da Baía (BRITO *et al.*, 2006). Sabe-se que a região selecionada para a coleta em nossos estudos é altamente impactada sendo nítida a presença de óleo nas amostras coletadas, constituindo um ambiente adequado para recuperação de organismos resistentes a hidrocarbonetos derivados de petróleo e para o estudo de genes catabólicos. A diversidade de comunidades microbianas adaptadas à presença de óleo deve ser abordada sob dois aspectos: estudos dependentes de cultivo, isolamento e enriquecimento e, estudos independentes de cultivos, genética molecular, genômica e metagenômica (VAN HAMME, *et al.*, 2003). Bibliotecas metagenômicas rDNA 16S oriundas de fragmentos de DNA ambiental têm como principal finalidade caracterizar organismos ainda não isolados (SEBAT *et al.*, 2003.)

A fim de elucidar os componentes microbianos adaptados às condições de poluição, o presente trabalho teve como proposta inicial fazer um breve

levantamento da diversidade bacteriana presente na amostra de sedimento impactado com petróleo.

Material e Métodos

Coleta das amostras de sedimento de mangue

Os sedimentos utilizados neste trabalho foram coletados no mês de março de 2006 no Canal do Fundão na região da Ilha do Fundão, Rio de Janeiro (RJ). As coordenadas de coleta foram orientadas pelo equipamento GPS GARMIN eTrex com “waypoint” sob as seguintes coordenadas geográficas: latitude 22°51'18.4” Sul e longitude 43°14'11.4” Oeste. A região se caracteriza por receber grandes volumes de esgoto doméstico além de contaminações por hidrocarbonetos derivados de petróleo. A coleta foi realizada segundo padrões US-EPA (1989), com distribuição aleatória de amostragem (Figura 12). Após a coleta, alíquotas dos sedimentos foram encaminhadas para análises físico-químicas com o intuito de determinar a concentração de metais pesados (Tabela 2).

Figura 12. Canal do Fundão (background) (A). Artigos para coleta: termômetro (B); trado (C); resíduos urbanos na região de coleta (D); sacos de coleta estéreis (E); alíquota da amostra de sedimento em placa de Petri (F).

Os valores do potencial hidrogeniônico (pH) e temperatura foram mensurados logo após a coleta. Alíquotas das amostras foram acondicionadas em frascos de vidro e encaminhadas aos laboratórios da Embrapa Solos (RJ) para análises de concentração de HPA's em CG-EM (Tabela 3). O material de estudo foi armazenado a 5°C até o início dos procedimentos de montagem dos microcosmos.

Tabela 2. Teor total de metais (método: SW 3051 EPA-US).

Metais (mg/Kg ⁻¹)	CONAMA 2004 ^a	CETESB	CETESB ^c	Sedimento
Cd	0,6 – 3,5	3,3	6,1	3,7
Pb	35 – 91,3	181	312	239
Cu	35,7 – 197	47	79	357
Cr	37,3 – 90	120	190	78,9
Ni	18 – 35,9	35	60	51,3
Zn	123 – 315	154	257	780
Co	-	29	53	

Resolução 344 – (a) Valores para disposição do material em água salina-salobra, limiar abaixo e acima do qual prevê provável efeito adverso à microbiota; (b) Valores orientadores para risco de saúde humana e ambiental.; (c) Valores de alerta para metais em solos.

Tabela 3. Concentração de hidrocarbonetos poliaromáticos (HPA) das amostras coletadas Canal do Fundão (RJ).

HPA	mg/kg ⁻¹
Naftaleno	0,11450
Acenaftaleno	0,02321
Acenafteno	0,10884
Fluoreno	0,06918
Fenantreno	0,14430
Antraceno	0,09695
Fluorantreno	0,61746
Pireno	0,57981
Benzeno(a)antraceno	0,15914
Criseno	0,10036
Benzo(b)fluoranteno	0,21619
Benzo(k)fluoranteno	0,05284
Benzo(a)pireno	0,14091
Indeno(1,2,3...cd)pireno	0,08089
Dibenzo(a,h)antraceno	0,02040
Benzo(ghi)perileno	0,09038
Total	2,61537

Montagem dos microcosmos e enriquecimento direcionado para degradadores de HTP e HPA's

O experimento em microcosmo foi montado em triplicatas de amostras, controle (sem adição de petróleo) e branco (sem amostra), para respirometria (Figura 13). O monitoramento de produção de CO₂ foi feito diariamente em um prazo de 60 dias. Além dos ensaios de respirometria mensurados diariamente, em cada quinzena é retirada três triplicatas do experimento, controle, 3% de contaminação e 10% de contaminação, e foi determinada a atividade enzimática em cada período.

Figura 13. Imagem do experimento em microcosmo. Amostras controle, 3 e 10% de petróleo (A, B, C, respectivamente). Vista geral das jarras de ensaio (D), detalhe de jarro montado (E,F,G,H).

Após a adição do petróleo nos experimentos, o pH foi medido separadamente sendo a amostra controle mais alcalina em relação aos tratamentos (3 e 10% de petróleo) (Tabela 4).

Tabela 4. Controles de pH no início do tratamento (t=0).

Amostra	pH
Controle	7,78
3% Petróleo	6,76
10% Petróleo	6,79

Determinação da taxa de respiração microbiana

Alíquotas de 150 g de sedimentos submetidas aos diferentes tratamentos foram colocadas em frascos hermeticamente fechados e submetidas à captura do carbono via CO_2 liberado pela respiração microbiana. A captura foi realizada pela reação química do C- CO_2 com uma solução aquosa de NaOH 0,5 M após 24 horas de exposição. Após este período de exposição era adicionado 1 mL de BaCl_2 0,5 m para parar a reação. A quantificação do carbono mineralizado é feito por meio de titulação com HCl 0,2 M, utilizando-se fenolftaleína como indicador do ponto de viragem. O

carbono mineralizado é expresso em mg de C-CO₂ 150 g⁻¹ de amostra, com base na seguinte fórmula:

$$C - CO_2 = (B - A) \times Eq \times M$$

onde,:

B= valor gasto de HCl na titulação do branco;

A= valor gasto de HCl na titulação da amostra;

Eq= equivalente do carbono = 6;

M= molaridade do HCl utilizado na titulação.

A respiração microbiana na amostra representa a soma de todos os processos metabólicos produzindo CO₂ como resultado da captura de O₂ por microrganismos metabolicamente ativos na amostra (ANDERSON, 1982).

Determinação da atividade de enzimas microbianas em amostras de sedimentos

A determinação das enzimas de metabolismo microbiano em amostras de sedimentos seguiu metodologia descrita por Alef e Nannipieri (1998) para determinação de enzimas em amostras de solo.

A atividade de beta-glicosidade foi determinada pelo método colorimétrico com paranitrofenol, liberado pela beta-glicosidase quando o solo é incubado com paranitrofenil-beta-D-glicosídeo. O paranitrofenol liberado é extraído por filtragem e determinado colorimetricamente. O tolueno foi utilizado para inibir a atividade microorgânica durante a incubação, sendo o dado obtido resultado apenas da atividade de enzimas livres ou imobilizadas no solo.

Para fosfatase alcalina, os procedimentos foram baseados na hidrólise de p-nitrofenil-fosfato, liberando p-nitrofenol, o qual é extraído quantitativamente do solo e medido colorimetricamente. O tolueno também é utilizado para inibir a atividade microbiana no momento da incubação.

A análise da atividade desidrogenase baseia-se na atividade oxirredução, que catalisa a transferência de elétrons de uma molécula para outra. Para verificar a atividade da enzima desidrogenase foi utilizado cloreto de trifeniltetrazólio (TTC) para um sistema de tamponamento e extração do produto enzimático mediante a adição de acetona e determinado por colorimetria.

A quantificação de lipase em sedimento foi realizada utilizando-se o método proposto por Marguesin e colaboradores (2002). O extrator utilizado foi paranitrofenilbutirato (pNPB), a reação foi interrompida por imersão em banho de gelo e foi feita a leitura espectrofotométrica.

Enriquecimento direcionado para organismos degradadores de hidrocarbonetos derivados de petróleo.

Foram inoculadas 5 g de sedimento de mangue de cada tratamento impactado em meio mineral 2% salino (Tabela 5), o que chamamos de pré-enriquecimento. O pré-enriquecimento foi deixado a 30 °C sob agitação 120 rpm por 24h. Passada as 24h foi aliqotado 2 mL do pré-enriquecimento e inoculado em 45 mL meio mineral salino (Zinder – modificado + metais traço (Tabela 6) e 3 mL de óleo diesel comercial filtrado em filtro 0,22 µm, o que chamamos de enriquecimento (Figura 14). Os enriquecimentos foram feitos em triplicatas, onde cada réplica foi removida das condições de climatização em 24h, 72h e 7 dias e preservadas em geladeira.

Tabela 5. Meio Mineral 2% salino (Zinder - modificado).

Componente	Quantidade
NH ₄ Cl	0,5 g
KH ₂ PO ₄	0,4 g
MgCl ₂ . 6 H ₂ O	0,1 g
CaCl ₂ .2 H ₂ O	0,05 g
NaCl	17 g
Solução Traço de Metais (Tabela 6)	10 mL
H ₂ O MILLI-Q	q.s.p. 1000 mL

Tabela 6. Soluções-traço para Meio Mineral 2% salino (Zinder - modificado).

Componente	Quantidade
FeSO ₄ . 7 H ₂ O	0,556 g
MnSO ₄ . H ₂ O	0,086 g
CoCl ₂ . 6 H ₂ O	0,17 g
ZnSO ₄ . 7 H ₂ O	0,21 g
H ₃ BO ₃	0,19 g
MiCl ₂ . "H ₂ O	0,020 g
Na ₂ MoO ₄ . 2 H ₂ O	0,01 g
H ₂ O Milli- Q	q.s.p. 1000 mL

Uma alíquota de 2 mL foi retirada de cada enriquecimento e levadas a microscopia para fixação do material e observação com corante DAPI. O volume restante foi preservado a 4 °C até o dia de extração do DNA. As amostras de enriquecimento armazenadas foram centrifugadas a 8.000 g por 20 minutos e o precipitado foi armazenado em gelo até o início dos procedimentos de extração de DNA.

Figura 14. Ensaios de pré-enriquecimento de amostras de sedimento de mangue provenientes de microcosmos impactado com petróleo.

Isolamento e identificação de colônias a partir de enriquecimentos

O isolamento de microrganismos, apesar de recuperar uma fração muito pequena da microbiota nativa de determinado ambiente, se torna um método útil e rápido quando se quer avaliar o potencial metabólico de determinados microrganismos. Sendo assim, junto aos demais experimentos, foram plaqueados em meio mineral 2% salino (Tabela 5) com adição de ágar (15 g/L) para solidificar o meio. Foi adicionado 50 μL de alíquotas de enriquecimento provindos de sedimento impactado com 3 e 10% de petróleo.

Para a triagem de bactérias degradadoras de HPA's foi adicionado ao meio, como única fonte de carbono, solução de naftaleno 250 mg mL^{-1} , diluída em acetona. As placas foram deixadas em fluxo laminar por 2h para evaporar a acetona e armazenadas em geladeira. O tempo de crescimento foi de 7 dias a 30 °C. Foi feita a seleção das colônias por aspectos morfológicos (Figura 16) e coloração de Gram para confirmação de pureza das colônias isoladas.

As colônias isoladas foram preservadas por ultra-congelamento a -80 °C. A identificação dos isolados foi feita por análise de sequência do gene RNA ribossomal 16S após extração de DNA (YONG e BLAKESLEY, 1991) e amplificação do rDNA com primers universais para o domínio Bactéria (LANE, 1991).

Visualização por DAPI de amostras de enriquecimento aeróbio com óleo diesel - microscopia de epifluorescência

As amostras de enriquecimento provenientes do primeiro repique dos microcosmos apresentavam 10% (v/v) de óleo diesel esterilizado além dos resíduos de petróleo proveniente da massa de sedimento inoculada. Esses interferentes dificultaram a aplicação dos protocolos comumente encontrados na literatura. Além disso, os trabalhos publicados não detalham os passos de limpeza da amostra, o que certamente inviabilizam a geração de dados de contagem seguros. Deste modo foi necessário desenvolvemos nosso próprio método de limpeza, preservando as características particulares dos enriquecimentos.

Para a observação de bactérias degradadoras presentes na amostras, utilizamos corante 4'-6'-diamidino-2-fenilidol (DAPI), partindo de uma solução estoque de concentração 0,5 mg/mL. A solução estoque foi diluída para solução de uso com concentração de 0,4 µg/mL. A solução de uso foi preparada no momento de confecção das lâminas. Foram tiradas alíquotas dos enriquecimentos aeróbios (15, 30, 45 e 60 dias de experimento em microcosmos) e visualizadas observadas em microscópio *Olympus* BH com filtro para DAPI, rodamina / CY3 e autofluorescência 420 nm.

A contagem dos organismos nos campos observados seguiu recomendações descritas por Kirchman e colaboradores (1982).

O protocolo otimizado (Anexo E) para observação com DAPI mostrou ser eficiente na observação do crescimento de microrganismos de amostras ambientais complexas (Figura 15).

Figura 15. Observação com DAPI. Bactérias de enriquecimento aeróbio com 3% de óleo diesel. Material controle (A) e cultivados 24h (B); 72h (C) e 7dias (D). Aumento de 60X.

Caracterização da diversidade taxonômica nas amostras de sedimento de mangue da Baía de Guanabara

A caracterização da composição de espécies microbianas contida no metagenoma das amostras de sedimento de mangue estudadas foi feita através da análise filogenética de sequências do gene RNA ribossomal. Esta caracterização contribuiu para se conhecer a diversidade microbiana representada na biblioteca e estabelecer correlações entre esta e sequências gênicas de interesse obtidas nas análises de triagem de genes catabólicos. Os resultados foram completados com 214 sequências analisadas.

A extração do DNA total seguiu a metodologia de lise direta de células, descrita por Rondon e colaboradores (2000), com algumas adaptações. A integridade dos fragmentos foi observada em gel de agarose 1%. Aproximadamente 50 ng do DNA extraído foram usados como molde para reação de PCR com volume final de 25 µL. Os *primers* utilizados para a amplificação do gene rDNA 16S da amostra foram: 27f [5'- AGA GTT TGA TCM TGG CTC AG - 3'] e 1100r [5'- AGGGTTGCGCTCGTTG- 3'] (LANE,1991). O programa de amplificação utilizado consistiu na primeira desnaturação a 95°C por 2 min.; 10 ciclos a 94 °C/1 min., 62 °C/30 seg. e 72 °C/2min., com diminuição de 0,5 °C por ciclo; e 20 ciclos a 94 °C/1 min., 57 °C/30 seg. e 72 °C/2min., extensão final a 72 °C por 3 min. Após a amplificação do rDNA 16S os produtos foram purificados com kit QiaexII™ (QIAGEN®) e quantificados em gel de agarose 1% (Figura 18). A partir dos produtos de PCR purificados, foi construída uma biblioteca de rDNA 16S em vetor plasmidial de clonagem de pGEM-T Easy em *E. coli* JM-109.

Os fragmentos rDNA 16S clonados foram seqüenciados utilizando o par de *primers* 10f (5' - TTT AAT GAA GTA GTT GCA ATC TGG - 3') e 1100r (5'- AGG GTT GCG CTC GTT G - 3'), adicionados na concentração de 3,2 pmoles por reação (Tabela 7).

Tabela 7. Reação de seqüenciamento para 16S rDNA (desnaturação: 95 °C/20 segundos; 50 °C/15 segundos; 60 °C por 1 minuto em 30 ciclos de repetição).

Reagentes	Volume
DNA (100 ng a 200 ng)	3,0 µl
Primer 10f ou 1100r	0,16 µl
Mix ET	40,0 µl
H ₂ O	2,84 µl
Volume final	10,0 µl

As seqüências obtidas, através do uso de seqüenciador automático MegaBace de 96 capilares (MolecularDynamics), foram alinhadas para a formação de contigs, usando-se o pacote Phred/Phrap/CONSED versão Linux (GODON *et al.*, 1997) para a montagem dos *contigs*.

A seqüência de contig, com aproximadamente 1000 pb, foi então submetida à comparação com seqüências dos bancos de dados Genbank (<http://www.ncbi.nlm.nih.gov/BLAST/>) e Ribosomal Data Project II 9.0 (<http://rdp.cme.msu.edu/index.jsp>). As seqüências recuperadas dos bancos de dados foram alinhadas no programa CLUSTAL X (THOMPSON *et al.*, 1997) e as análises filogenéticas foram conduzidas utilizando o programa MEGA versão 2.0 (KUMAR *et al.*, 2001).

A matriz de distância evolutiva foi calculada com o modelo de Kimura (KIMURA, 1980) e a construção da árvore filogenética a partir das distâncias evolutivas foi feita pelo método de *Neighbor-Joining* (SAITOU e NEI, 1987), com valores de *bootstrap* calculados a partir de 1.000 re-amostragens, utilizando o software incluído no programa MEGA 2.0.

Resultados e Discussão

Quantificação C-CO₂ produzido em microcosmos

Comunidades microbianas resistentes a petróleo vêm sendo extensamente estudadas nos últimos anos, principalmente no que diz respeito à capacidade de degradar ou biotransformar componentes específicos do óleo. Uma forma de avaliar o potencial de biodegradação de uma comunidade microbiana de sedimento é através de experimentação em microcosmos, determinando-se a respiração microbiana basal (RMB), que representa o resultado dos processos metabólicos de todos os organismos vivos presentes na amostra.

A respiração microbiana do sedimento representa a soma de todos os processos metabólicos produzindo CO₂ como resultado da captura de O₂ por microrganismos metabolicamente ativos na amostra (ANDERSON, 1982). A Figura 16 mostra a respiração acumulada (μg de C-CO₂ 150 g⁻¹ de amostra) das amostras de sedimento de mangue submetidos a diferentes tratamentos. Nos tratamentos onde foram adicionados óleo cru (3 e 10%) foi possível visualizar um crescente aumento de liberação de CO₂, comparado ao sedimento controle.

Finalizado o primeiro período de observação (30 dias), pode-se observar que o sedimento submetido a 10% de óleo cru superou o volume acumulado de CO₂ do sedimento submetido a 3% de óleo cru. Esta tendência foi observada até o sexagésimo dia (Figura 16). Tal observação pode estar relacionada a adaptação do consórcio microbiano à toxicidade inicialmente imposta e a liberação de compostos intermediários que favoreceram o co-metabolismo microbiano nesse período. Ainda, pode ser observado que a liberação de CO₂ dos tratamentos com 3% e 10% de óleo cru se manteve muito superior, praticamente o dobro, ao produzido pelos controles nestes dois períodos. Estes resultados comprovam uma intensa taxa de mineralização do carbono orgânico proveniente dos hidrocarbonetos presentes no óleo cru.

Figura 16. Produção de Carbono (CO₂) acumulado nos períodos de 15 dias (A); 30 dias (B); 45 dias (C) e 60 dias (D). Eixo X: Tempo em dias. Eixo Y: µg de C-CO₂ 150 g⁻¹ de amostra.

Determinação de enzimas de metabolismo microbiano em amostras de sedimento

A oxidação biológica de compostos orgânicos está conectada com a desidrogenação (liberação de hidrogênio), conduzida pela ação de enzimas denominadas desidrogenases. Estas enzimas estão presentes somente em organismos ativos e sua atividade está relacionada com a presença de compostos orgânicos, como hidrocarbonetos. Os resultados apresentados na Tabela 8 para os primeiros 15 dias do experimento de microcosmos demonstram valores superiores para as réplicas de 3% e 10% em relação ao controle. Esta mesma tendência também foi observada aos 30 dias, contudo com valores absolutos menores.

Tabela 8. Valores para determinação de atividade de enzimas em sedimentos.

Enzima	Desidrogenase		Beta-glicosidase		Fosfatase alcalina		Lipase	
Tempo (dias)	15	30	15	30	15	30	15	30
Controle	262,61	257,55	106,19	99,36	89,86	84,44	99,58	101,61
3%	278,52	286,76	111,72	106,55	96,85	115,74	110,55	165,81
10%	312,85	282,49	108,26	103,66	94,74	125,18	107,22	198,92

A beta-glicosidase é uma hidroxilase diretamente relacionada ao ciclo de carbono. Os resultados apresentados demonstram uma tendência de diminuição da atividade de beta-glicosidase em função do tempo. Essa tendência foi também observada por Ceccanti e colaboradores (2006) em estudos com solos. Em modo inverso, a concentração de fosfatase alcalina foi crescente nos tratamentos com adição de petróleo.

A lipase é uma enzima produzida por grande parte dos microrganismos presentes em solos impactados com óleos, capaz de degradar lipídeos em glicerina e ácidos graxos. Na Tabela 8 pode-se observar a diferença na concentração da lipase entre os controle e os tratamentos com 3 e 10% de petróleo no período de estudo. Observa-se também um pequeno aumento da concentração de lipase nos tratamentos entre o 15° e 30° dia.

Os estudos realizados para detecção de enzimas microbianas em sedimentos mostram um metabolismo ativo de desidrogenação, indicando um impacto causado pelos hidrocarbonetos do óleo aos microrganismos dos sedimentos. Assim como a adição de petróleo influenciou na taxa de respiração da comunidade presente, a produção de enzimas metabólicas foi afetada, passando pelo mesmo período de adaptação frente aos compostos presentes no óleo cru.

A atividade de beta-glicosidase reflete a presença de elevadas concentrações de matéria orgânica do sedimento/solo, sendo que a sua atividade está estritamente ligada à utilização desses compostos como

substratos. A concentração de beta-glicosidase diminuiu durante o período de observação. Esta tendência já foi reportada por Ceccanti e colaboradores (2003) e reflete possivelmente uma depleção de nutrientes em função da alta atividade metabólica inicial, seguida de uma fase de diminuição da biodisponibilidade de nutrientes essenciais ao metabolismo. Isto fica mais evidente quando observamos os resultados obtidos para atividade da fosfatase alcalina. A fosfatase alcalina está ligada ao ciclo do fósforo, elemento essencial para reações de biossíntese e metabolismo primário dos microrganismos. Uma vez que ocorra a depleção deste elemento, a atividade da enzima fosfatase alcalina é aumentada, para liberação de fósforo orgânico e amenização dos efeitos da carência deste elemento. Assim, como pode ser visto na Tabela 8, a atividade desta enzima foi aumentada em função do tempo, possivelmente para suprir a necessidade de fósforo pelo metabolismo, ao passo que pode-se verificar uma diminuição da atividade da beta-glicosidase. Finalmente, a lipase apresentou aumento de concentração até o final do período de ensaio.

Isolamento e identificação de bactérias a partir dos enriquecimentos

Um total de 34 isolados foram obtidos a partir dos enriquecimentos aeróbios (Figura 17) utilizando os protocolos de isolamento seletivo em meio contendo naftaleno em condições aeróbias.

Os isolados purificados foram submetidos a análise filogenética de rDNA 16S para identificação (Anexo I), conforme descrito anteriormente. A Figura 18 mostra a árvore filogenética construída a partir de sequências do rDNA 16S dos isolados e sequências afiliadas recuperadas dos bancos de dados Genbank e RDP. A análise filogenética revelou que 20 isolados foram relacionados com sequências de Proteobacteria e 14 com Actinobacteria. Os agrupamentos foram suportados por valores elevados de *bootstrap*.

Figura 17. Visualização de colônias isoladas em meio mineral salino (2% Cloreto de sódio) e naftaleno (250 mg mL⁻¹).

Dos isolados afiliados ao grupo Proteobacteria, 11 foram agrupados com Alfa-Proteobacteria, mais especificamente relacionados aos gêneros *Rhizobium*, *Pseudaminobacter*, *Novosphingobium* e *Thalassospira*. Outros 6 isolados foram relacionados a Gama-Proteobacteria e formaram um agrupamento suportado por alto valor de *bootstrap* com espécies do gênero *Pseudomonas* (Figura 19).

Dos isolados afiliados ao grupo Actinobacteria, 3 foram relacionados ao gênero *Rhodococcus*, sendo os isolados isob2 e isob1 relacionados a *Rhodococcus pyridinovorans*, e o isolado isoa2 a *R. erythropolis*. Outros 4 isolados, iso20a, iso21c, iso22c e iso22b, foram posicionados filogeneticamente próximos ao gênero *Mycobacterium*. Os demais isolados, iso22a e iso22g, foram recuperados com sequências de espécies de *Dietzia*.

Como esperado, pela afiliação filogenética dos isolados, podemos observar organismos ou grupos taxonômicos já descritos na literatura em associação com a degradação de petróleo em amostras impactadas ou em óleo bruto.

Figura 18. Árvore filogenética mostrando a relação entre os isolados obtidos em meio mínimo + HPA (naftaleno) e linhagens de microrganismos relacionados com base na sequência de rDNA 16S. (Valores de Bootstrap $\geq 69\%$ são indicados na árvore).

Caracterização da diversidade bacteriana de amostras de sedimento contaminado da Baía de Guanabara

Um total de 214 sequências de rDNA 16S (Anexo II) foram obtidas a partir das amostras de sedimentos. Os *contigs* finalizados foram comparados com sequências depositadas no GenBank utilizando o programa BLASTn (ALTSCHUL *et al.*, 1997). Todos apresentaram alinhamento parcial com sequências de rDNA 16S provenientes de diferentes bactérias. Foram recuperadas sequências com similaridade $\geq 90\%$ com os contigs obtidos. O programa ClustalX (THOMPSON *et al.*, 1990) foi utilizado para o alinhamento das sequências e uma matriz de similaridade construída no software MEGA (*Molecular Evolutionary Genetics Analysis*) versão 3.1 (KUMAR *et al.*, 2004).

A biblioteca genômica de rDNA 16S de sedimento de mangue contaminado com petróleo apresentou clones relacionados a Acidobacteria, Actinobacteria, Chlorobi, Chloroflexi, Deferribacteres, Firmicutes, Proteobacteria e Spirochaetes (Figura 19 A).

Sessenta e cinco clones (32%) foram considerados não-afiliados, ou seja, não pertencendo a nenhuma Divisão ou Filo já descrita dentro do Domínio Bactéria. A grande maioria dos clones se mostrou afiliada ao Filo Proteobacteria, sendo a Classe Delta-Proteobacteria a mais abundante (Figura 19 B).

Figura 19. A distribuição dos clones e suas proporções entre os diversos grupos taxonômicos do Domínio Bactéria e os clones não afiliados (A). Proporção de clones no Filo Proteobacteria das amostras de sedimento de mangue (B).

As análises filogenéticas da biblioteca de rDNA 16S representam a estrutura original das comunidades microbianas presente nas amostras de sedimento coletadas. A cobertura da diversidade foi medida pelos índices de *GOOD* e pelo índice de *ACE*.

$$\text{Cobertura de Good: } C = 1 - F_1/N$$

onde:

F_1 é o número de OTU's representado por apenas 1 clone (singleton)

N é o número total de clones.

$$\text{Cobertura de ACE: } C = 1 - F_1/N_{\text{raros}}$$

onde:

N_{raros} é o número total de clones em OTU's raras, ou seja, em UTOs que têm 10 ou menos clones.

Ambos os índices apresentaram valores igual a 0,54, o que demonstra que uma cobertura parcial da diversidade original dos sedimentos foi obtida. Valores próximos a 1,0 indicam uma cobertura total da diversidade na amostra.

Assim, podemos inferir que os sedimentos possuem uma abundância e riqueza de táxons maior que a representada nas análises filogenéticas realizadas.

A árvore filogenética ilustrada na Figura 20 representa a relação entre as OTU's do Filo Proteobacteria da biblioteca. A maioria das sequências apresentaram índices de similaridade elevados (*cutoff* 97%) com sequências de organismos não-cultivados, em especial de Delta-Proteobacteria. A [única exceção foi o clone P2E11, que se afiliou a *Desulfosarcina* sp. (AB294892). No entanto, as demais classes apresentam representantes com gêneros definidos como, por exemplo, o clone P2A3, com 99% de similaridade com *Delftiatsuruhatensis* (EF421404); P3B12 e P2H3 com 100% de similaridade com sequências de *Alteromonas* sp. (EF587979 e EU100407, respectivamente); P2A11 e P3F10, com 99% de similaridade com *Sulfurimonas* sp. (AY225612) e *Arcobacter* sp. (AB250570), respectivamente; e o clone P3E04, com 100% de similaridade com *Sulfovorum* sp. (EF061935).

Representantes dos demais Filos tiveram poucas sequências representativas em nossas análises (Figura 19). Firmicutes foi o segundo grupo mais representativo, com clones distribuídos entre a Classe *Clostridia*. *Aminobacterium* sp. (AB175355) e *Acholeplasma* sp. (DQ337046) foram similares aos clones P1G2 e P2H2, P1E9, com valores superiores a 87%.

A maior parte das sequências obtidas foram similares a sequências depositas em GeneBank provenientes de organismos não-cultivados, detectados em ambientes de estudo similares ao nosso, incluindo sedimentos marinhos e de mangue impactados por petróleo (Anexo III). Estes resultados sugerem que a composição das comunidades degradadoras em diferentes ambiente contém ainda um grande número de microrganismos desconhecidos, ainda não isolados e caracterizados, reforçando a importância da abordagem metagenômica empregada neste estudo como uma estratégia para acessar novos genes e enzimas na microbiota ambiental.

Figura 20. Análise filogenética do gene rDNA16S a partir de sedimento de mangue impactado com petróleo. (Distância evolutiva baseada no modelo Kimura 2 parâmetros e Neighbour-Joining).

CAPÍTULO 2 - CONSTRUÇÃO E TRIAGEM GENÉTICA DE BIBLIOTECAS METAGENÔMICAS DE ALTO PESO CONSTRUÍDAS A PARTIR DE AMOSTRAS DE SEDIMENTOS IMPACTADAS POR HIDROCARBONETOS DE PETRÓLEO

“Para uma visão sistêmica de integração de fenômenos e de tecnologia da Informação, a genética como Física da Informação.”

Prof. Dr. Joaquim Machado

2.1 Introdução

O termo metagenômica tem sido aplicado em estudos que utilizam amostras ambientais, e se baseia na construção e análise de bibliotecas gene específicas ou funcionais geradas a partir de genomas coletivos presentes em determinada amostra (OSBOURNE *et al.*, 2000). Os estudos iniciais de bibliotecas metagenômicas com o intuito de clonagem e expressão de vias metabólicas iniciaram-se aproximadamente há uma década, com trabalhos exaustivos de amostragem e concentração de biomassa microbiana para obtenção de DNA genômico com qualidade para clonagem direta em vetores BAC (STAIN *et al.*, 1996; BEJÀ *et al.*, 2000; RONDON *et al.*, 2000).

No entanto, novas tecnologias enzimáticas e engenharia genética se desenvolveram durante este período sanando as dificuldades de isolamento e purificação de DNA de amostras de difícil acesso como, por exemplo, kits de amplificação de DNA genômico diretamente do DNA extraído da amostra ambiental (MoBio - <http://www.mobio.com/products/index.php>) e clonagem direta em fosmídeos (Epicentre - <http://www.epibio.com>).

Com o desenvolvimento de recursos biotecnológicos para aplicação em pesquisa e desenvolvimento e a geração de novos equipamentos de sequenciamento (Pirosequenciamento) e programas de análise de sequências, a microbiologia atual atravessa por uma nova era de acesso imediato a recursos naturais, principalmente no que diz respeito a comunidades microbianas não cultivadas. As informações contidas nesses genomas representam fontes inesgotáveis de produtos para aplicações biotecnológicas (RIESENFELD *et al.*, 2004; HANDELSMAN, 2005).

Com o intuito de explorar a diversidade de genes envolvidos na biotransformação de hidrocarbonetos derivados de petróleo, esse capítulo tem como objetivo apresentar os métodos de extração e clonagem de DNA de alto peso molecular e fixação dos clones em membranas de náilon (macroarrays) para detecção e identificação dos clones com homologia à sonda *alkB1*.

2.2 Material e Métodos

2.2.1 Extração de DNA de alto peso molecular

A extração de DNA foi feita com o kit *Fast DNA spin Kit* para solo (Q-BIO gene, BIO-101), gerando fragmentos de DNA que variam de 5 a 35 Kb em concentração, tamanho e pureza aceitáveis para clonagem direta e vetor de clonagem pCC1FOS (Figura 2.1). As amostras de sedimento de mangue foram provenientes de enriquecimentos a partir do microcosmos.

2.2.2 Procedimentos de clonagem de DNA metagenômico

As etapas para clonagem direta em vetor pCC1Fos seguiram estritamente o protocolo do fabricante (www.epibio.com), conforme descrito a seguir. O protocolo para clonagem se divide em 7 passos iniciais:

1. Fragmentação do DNA inserto até aproximadamente 40 Kb;
2. Reparo das extremidades “blunt end”;
3. Seleção dos fragmentos reparados LMP (Low Melting Point) em PFGE;
4. Purificação dos fragmentos selecionados a partir de gel de agarose L.M. 1%;
5. Ligação dos fragmentos “blunt end” ao vetor pCC1FOS;
6. Empacotamento do sistema de ligação e transfecção em EPI300 T-1^R, e;
7. Seleção de clones em placas LB/Clorafenicol

Os procedimentos de clonagem foram iniciados pela reação de reparo das extremidades “*blunt end*”, uma vez que o DNA já apresentava o tamanho recomendado pelo sistema. Os volumes das alíquotas de DNA extraídos foram unidos para que se atingisse a concentração recomendada. A quantificação foi feita em programa *BIORAD quantify*, onde constatamos ter aproximadamente 100 µg de DNA para a reação de reparo das extremidades.

Dando prosseguimento aos procedimentos de clonagem com um gel preparativo de agarose low melting 1% em eletroforese horizontal por aproximadamente 20h. A eletroforese foi realizada em cuba de aproximadamente 20 cm a 35V. Para evitar a degradação do DNA inicial por aquecimento do tampão da cuba de eletroforese, o aparato para a corrida foi montado dentro de câmara fria a uma temperatura aproximada de 4°C. Em seguida o gel foi cortado ao meio e a parte que continha os padrões de peso molecular foi corada com solução de brometo de etídio, e a parte do gel que continha a amostra foi mantida sob refrigeração (Figura 2.2). Após a reação de reparo das extremidades, prosseguiu-se a reação de ligação ao vetor, seguindo as recomendações do fabricante.

Outro parâmetro usado foi a relação entre o tamanho aproximado do inserto recuperado e o vetor, segundo a equação que segue:

$$M_i = T_i \times M_v / T_v$$

onde:

M_i = massa do inserto

M_v = massa do vetor

T_i = tamanho do inserto

T_v = tamanho do vetor

Por exemplo: $M_i = 20.000 \text{ pb} \times 500 \text{ ng} / 8,1 \text{ pb} \Rightarrow 1,234 \text{ } \mu\text{g}$ de inserto

O volume de 80 μL da reação de reparo das extremidades foi concentrado em *speed vaccum* (Eppendorf) para volume final de 30 μL . Portanto, a concentração de DNA que era de 200 ng/ μL , aumentou aproximadamente duas vezes e meia (~500 ng/ μL).

O procedimento de empacotamento dos clones foi realizado pelos seguintes passos:

1. Foi descongelado um tubo de MaxPlax® packing em banho de gelo e adicionado 25 μL aos 10 μL de cada reação de ligação. A

reação foi homogeneizada com pipeta evitando formação de bolhas.

2. A reação foi incubada a 30 °C por 90 minutos.
3. Após a incubação, foi adicionado mais 25 µL MaxPlax® packing e incubado novamente a 30 °C por 90 minutos.
4. Foi adicionado tampão de diluição de fagos PDB (10 mM Tris-HCl [pH 8.3], 10 mM NaCl, 10 mM MgCl₂) para volume final de 1 mL.
5. Foi adicionado 25 µL de clorofórmio e armazenado a 4 °C .

Os fagos foram armazenados a 4 °C até o momento da transfecção em *E.coli* EPI-100 T^R . Para transfecção, preparou-se um pré-inóculo das células *coli* EPI-100 T^R, deixado a 37 °C por 24h. Foi feito um repique para inoculo, 5 mL em 45 mL de LB/MgSO₄, e monitoramos o crescimento em espectrofotômetro Eppendorf até atingir a OD₆₀₀ = 0,9. Em seguida incubamos em estufa pré-aquecida a 37 °C por 20 minutos. O volume foi distribuído em placas contendo LB / Clorafenicol (12,5 µg/mL).

As placas permaneceram em estufa a 37 °C por 20h e em seguida feita a contagem das colônias. As colônias foram coletadas com palitos de madeira estéreis e crescidos em meio Circle Grow + Clorafenicol 12,5 µg/mL por aproximadamente 20h. Foi feita a estocagem da biblioteca em placas de poliestireno 96 poços, estéreis, com adição de ½ volume de glicerol 40% armazenados em ultra-congelamento.

2.2.3 Macroarrays - arranjo e fixação das bibliotecas metagenômicas em membranas de náilon

Como primeiro procedimento foi feita uma cópia da biblioteca metagenômica em placas de 96 poços da marca Genetix. A cópia em placas da marca Genetix fez-se necessária para o reconhecimento e o funcionamento do robô Q-Bot (Genetix) utilizado nos procedimentos de rearranjo de clones em placas de 384 poços (Genetix). Para repique da biblioteca foi adicionado em cada poço 80 µL de meio Circle Grow (CG) + 12,5 µg/mL de Clorafenicol (Sigma). Para o rearranjo em placas de 384 poços, (Figura 2.5) foi utilizada a mesma formulação do meio de crescimento em volume de preenchimento de 30 µL de meio em cada poço. Em ambas as cópias feitas em placas genetix

tiveram o meio de crescimento suplementado com 8% (v/v) de glicerol (Merck). As placas 384 poços foram incubadas a 37°C por 16h (ORION, Estufa de cultura 502, FAMEN – Controlador Modelo A-LT, São Paulo, Brasil) para o crescimento dos clones.

Foram feitos dois macroarranjos de sedimento de mangue impactado com petróleo (Tabela II.1). Os macroarranjos foram configurados em 4X4, feitos em duplicata com cópia em triplicata, finalizando 3 membranas.

Previamente à transferência dos clones, filtros de náilon (PerForma II – Genetix, d= 22,5 X 22,5 cm²) umedecidos em meio CG com 12,5 µg/mL de Clorafenicol e colocados sobre papel Whatman 3M umedecidos no mesmo meio e previamente autoclavados. As membranas de náilon e o papel Whatman foram acomodados nas placas de crescimento do kit do robô Q-BOT, de modo a remover todas as bolhas de ar formadas durante a montagem. As bolhas prejudicam o crescimento bacteriano, não favorecendo a uniformidade de massa celular durante o crescimento controlado.

No processo de transferência (Figura 21), o robô imergiu sua cabeça de 384 pinos (0,4 mm) por 3 segundos nas placas contendo os clones crescidos. Em seguida, imprimiu a membrana com leve pressão por 30 milissegundos, adicionando cerca de 20 a 60 nL de solução contendo bactérias viáveis. Cada clone foi impresso na distância de 1,125 µm um em relação ao outro. Este processo se repetiu nos seis campos, capacidade total da membrana (PerForma II – Genetix, d= 22,5 X 22,5 cm²). A cada ciclo de impressão o robô lava automaticamente os pinos de impressão em água sanitária, seguida de lavagem em água milli-Q, e por fim etanol absoluto (Merck) e jato de ar por 6 segundos.

Figura 21. Esquema resumo de endereçamento de clones em membrana, utilizada para a construção dos macroarranjos (4X4).

Após a impressão, as membranas foram incubadas a 37 °C de 1 a 4h em estufa para o crescimento uniforme das colônias. Esse passo é crítico, à medida que um crescimento demasiado gera a contaminação do arranjo.

2.2.3.1 Processamento das membranas para fixação de DNA

Após o crescimento das colônias as membranas foram colocadas sobre papel de filtro Whatman 3M umedecido em solução desnaturante (1,5 M NaCl (Merck) e 0,5M NaOH (Sigma), onde permanece por 4 minutos em temperatura ambiente. Em seguida o conjunto membranas + filtros foram colocados sobre outro papel Whatman 3M umedecidos em solução desnaturante, onde permaneceram por mais 4 minutos aparados por uma placa de vidro dentro de banho-maria com água fervente. Posteriormente as membranas foram transferidas para outro papel Whatman 3M com solução neutralizante (1,5 M NaCl, 1 M Trizma (Sigma) pH 7,0), onde permaneceram por mais 4 minutos. Transcorrido o tempo, as membranas foram transferidas para uma solução de proteinase K (5 M Na Cl, 1 M Tris (Invitrogen) e HCl (Merck) pH 8,0,), 0,5 M EDTA pH 8,0, 10X Sarcosil (Sigma), 1,5% (p/v) de proteinase K (Invitrogen). A proteinase K foi adicionada à solução momentos antes do uso. As membranas foram acomodadas em recipiente de vidro de modo que o lado de crescimentos das colônias ficasse para baixo, com a face totalmente submersa. As membranas foram incubadas por 1h a 37 °C em agitados orbital (INNOVA 4230, Refrigerator Incubator Shaker, New Brunswick Scientific, Edson, NJ, EUA) a 50 rpm. Após a incubação, o excesso de solução foi retirado com auxílio de papel filtro nas duas faces.

Após o tratamento das membranas, fixou-se o material genético das bactérias nas membranas utilizando o aparelho *Ultraviolet Crosslinker* CL-508, através da exposição a 120.000 microjoules por cm² por 30 segundos.

2.2.4 Construção das sondas de genes catabólicos

A construção das sondas teve início após o rearranjo e fixação dos clones em membrana de náilon (macroarrays). As sondas foram construídas a partir de reação de PCR otimizada para cada par de iniciadores (primers). Os

genes relacionados à degradação de *n*-alcanos foram selecionados para realização dos procedimentos iniciais de hibridização.

Foram feitas extrações de DNA total das linhagens portadoras dos genes catabólicos (para *alkB1*- *Pseudomonas butanovora*^T DSM 2080) com lise direta segundo PIZA *et al.*, 2004. Aproximadamente 50 ng do DNA extraído foram usados como molde para reação de PCR com volume final de 25 µL (Tabela 9). Os primers utilizados para a amplificação dos genes *alkB1*, utilizado nesse estudo, estão descritos na Tabela 10. Nas otimizações foram feitas reações em gradiente de temperatura entre 55 °C a 65 °C.

Tabela 9. Conjunto de primers utilizados para amplificação de genes de monoxigenases de *n*-alcanos.

Primer	Sequência (sentido 5' – 3')	Tamanho (pares de bases)	Temperatura de anelamento
<i>alkB1</i> (F)	AACACAGCCACGAACCTCGGGC	548 pb	56 °C
<i>alkB1</i> (R)	CGCATGGTGATCGCTGTGCCGCTGC		

O programa de amplificação utilizado consistiu na primeira desnaturação a 95 °C por 2 min.; 30 ciclos em temperatura de anelamento definida para cada par de primers 1 min., 72 °C / 2 min e extensão final a 72 °C por 5 min.

Tabela 10. Reação de PCR para preparação de sondas de genes catabólicos.

Reagentes	Volumes
Tampão de enzima 1 X	2,5 µL
dNTP 200 µM	0,2 µL
Primer (F) 1,2 µM	1,5 µL
Primer (R) 1,2 µM	1,5 µL
DNA polimerase (Invitrogen) (1 U)	0,4 µL
MgCl ₂	0,75 µL
H ₂ O	17,15 µL
DNA molde (50 ng/ µL)	1 µL

Após a amplificação dos fragmentos os produtos foram purificados com kit GFX DNA purification kit (Invitrogen) a partir de gel de agarose 1,2% (Figura 2.7). Os produtos de PCR *alkB1* purificados foram utilizados como sondas em posteriores ensaios de hibridização.

2.2.5 Screening genético para detecção de genes catabólicos em macroarranjos de bibliotecas metagenômicas de grandes vetores

Os ensaios de hibridização foram realizados conforme a metodologia descrita por Destefano *et al.* (2003) e Gonçalves e Rosato (2000). Os ensaios de hibridização DNA-DNA foram feitos com os clones obtidos a partir da clonagem direta do DNA metagenômico em fosmídeos impressos em membrana de náilon.

Os produtos de PCR do gene *alkB1*, amplificado a partir de DNA genômico de *Pseudomonas butanovora* DSM 2080^T, foram marcados com o Kit de marcação não-radioativa DIG DNA Labelling and Detection Kit (Roche Applied Science). O controle da marcação foi feito utilizando o DNA - controle (5 µL) contido no Kit.

O DNA (300 a 500 ng) foi diluído para um volume total de 15 µL, desnaturado a 98 °C por 10 min e transferido imediatamente para o gelo. Ao DNA desnaturado foram adicionados 2,0 µL de hexanucleotídeos, 2,0 µL da mistura de dNTP's e 1,0 µL da enzima Klenow e, em seguida, incubado a 37 °C por 22h em banho-maria. A reação de marcação foi interrompida com a adição de 2,0 µL de EDTA 0,2 M, pH 8,0. O DNA marcado foi precipitado adicionando-se 2,5 µL de cloreto de lítio (LiCl) 4 M e 75 µL de etanol absoluto a -20 °C; e mantido por 30 min a -70 °C. Posteriormente, foi centrifugado a 12000 rpm por 15 min a 4 °C e lavado, duas vezes, com 50 µL de etanol 70% gelado, centrifugando-se a 12000 rpm por 2 min a 4 °C. O DNA marcado foi mantido a temperatura ambiente para secar e, em seguida, foi dissolvido em 50 µL de tampão TE (Tampão Tris/EDTA; 10 mM tris-HCl pH 8,0, 1 mM EDTA). O DNA marcado foi diluído 1:10, 1:100 e 1:1000 em TE. O DNA controle foi diluído 1:5 e a partir dessa diluição foram feitas diluições de 1:10 até 1:10000. Uma alíquota (1 µL) de cada DNA, inclusive o concentrado, foram aplicados em

membrana de nitrocelulose e após secagem do DNA, a membrana foi queimada por 2h a 80 °C. A eficiência da marcação da sonda foi verificada de acordo com o procedimento de detecção imunológica (Figura 22).

Figura 22. Procedimento de detecção imunológica para certificar a qualidade de marcação e a diluição do DNA sonda adequada para os procedimentos posteriores de hibridização.

As membranas foram incubadas por 5 min em tampão de ácido maléico (ácido maléico 0,1 M, NaCl 0,15 M, pH 7,5). Em seguida transferidas para 30 mL de solução bloqueadora 1X (solução bloqueadora 10X em tampão de ácido maléico) e mantidas sob baixa rotação (30 rpm) por 30 min. O anticorpo antidigoxigenina conjugado com fosfatase alcalina (anti-DIG-AP) foi diluído em solução bloqueadora 1X (6 μ L para 30 mL), adicionado às membranas e agitado, sob baixa rotação (30 rpm) por 30 min. Após este período, foram efetuadas duas lavagens de 15 min com 50 mL de tampão de ácido maléico, sob forte agitação (60 rpm) e, então, as membranas foram equilibradas por 5 min com 50 mL de tampão de detecção (Tris-HCL 10 mM pH 9,5; NaCl 0,1 M; $MgCl_2$ 10 mM).

As membranas foram embebidas em solução SSC 6X até completa hidratação e transferidas para as garrafas de hibridação. Aproximadamente 0,2 mL de solução de pré-hibridação (SSC 5X; N-laurilsarcosina 0,1%; SDS 0,02%; formamida 50%, agente bloqueador 2X) pré-aquecida a 55 °C foi adicionada para cada cm^2 de membrana e incubada por 6h sob leve agitação a 55 °C, em forno de hibridação (Fisher Scientific). Após este período, as membranas foram incubadas por 18h a 55 °C, sob leve agitação. Após hibridação, as membranas foram lavadas 2 vezes, sob forte agitação, por 15 min em SSC 2X e SDS 0,1% à 60 °C, num volume suficiente para cobri-las.

Em seguida foram realizadas duas lavagens de 15 min em SSC 0,1X e SDS 0,1%, sob forte e constante agitação, em forno de hibridação a 65 °C.

Após as lavagens das membranas a detecção da hibridação foi feita conforme descrito acima, exceto na utilização do substrato, o qual foi substituído pelo substrato quimioluminescente CSPD (fosfato de dissódio-3- (4-metoxiespiro-{1,2-dioxetano-3,2'-(5'-cloro)-tríciclo-[3.3.1.1^{3,7}]-decano}-4-il)-fenil; DIG Luminescent Detection Kit, Roche Diagnostics). Foram adicionados às membranas 10 mL do tampão de detecção contendo 100 µL de CSPD. Após 5 min de incubação, sem agitação, as membranas foram colocadas em sacos plásticos, seladas e expostas a filmes de auto-radiografia (Hyperfilm-Amersham) por 30 min. Essa última etapa foi realizada ao abrigo da luz.

Para a avaliação das membranas e detecção dos clones positivos foi utilizado medidas o software BZScan (<http://tagc.univ-mrs.fr/bioinformatics/bzscan/bzscan-outputs/node15.html>). O software gerou uma imagem qualitativa para cada spot que foi avaliado segundo a similaridade da mancha, e uma qualidade métrica quantitativa sobre uma escala de 0 a 1. Posteriormente o BZScan foi usado o programa para avaliar a qualidade do sinal medido, comparado a um local determinado pelo experimentador. Os procedimentos geraram sinais de diferente intensidade entre fracos, ou duvidosos dentre os clones fixados. Portanto, foram estipulados três spots de alta qualidade para normalização.

Para minimizar variações decorrentes de diferenças técnicas na qualidade do DNA, marcação da sonda, e hibridização entre as condições experimentais, de transformação e normalização foram observadas as triplicatas impressas na mesma membrana com arranjo 4X4.

2.3 RESULTADOS E DISCUSSÃO

2.3.1 Clonagem direta de DNA de Alto Peso Molecular

A extração de DNA com qualidade e tamanhos exigidos para clonagem direta em vetores fosmidiais é tema de discussão em diversos trabalhos que desenvolvem esse tipo de metodologia (MALTSEV *et al.*, 2005). A estratégia de enriquecimento prévio das amostras de sedimento impactadas com petróleo foi uma alternativa eficiente, que permitiu a recuperação de DNA de alto peso molecular para a clonagem em foscídeos. Os fragmentos extraídos das amostras variam de 30 a 35kb, recuperados em gel preparativo 1% de agarose *low melting* (Figura 23).

Figura 23. Esquema de gel preparativo para seleção de fragmentos de alto peso molecular. eluição de fragmentos a partir de gel de agarose low-melting. 1- Ladder 1kb plus; 2- λ -Hind III; 3- DNA controle kit pCC1FOS; 4- alíquota Amostras; 5 vazia; 6 a 10 – Amostras.

2.3.2 Validação dos clones da bibliotecas de alto peso molecular

O DNA fosmidial foi isolado utilizando métodos de mini-preparação com kit *FosmidMax DNA purification kit* (Epicentre), segundo protocolo do fabricante, com modificações feitas nesta etapa para melhor recuperação dos vetores. Este DNA foi empregado em restrição enzimática dos vetores e posterior detecção do inserto clonado em gel de agarose (Figura 24). As letras (A, B e C) representam fragmentos gerados pela digestão parcial dos fasmídeos fechados. Em **A**, o fragmento inteiro, parcialmente digerido, contendo vetor + inserto, de tamanho aproximado a 40 Kb. Em **B**, o inserto (DNA genômico) clonado. Em **C**, o fragmento que representa o vetor pCC1FOS com aproximadamente 8,2 Kb.

Figura 24 – Digestão enzimática de produtos de miniprep. Agarose 1%. (1) DNA λ – 100 ng; (2) DNA λ – 200ng; (3) DNA λ digerido EcoRI ; (4) DNA λ digerido Pst I; (5) Ladder 1 Kb; (6) DNA inserto controle; (7) vetor aberto pcc1FOS, 1 μ g; clone C3 (SED3) digeridos BamHI, BglI e XhoI, respectivamente; (8,9,10).

Outra forma utilizada para validar as bibliotecas metagenômicas obtidas foi utilizando a triagem com sondas específicas e triagem funcional por HTS. A triagem genética foi feita pela hibridização dos clones fixados em macroarrays, utilizando como sonda produtos de PCR marcados do gene *alkB1* (Figura 25). Para o procedimento de validação e triagem foram selecionados 17 possíveis hits positivos entre os clones fixados (864 clones), o que totaliza 1 hit positivo a cada 54 clones hibridizados.

Os clones detectados foram mapeados segundo re-endereçamento (placas 384 poços) e representam os clones originais (placa 96 poços) (Tabela 11).

Figura 25. Membrana de macroarray da biblioteca metagenômica de enriquecimentos a partir de sedimentos impactados com petróleo. Círculos em vermelho destacam clones selecionados como *hits* positivos (portadores do gene *alkB1*).

Tabela 11. Clones identificados preservados em microplacas 96 poços portadores de genes *alkB1*.

Placa	Clones
Metagenoma SED 1	D1e H8
Metagenoma SED 2	H2
Metagenoma SED 4	F8, E7, G1
Biblioteca SED aeróbio 27/12/06	F1, H5, H7
Biblioteca SED aeróbio 27/12/06;	F1, F3, G2
04 Biblioteca SED aeróbio 27/12/06	H1 e H2
05 Biblioteca SED aeróbio 27/12/06	F3, H7, G2

Um total de 17 clones apresentaram sinal de hibridização com a sonda *alkB1*. Os clones selecionados por hibridização foram submetidos aos ensaios para detecção de monoxigenases do tipo *Baeyer-Villiger* (Capítulo 3). A triagem da biblioteca em *macroarray* com as sondas para *alkB1* representou uma estratégia para a identificação de clones com sequências gênicas de monoxigenases, permitindo a otimização dos ensaios de HTS que seriam realizados para detecção de enzimas funcionais. Caso a triagem genética não tivesse sido realizada, a detecção de enzimas funcionais precisaria ter sido realizada num total de 864 clones. Como resultado da avaliação, apenas 17 clones foram analisados, num total de 1,97% da amostra original.

CAPÍTULO 3 - TRIAGEM FUNCIONAL DE BIBLIOTECAS METAGENÔMICAS DE ALTO PESO CONSTRUÍDAS A PARTIR DE AMOSTRAS DE SEDIMENTOS IMPACTADAS POR HIDROCARBONETOS DE PETRÓLEO

“A evolução molecular percorre caminhos adaptativos rumo a *picos* que representam quão bem as proteínas executam tarefas catalíticas específicas.”

Prof. Dr. Joaquim Machado

3.1 INTRODUÇÃO

Ensaio enzimáticos são protocolos experimentais que tornam visíveis as reações químicas catalisadas por enzimas. No formato de triagem de alto desempenho (HTS), esses ensaios são ferramentas essenciais para a exploração da biodiversidade na busca de novas enzimas (MANTOVANI, 2007). Os ensaios no formato HTS utilizam um substrato capaz de liberar um cromóforo ou fluoróforo como consequência da ação enzimática frente a um substrato, possibilitando detectar as reações enzimáticas de interesse (VASCONCELLOS, 2006).

A biodegradação de alcanos e aromáticos envolve sistemas enzimáticos complexos, com atividade hidrolítica responsáveis pela ação primária aos substratos. Deste modo, o objetivo desse capítulo foi acessar novas enzimas utilizando bibliotecas metagenômicas construídas a partir de enriquecimentos de sedimentos contaminados com petróleo através da análise funcionalmente dirigida, ou seja, acessar sequências genômicas totalmente desconhecidas e detectar somente enzimas que são ativas através de técnicas de triagem de alto desempenho para detecção de atividades hidrolíticas como esterases, lipases e epóxido-hidrolases utilizando substratos fluorogênicos.

3.2 MATERIAL E MÉTODOS

3.2.1 Método de triagem em HTS para bactérias isoladas a partir dos enriquecimentos direcionados

Para realização dos ensaios, os microrganismos foram cultivados em NA e incubados a 35 °C por 48h (em placas de crescimento de 2,0 mL). A massa úmida das colônias desenvolvidas foram removidas, transferidas para tubos de polipropileno tipo *Eppendorf* esterilizados e ressuspensas em solução tampão borato pH 8,0 a 20 mM. Todos os experimentos foram realizados em microplaca de polipropileno com 96 cavidades e monitoradas por leitor de fluorescência (*Flashscan 530 Analytic Jena* – Figura 26) durante 10h, utilizando filtros com comprimento de onda de excitação de 365 nm e as leituras feitas a 460 nm (Figura 23)

Figura 26. **A** - microplaca de 96 poços com 200 μ L de capacidade, **B** - leitor de fluorescência *Flashscan 530 Analytic Jena* com filtro de excitação λ_{ex} =365 nm e leitura λ_{em} =460 nm.

Os ensaios foram montados utilizando concentrações bacterianas de 1,0 mg/mL na presença de BSA a 2,0 mg/mL em solução tampão, solução aquosa de periodato de sódio a 1,0 mM e sondas a 100 mM. Todos os ensaios foram realizados em duplicata. Adicionalmente foram realizados controles negativos (substrato + BSA + NaIO₄ em solução tampão sem a presença de microrganismo) para controle da intensidade de fluorescência liberada pela hidrólise não enzimática dos substratos, controles positivos (controles positivos + BSA + NaIO₄ + suspensão celular em solução tampão) para avaliar uma possível interferência na fluorescência do produto resultante da reação de biotransformação, além do controle microbiano (BSA + NaIO₄ + suspensão celular em solução tampão) para monitorar a possível auto-fluorescência dos microrganismos.

O tratamento dos dados foi feito pela diferença entre a média dos valores de intensidade de fluorescência dos ensaios com microrganismos (E) e média dos valores de intensidade de fluorescência emitidos pelo controle negativo (CN). Com base no valor máximo de emissão de fluorescência do ânion umbeliferona em pH 8,8 de 2500 unidades de fluorescência, somente os ensaios que resultaram em valores de intensidade de fluorescência superiores a 100 unidades, ou seja, acima de 5% de conversão, foram considerados positivos.

3.2.2 Triagem enzimática de bibliotecas metagenômicas

A triagem enzimática das bibliotecas metagenômicas iniciou com a otimização dos ensaios para que fosse estabelecido o volume mais adequado do meio de crescimento celular e sobrenadante nos experimentos. As concentrações de 10, 30, 50, 70 e 100 μL da mistura células + sobrenadante foram testadas frente às sondas e também ao controle positivo (Tabela 12).

1. 10 μL de sonda ou controle positivo 2,0 mM; 90 μL de tampão borato pH 8,0; 10 μL de células + sobrenadante ; 10 μL NaIO_4 20 mM e 170 μL 2,4 mg / mL.
2. 10 μL de sonda ou controle positivo 2,0 mM; 70 μL de tampão borato pH 8,0; 30 μL de células + sobrenadante ; 10 μL NaIO_4 20 mM e 170 μL 2,4 mg / mL.
3. 10 μL de sonda ou controle positivo 2,0 mM; 50 μL de tampão borato pH 8,0; 50 μL de células + sobrenadante ; 10 μL NaIO_4 20 mM e 170 μL 2,4 mg / mL.
4. 10 μL de sonda ou controle positivo 2,0 mM; 30 μL de tampão borato pH 8,0; 70 μL de células + sobrenadante ; 10 μL NaIO_4 20 mM e 170 μL 2,4 mg / mL.
5. 10 μL de sonda ou controle positivo 2,0 mM; 100 μL de células + sobrenadante ; 10 μL NaIO_4 20 mM e 170 μL 2,4 mg / mL.

As reações foram feitas tanto com as células repouso, ou seja, as células foram ressuspensas em solução tampão, como também com o sobrenadante com a finalidade de verificar se a enzima em questão estaria sendo expressa intracelular ou exudada no meio. Reações de biotransformação foram feitas para confirmação dos resultados positivos obtidos nos ensaios de triagem enzimática.

Para a confirmação dos resultados obtidos por HTS, os microorganismos foram cultivados em erlenmeyer de 500 mL contendo 250 mL de meio circle grow contendo MgSO_4 10 mM e cloranfenicol 12,5 μg / mL à 37 °C por 12h. Para realização das reações de biotransformação as células foram separadas por centrifugação á 5000 rpm e 18 °C por 10 min. Em seguida cerca de 1,0 g de massa celular úmida foram ressuspensas em 50 mL tampão fosfato pH 7,0 e o sobrenadante armazenado separadamente. Posteriormente foram adicionados 20 μL dos substratos em questão nos erlenmeyers contendo as células em tampão, e 40 μL nos recipientes contendo 250 mL do sobrenadante separados anteriormente das células. As reações foram mantidas sob agitação

overnight á 28 °C e em seguida extraídas com acetato de etila. As aliquotas foram analisadas por CG-EM.

Tabela 12. Otimização da concentração de células + sobrenadante nos ensaios HTS.

Microrganismos	Volume (µL)	Intensidade de fluorescência		
		Ensaio	Controle (+)	Controle (-)
Metagenoma sedimento 2 clone E2	10	87	1220	30
	30	140	950	20
	50	47	589	18
	70	35	419	16
	100	28	382	15
Metagenoma sedimento 2 clone B1	10	107	1291	30
	30	141	1000	20
	50	50	805	18
	70	34	400	16
	100	26	299	15

3.3 RESULTADOS E DISCUSSÃO

A Tabela 13 apresenta os isolados que apresentaram valores expressivos na análise de triagem enzimática de alto desempenho. Destacaram-se os isolados 5, 7 e 27B, que apresentam valores superiores a 10% de conversão do substrato, chegando a valores próximos a 30%. Os valores apresentados na Tabela 13 comprovam que os microrganismos isolados apresentam vias de produção de enzimas do tipo monoxigenases, lipases e esterases, as quais participam ativamente na degradação de óleo e derivados.

Esses dados enriquecem os resultados obtidos em análises de enzimas de sedimento e respirometria das amostras de sedimento, enriquecidas com petróleo, as quais apontam mineralização dos compostos hidrocarbonetos pela atividade de microrganismos autóctones totais.

Tabela 13. Resumo dos resultados obtidos dos ensaios de atividade enzimática em bactérias isoladas de enriquecimentos com óleo diesel.

Isolado	EP1	EP2	ES1	ES2	LIP	BV1	BV2	BV3
22B		169	137	264				
6		140	126					
20A			195	146				
11		139	169					
22J			198	145				
10			168					
21F	176							
22C	179							
21C			151					
5			362					
7			362					
20B		110		110				
16B					417			
27B		443	229					561

Os resultados indicam os valores de intensidade de fluorescência observado para cada isolado. Os ensaios considerados positivos foram aqueles que apresentaram valores superiores a 100 unidades, ou seja, acima de 5% de conversão considerando o máximo de fluorescência do ânion umbeliferona em pH 8,0 de 1900 unidades de fluorescência.

3.3.1 Triagem enzimática de biblioteca metagenômicas

No caso da biblioteca construída a partir de sedimento (864 clones), a avaliação por HTS revelou 3 *hits* positivos para a hidrólise de acetatos (sonda 1a, Figura 8), 1 *hit* para a hidrólise de propionatos (sonda 1b, Figura 8) e 3 hits para a hidrólise de epóxido (sonda 2a, Figura 8). Os resultados estão resumidos na Tabela 14.

Tabela 14. Resultados obtidos pela triagem enzimática para hidrolases da biblioteca metagenômica construída a partir de sedimento utilizando as sondas fluorogênicas 1a, 1b, 1c, 2a e 2b.

Placa	<i>Hit</i> positivo (% de conversão com relação ao controle positivo – sonda ^a)
Metagenoma SE 05/12/06	B4 (40 – 1a); F4 (30 – 1a)
Met. Sed. 4G5	D7 (80 – 2a); E7 (87 – 2a); G3 (82 – 2a)
Sedimento aeróbio 2	-
Metagenoma sedimento 2	-
Sedimento aeróbio 1	-
Sedimento aeróbio 3	-
Met. Sed. 1	B6 (20 – 1b)
Biblioteca sedimento aeróbio 27/12/06	H5 (60 – 1a)
Sedimento aeróbio 27/12/06	-

^aAs sondas estão referenciadas segundo a legenda da Figura 8.

Esses resultados revelam uma frequência de 1 *hit* positivo para cada 216 (1:216) clones capazes de codificar funcionalmente esterases/lipases, e uma frequência de 1:288 clones testados no caso de codificar genes que expressem epóxido-hidrolases.

Diante dos resultados, foi feita a confirmação da atividade hidrolítica por biocatálise convencional. O foco dessa etapa do trabalho foi testar a habilidade de resolução dos *hits* detectados para hidrólise de ésteres frente aos derivados de feniletanol (Figura 27, substratos 17 e 18). No caso dos *hits* detectados para hidrólise de epóxidos, o objetivo foi testar a resolução do epoxiglicidil, que é um importante bloco construtor para a síntese de fármacos (Figura 27, substrato 20).

Figura 27. Resolução cinética dos substratos 17, 18 e 19.

O *hit* met. sed. 1 B6 que apresentou atividade para a hidrólise de propionatos (sonda 1b, Figura 6) nos experimentos de HTS testados frente ao substrato 18. Os resultados obtidos foram confirmados através da observação da formação do álcool correspondente. Além disso, esse clone apresentou-se altamente seletivo ($E > 100$), hidrolisando seletivamente um dos substratos com alta taxa de conversão e excesso enantiomérico (Tabela 15).

Tabela 15. Reação de hidrólise dos substratos 17 e 18 utilizando como biocatalisadores os hits positivos testados por HTS.

<i>Hit</i>	Substrato	Condição	% conversão ^a	% e ^a e ^b	<i>E</i>	Configuração absoluta
met. sed. 1 clone B6	18	Células	42	97	>100	<i>R</i>
		Sobrenadante	-	-	-	-
metagenoma SE 05/12/06 clone B4	17	Célula	20	> 99	>100	<i>R</i>
		Sobrenadante	-	-	-	-

^aCálculo da porcentagem de conversão em relação ao padrão interno.

A confirmação da configuração absoluta do produto da hidrólise foi feita a partir da co-injeção do álcool racêmico com o padrão enantiomericamente enriquecido *R*. A hidrólise preferencial do éster de configuração *R* concorda com a 'Regra de Prelog', que é baseada no tamanho dos grupos substituintes ligados ao estereocentro.

O *hit* metagenoma SE 05/12/06 B4 a reação com o sobrenadante não apresentou conversão do éster. A reação com as células levou à formação enantiosseletiva do álcool (*R*)-19 que é produto da hidrólise do substrato 17, porém nesse caso a taxa de conversão foi baixa (Tabela 15).

DISCUSSÃO GERAL

As amostras de sedimentos estudadas apresentaram uma resposta adaptativa natural da comunidade microbiana nos ensaios de microcosmos enriquecidos com petróleo, demonstrando atividade de mineralização do petróleo, como observado pela evolução nas taxas de geração de CO₂ em relação ao controle no período de 60 dias. Estes valores foram comparáveis a resultados da literatura em estudos similares de respirometria em solos e sedimentos impactados com petróleo e metais pesados (ERICKSSON *et al.*, 2001; NAKATSU *et al.*, 2005).

Nos ensaios realizados para verificar a presença de enzimas microbianas nos sedimentos, detectou-se concentrações de lipases microbianas em sedimentos proveniente da microbiota adaptada nas amostras estudadas, indicando a presença de microrganismos degradadores na amostra. Os isolamentos feitos a partir de enriquecimentos de amostras do microcosmos permitiram a recuperação de microrganismos aeróbios capazes de metabolizar compostos do petróleo, utilizando diesel como única fonte de carbono no meio seletivo. A presença de lipases nas bactérias isoladas dos enriquecimentos foi confirmada posteriormente pelos ensaios enzimáticos de HTS. Nestes ensaios, o isolado iso16B, relacionado a *Thalassospira lucentensis* QMT2^T, foi produtor de lipase (Tabela 13). *Thalassospira* é citado na literatura como isolado de ambiente marinho pertencente ao grupo das Alfa-Proteobacteria (LOPEZ *et al.*, 2002). Contudo, estes autores não verificaram atividade de lipase em ensaios utilizando Tween 80% para o isolado QMT2^T. Contudo, outros isolados com elevada similaridade de rDNA 16S com a linhagem iso16B (99%), incluindo iso27B e iso27C, também não apresentaram atividade de lipase nos HTS realizados. Do mesmo modo, as linhagens iso19B e iso19C, que possuem 98% de similaridade na sequência de rDNA 16S em relação à cepa QMT2^T, não apresentaram atividade para lipase e esterase.

Esterases e lipases são enzimas amplamente distribuídas entre diferentes táxons bacterianos, como observado em estudos anteriores. Vasconcellos (2006) fez estudos enzimáticos utilizando microrganismos

isolados de enriquecimentos de petróleo e pode observar níveis elevados na produção dessas enzimas por diversas espécies do gênero *Bacillus*, *Halomonas* e *Pseudomonas*. Em nossos estudos, utilizando os mesmos substratos que Vasconcellos (2006) (Figura 8), verificamos atividades para 17 dos 34 isolados obtidos das amostras de sedimentos enriquecidos, com capacidade de conversão preferencial das sondas ES1 (esterase), ES2 (esterase) e LIP (lipase), nesta ordem. Isolado isso 16B apresentou elevada produção de lipases em relação aos demais. Portanto podemos concluir que o substrato de escolha para atividade enzimática da maioria das bactérias isoladas nesse estudo foi o substrato ES1, sendo que o isolado 16B apresentou atividade em ambos, ES1 e LIP.

Os ensaios de enriquecimento com diesel permitiram o desenvolvimento de uma microbiota degradadora nos microcosmos. Esses resultados puderam ser confirmados pela observação em microscopia com o marcador DAPI. Os períodos de enriquecimento geraram um aumento significativo do número de indivíduos durante as fases de incubação de 24, 72 horas e 7 dias, viabilizando a extração direta de DNA de alto peso molecular para clonagem em vetores de alta capacidade. Os enriquecimentos possibilitaram também o isolamento de bactérias resistentes ao petróleo. Dentre os 34 isolados identificados, quatro grupos foram representativos: Alfa-Proteobacteria (32,2%), Gama-Proteobacteria (20,3%), Firmicutes (5,8%) e Actinobacteria (41,7%). Neste aspecto podemos verificar grandes diferenças entre a distribuição de organismos autóctones colonizadores do sedimento e a coleção de microrganismos recuperados a partir dos enriquecimentos com óleo diesel e naftaleno.

Podemos verificar na Figura 18 uma baixa representatividade de Actinobacteria (0,5%) nas sequências de rDNA 16S analisadas. No entanto, este representou o maior grupo de isolados recuperados a partir dos isolamentos (41,7%). Além disso, os sequenciamentos demonstraram a presença marcante da classe Delta-Proteobacteria entre as comunidades analisadas. Contudo, nenhum representante deste grupo foi recuperado nos isolamentos.

Frente aos resultados obtidos no estudo, pudemos concluir que a estratégia de enriquecimento das amostras de sedimento impactado com petróleo foi fundamental para a obtenção de DNA de alto peso molecular em concentração adequada para os ensaios de clonagem em foscídeos. Esta estratégia permitiu, ainda, o isolamento de microrganismos degradadores com atividades enzimáticas interessantes para aplicações industriais. Contudo, a maior contribuição do estudo foi na obtenção e triagem da biblioteca metagenômica, a qual permitiu a identificação de clones com as propriedades enzimáticas desejadas de lipase e esterase com elevada afinidade pelo substrato, e também com elevado excesso enantiomérico e enantioespecificidade (Tabela 15).

Neste sentido, podemos afirmar que este representa o primeiro estudo do gênero no país, englobando desde a construção de bibliotecas de alto peso a partir de amostras ambientais impactadas, até a triagem genética e ensaios de biocatálise convencional para detecção de enzimas para aplicações biotecnológicas (SIMÕES *et al.*, resultados não publicados, Anexo D).

REFERÊNCIAS BIBLIOGRÁFICAS*

ALEF, K.; NANNIPIERI, P. **Methods in applied soil microbiology and biochemistry**. London: Academic Press, 1998.

ALLAN, J. P.; ATEKWANA, E. A.; DURIS, J. W.; WERKEMA, D. D.; ROSSBACK, S. The Microbial Community Structure in Petroleum-Contaminated Sediments Corresponds to Geophysical Signatures. **Appl. Environ. Microbiol.**, v. 73, n. 09, p 2860–2870, 2007.

ALLEN, R. H.; JAKOBY, W. B. Tartaric acid metabolism. IX. Synthesis with tartrate epoxidase. **J. Biology. Chem.**, v. 244, n. 8, p. 2078–2084, 1969.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W.; LIPMAN, D. J. Basic local alignment search tool. **J. Mol. Biol.**, v. 215, n. 3. p. 403-410, 1990.

ALVES, L.; MESQUITA, E.; GIRIO, F. M. Desulfurização bacteriana de combustíveis fósseis. **Boletim da Sociedade Portuguesa de Biotecnologia**, n. 62, p, 6-8, 1999.

AMANN, R. I.; LUDWING, W.; SCHEIFER, K. H. PHYLOGENETIC. Identification *in situ* detection of individual microbial cells without cultivation. **Microbiology**, v. 59, p. 143-159, 1995.

ANDERSON, J. P. E. Soil Respiration. In: PAGE, A. L. L, (Ed.). **Agronomy Monograph**: Chemical and biological Properties, Madison: American Society of Agronomy and Soil Science Society, 1982.

APPLEGATE, B. M.; MATRUBUTHAM, U.; SANSERVERINO, J.; SAYLER, G. S. Biodegradation gene as marker genes in microbial ecosystems. **Mol. Microbiol. Ecol. Man.**, v.6, n. 18, p. 1-14, 1995.

ATLAS, R. M. Diversity of microbial communities. **Adv. Microbiol. Ecol.**, v. 7, p. 1-47, 1984.

* De acordo com:

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

ATLAS, R. M.; HOROWITZ, A.; KRICHEVSKY, M.; BEJ, A. K. Response of microbial populations to environmental disturbance. **Microbiol. Ecol.**, v. 22, p. 249-256, 1991.

BADALASSI, F.; NGUYER, H. K.; CROTTI, P.; REYMOND, J. L. **Helv. Chem. Acta**, v.85, p.3090-3098, 2002.

BADALASSI, F.; WAHLER, D.; KLEIN, G.; CROTTI, P.; REYMOND, J. L. **Ang. Chem. Int. Ed.**, v.39, 4067-4070, 2000.

BARNS, S. M.; FUNDYGA, R. E.; JEFFRIES, M. W.; PACE, N. R. Remarcable archeal diversity detected in Yellowstone National Park hot spring environment. **Proc. Natl. Acad. Sci.**, v. 91, p. 1609-1613, 1994.

BEDESSEM, M. E.; SWOBODACOLBERG, N. G.; COLBERG, P. J. S. Naphthalene mineralization coupled to sulfate reduction in aquifer-derived enrichments. **FEMS Microbiol. Lett.**, v. 152, p. 213–218, 1997.

BÉJÀ, O. To BAC or not to BAC: marine ecogenomics. **Curr. Opin. Biotechnol.**, v.15, p. 187–190, 2004.

BÉJÀ, O.; SUZUKI, M. T.; KOONIN, E. V.; ARAVIND, L.; HADD, A.; NGUYEN, L. P.; VILLACORTA, R.; AMJADI, M.; GARRIGUES, C.; JAVANOVIVICH, S. B.; FELDMAN, R. A.; DELONG, E.F. Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage. **Environ. Microbiol.**, v.62, p. 516-529, 2000.

BÍCEGO, M. C. **Contribuição ao estudo de hidrocarbonetos biogênicos e do petróleo no ambiente brasileiro**. 156 p. Dissertação (Mestrado) - Instituto Oceanográfico, Universidade de São Paulo, São Paulo, 1988.

BONCH-SMOLOVSKAYA, E.A., MIROSHNICHENKO, M.L., LEBEDINSKY, A.V., CHERNYH, N.A., NAZINA, T.N., IVOILOV, V.S., BELYAEV, S.S., BOULYGINA, E.S., LYSOV, Y.P., PEROV, A.N., MIRZABEKOV, A.D., HIPPE, H., STACKEBRANDT, E., L'HARIDON, S., JEANTHON, C., Radioisotopic, culture-based, and oligonucleotide microchip analyses of thermophilic microbial communities in a continental high-temperature petroleum reservoir. **Appl. Environ. Microbiol.**, v69, p. 6143-6151., 2003.

BRESLLER, D. C.; FEDORSCK, P. M. Identification of disulfides from the biodegradation of dibenzothiophene. **Appl. Environ. Microbiol.**, v.67, n.11, p. 5084-5093, 2001.

BURNS, R. G. Enzyme activity in soil: location and a possible role in microbial ecology. **Soil Biol. Biochem.**, v.14, n.3, p.423-427, 1982.

CASTRO, H. F. DE; MENDES, A. A.; SANTOS, J. C. DOS; AGUIAR, C L. Modificação de óleos e gorduras por biotransformação. **Quím. Nova**, v.27, n.1, p.146-156, 2004.

CECCANTI, B. Il ruolo dei lombrichi (*Eisenia foetida*) nella bioremediation di um suolo contaminato da idrocarburi. **Convegno SISS Qualità del suolo, impatto antropico e qualità dei prodotti agricoli**. Siena, 2003.

CERNIGLIA, C. E. Microbial metabolism of polycyclic aromatic hydrocarbons. **Adv. Appl. Microbiol.**, v. 30, p. 31–71, 1984.

CERNIGLIA, C. E.; FU, P. P.; YANG, S. K. Metabolism of 7-methylbenz[a]anthracene and 7-hydroxymethylbenz[a]anthracene by *Cunninghamella elegans*. **Appl. Environ. Microbiol.**, v.44, n.3, p.682-9, 1982.

CERNIGLIA, C. E.; GIBSON, D. T.; VAN BALEN, C. Oxidation of naphtalene by *Cyanobacteria* and *Microalgae*. **J. Microbiol.**, v.116, p. 495-500, 1980.

CHAPMAN, R. E. **Petroleum geology-Development in the petroleum science**. Amsterdam: Oxford, 1983.

CHEUNGA, K. H.; GU, J. D. Mechanism of hexavalent chromium detoxification by microorganisms and bioremediation application potential: A review International. **Biodeter. Biodegrad.**, v. 59, p. 8-15, 2007.

CONRAD, R.; WEBER, M.; SEILER, W. Kinetics and electron transport of soil hydrogenases catalyzing the oxidation of atmospheric hydrogen. **Soil Biol. Biochem.**, v. 15, p.167-176, 1983.

CONTI, R.; RODRIGUES, J. A. R.; MORAN, P. J. S. Biocatálise: Avanços recentes. **Quím. Nova**, v. 24, p. 672-675, 2001.

COTTRELL, M.; MOORE, J.A.; KIRCHMAN, D.L. Chitinases from Uncultured Marine Microorganisms, **Appl. Environ. Microbiol.**, v. 65, p. 2553 – 2557, 1999.

COURTOIS, S.; FROSTENGARD, A.; GORANSSON, P.; DEPRET, G.; JEANNIN, P.; SIMONET, P. Quantification of bacterial subgroups in soil: comparison of DNA extracted from soil or from cells previously released by density gradient centrifugation. **Environ. Microbiol.**, v. 3, p. 431-439, 2001.

CRESPIN, E. **Detecção e Quantificação de Bactérias Degradadoras de Hidrocarbonetos em Amostras de Petróleo Utilizando Primers Grupo específicos.** f.n. 112. Dissertação (Mestrado em Genética) – Universidade de Campinas, Campinas, 2008.

CURY, J. C. **Atividade microbiana e diversidades metabólica e genética em solos de mangue contaminado com petróleo.** n.f. 147. Dissertação (Mestrado em Agronomia) – Escola Superior de Agronomia Luiz de Queiroz, Piracicaba, 2002.

CURY J. C. **Diversidade de Bacteria e Archaea em solos de mangue e marisma.** Tese (Doutorado em Agronomia) - Escola Superior de Agronomia Luiz de Queiroz, Piracicaba, 2006.

DÍAZ, E. Bacterial degradation of aromatic pollutants: a paradigm of metabolic versatility. **Int. Microbiol.**, v. 7, p. 173-180, 2004.

DICK, R. P. Soil enzyme activities as indicators of soil quality. In: DORAN, J. W. et. al. (Ed.). **Defining soil quality for a sustainable environmental.** Madison: SSSA, 1994.

DICK, W. A., Tabatabai, M. A. Bioremediation of contaminated soils. **Agronomy monograph.** Soil Science Society of America. v. 35, p. 315-338, 1999.

DUARTE, G. F.; ROSADO, A. S.; SELDING, L.; ARAUJO, W.; VANELSAS, J. D. Analysis of bacterial community structure in sulfurous-oil-containing soils and detection of species carrying dibenzothiophene desulfurization (*dsz*) genes. **Appl. Environ. Microbiol.**, v.92, p. 1052-1062, 2001.

DWORKIN, M.; FOSTER, J. W. Studies on *Pseudomonas methanica* (Sbhngen) nov. comb. **J. Bacteriol.**, v. 72, p. 646-659, 1956.

ERIKSSON, M.; KA, J. O.; MOHN, W. W. Effects of Low Temperature and Freeze-Thaw Cycles on Hydrocarbon Biodegradation in Arctic Tundra Soil. **Appl. Environ. Microbiol.**, v. 67, p. 5107-5112, 2001.

Farinas, E.T.; Butler, T.; Arnold, F. Directed enzyme evolution. **Curr. Opin. Biotechnol.**, v.12, p. 545, 2001.

FERRIS, M. J.; MUYZER, G.; WARD, D. M. Denaturing gradient gel electrophoresis profiles of 16S rRNA-defined populations inhabiting a hot spring microbial mat community. **Appl. Environ. Microbiol.**, v. 62, p. 340-346, 1996.

FILHO, J. G. M.; MENEZES, T. R.; OLIVEIRA, E. A.; IEMMA, M. B., Caracterização da Contaminação por Petróleo e seus Derivados na Baía de Guanabara: Aplicação de Técnicas Organogeoquímicas e Organopetrográficas. **Anu. Inst. Geoc.**, Rio de Janeiro, v. 26, 2003.

FUJIWARA, S. J. Extremophiles: Developments of their special functions and potential resources. **J. Biosci. Bioenerg.**, v. 94, p. 518-525, 2002.

FURUKAWA, K. Molecular genetics and evolutionary relationship of PCB-degrading bacteria. **Biodegradation**, v. 5, p. 289-300, 1994.

G, K. A.; POGREBINSKY, O. S.; MRANCHKO, G. T.; XI, L.; MONTICELO, D. J.; SQUIR, C. H. Molecular mechanisms of biocatalytic desulfurization of fossil fuels. **Nat. Biotechnol.**, v.14, p. 1705-1709, 1996.

GALAN, B.; DIAZ, E.; GARCIA, J. L. Enhancing desulphurization by engineering a flavin reductase-encoding gene cassette in recombinant biocatalysts. **Environ. Microbiol.**, v.2, p. 687-694, 2000.

GILLESPIE, D. E.; BRADY, S. F.; BETTERMAN, A. D.; CIANCOTTO, N. P.; LILES, M. R.; RONDON, M. R.; CLARDY, J.; GOODMAN, R. M.; HALDESMAN, J. Isolation of high molecular weight DNA from soil for cloning into BAC vectors. **App. Environ. Microbiol.**, v. 68, 4301-4306, 2008.

GIOVANNONI, S.J.; BRITSCHGI, T. B.; MOYER, C.L.; FIELD, K.G. Genetic diversity in Sargasso Sea bacterioplankton. **Nature**, v.3, p. 60-63, 1990.

GODDAR, J. P.; Reymond, J. L. Recent advances in enzyme assays. **Trend. Biotechnol.**, v. 22, p. 363-370, 2004.

GODON, J.J.; ZUMSTEIN, E.; DABERT, P.; HABOUZIT, F.; MOLETTA, R. Molecular microbial diversity of an anaerobic digester as determined by small-subunit rDNA sequence analysis. **Appl. Environ. Microbiol.**, v. 63, p.2802-2813, 1997.

GOLDEMBERG, J. From physics to development strategies. **Ann. Rev. Energ. Environ.**, v. 23, p.123- 128 , 1998.

GOLDEMBERG, J.; VILLANUEVA, L.D. **Energia, Desenvolvimento e Meio Ambiente**. 2^a ed. São Paulo: Editora Universidade de São Paulo, 2003.

GRAY, K.A.; MRACHKO, G.T.; SQUIRES, C.H. Biosulfurization of fossil fuel. **Curr. Opin. microbiol.**, v. 6, p. 229-235, 2003.

GREGORICH, E. G.; CARTELERT, B. H. Towards a minimum data set to assess soil organic matter quality in agricultural soils. **Can. J. Soil. Sci.**, v. 74, p. 367-385, 1994.

GUPTA, R.; BEG, Q. K.; LORENZ, P. Bacterial alkaline proteases: molecular approaches and industrial applications. **Appl. Microbiol. Biotechnol.**, v. 59, p. 15-32, 2002.

GUTIÉRREZ, M. C.; SLEEGERS, A.; SIMPSON, H. D.; ALPHAND, V.; FURSTOSS, R. The first fluorogenic assay for detecting a Baeyer–Villigerase activity in microbial cells. **Organ. Biomol. Chem.**, v. 1, p. 3500-3506, 2003.

HAMANN, C.; HEGEMANN, J.; HILDEBRANDT, A. Detection of polycyclic aromatic hydrocarbon degradation genes in different soil bacteria by polymerase chain reaction and DNA hybridization. **FEMS Microbiol. Lett.**, V. 171, p. 255-263, 1999.

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

HARWOOD, C. S.; BURCHHARDT, G.; HERRMANN, H.; FUCHS, G. Anaerobic metabolism of aromatic compounds via the benzoyl-CoA pathway. **FEMS Microbiol. Rev.**, v.22, p. 439-458, 1999.

HASAN, F.; SHAH, A .; HAMEED, A . Industrial applications of lipases. **Enz. Microbiol. Technol.**, v.3, p. 235-251, 2006.

ISHII, Y.; KONISHI, J.; OKADA, H.; HIRASAWA, K.; ONAKA, T.; SUZUKI, M. Operon structure and functional analysis of the gene encoding thermophilic desulfurizing enzymes of *Panibacillus* sp. A11-2. **Biochem. Biophys. Res. Commun.**, v. 270, p. 81-88, 2000.

JACKSON, J. B. C.; CUBIT, J. D.; KELLER, B. D.; BATISTA, V.; BURNS, K.; CAFF E, Y. H. M.; CALDWELL, R. L.; GARRITY, S. D.; GETTER, C. D.; GONZALEZ, C.; GUZON, H. M.; KAUFMAN, K. W.; KNAP, A. H.; LEVINGS, S. C.; MARSHALL, M. J.; STEGER, R.; THOMPSON, R. C.; WEIL, E. Ecological effects of a major oil spill on Panamanian coastal marine communities. **Science**, v. 243, N. 4887, p. 37-44, 1989.

KANNALY, R. A.; HARAYAMA, S. Biodegradation of high-molecular-weight polycyclic aromatic hydrocarbons by bacteria. **J. Biotech.**, v. 182, p. 2059-2067, 2000.

KAPUSCINSKI, J. DAPI: a DNA-specific fluorescent probe. **Biotech. Hystochem.**, v.70, p. 220–233, 1995.

KENNISH, M. **Estuarine and Marine Pollution**. Boca Raton: CRC Press. 1997.

KENNISH, M. J. **Practical handbook of estuarine and marine pollution**. Boca Raton: CRC Press, 1996.

KEPNER, R. L.; PRATT, J. R. Jr. Use of fluorochromes for direct enumeration of total bacteria in environmental samples: past and present. **Microbiol. Rev.**, v. 58, p. 603-615, 1994.

KIM, U. J. **Caltech Genome Research Laboratory**: a collection of the experimental procedures. Disponível em: <http://informa.bio.caltech.edu/protocol_index.html>. Acesso em: 30 Out. 2008.

KIM, U. J.; BIRREN, B. W.; SLEPAK, T.; MANCINO, V.; BOYSEN, C.; KANG, H.-L.; SIMON, M. I.; SHIZUYA, H. Construction and characterization of a human bacterial artificial chromosome library. **Genomics**, v.34, p. 213-218, 1996.

KIM, Y. -J.; CHOI, G. -S.; KIM, S. -B.; YOON, G. S.; KIM, Y. -S.; RYU, Y. -W. Screening and characterization of a novel esterase from a metagenomic library. **Prot. Express. Purif.**, v. 45, p. 315-323, 2006.

KIMURA, M. A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. **J. Mol. Evol.**, v. 16, p. 111-120, 1980.

KIRCHMAN, D. L.; SIGDA, J.; KAPUSCINSKI, R.; MITCHELL, R. Statistical analysis of the direct count method for enumerating bacteria. **Appl. Environ. Microbiol.**, v. 44, p. 376-382, 1982.

KLEIN, G.; REYMOND, J. L. An enantioselective fluorimetric assay for alcohol dehydrogenases using albumin-catalyzed b-elimination of umbelliferone. **Biol. Med. Chem. Lett.**, v. 8, p. 1113–1116, 1998.

KLEIN, G.; REYMOND, J. L. Enantioselective fluorogenic assay of acetate hydrolysis for detecting lipase catalytic antibodies. **Helvet. Chim. Acta**, v.55 , p. 400–408, 1999.

KUMAR S.; TAMURA, K.; JAKOBSEN, I. B.; NEI M. MEGA2: molecular evolutionary genetics analysis software. **Bioinformatics**, v. 17, n.12, p.1244-5, 2001.

LAMB, D. C.; KELLY, D. E.; MASAPHY, S.; JONES, G. L.; KELLY, S. L. Engineering of heterologous cytochrome p450 in *Acinetobacter* sp.: application for pollutant degradation. **Biochem. Biophys. Res. Commun.**, v. 276, p. 797-802, 2000.

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

LANE, D. J. 16S/23S rRNA sequencing. *In*: GOODFELLOW, M.; STACKEBRANDT, E. (Ed.). **Nucleic acid techniques in bacterial systematics**. Chichester: John Wiley and Sons, 1991.

LANE, D. L.; PACE, B.; OLSEN, G. J.; STAHL, D. A.; SOGIN, M. L.; PACE, N. R. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. **Proc. Nat. Acad. Sci.**, v. 82, p. 6955-6959, 1985.

LANE, D. J. 16S/23S rRNA sequencing. In: GOODFELLOW, M.; STACKEBRANDT, E. (Ed.). **Nucleic acid techniques in bacterial systematic**. Chichester: John Wiley & Sons, 1991. p. 115-147.

LEAHY, J. G.; COLWELL, R. R. Microbial degradation of hydrocarbons in the environment. **Microbiol. Rev.**, v. 54, p. 305-315, 1990.

LIDSTROM, J. E.; BARRY, R. P.; BADDOCK, J. F. Long-Term effect on microbial communities after a subarctic oil spill. **Soil Biol. Biochem.**, v.31, p.1677-1689, 1999.

LILES, M. R.; WILLIAMSON, L. L.; RODBUMRER, J.; TORSVIK, V.; GOODMAN, R. M.; HANDELSMAN, J. Recovery, purification, and cloning of high-molecular-weight DNA from soil microorganisms. **Appl. Environ. Microbiol.**, v.74, p.3302-3305, 2008.

LIU, Y. Y.; XU, J.; HU, Y. Enhancing effect of Tween-80 on lipase performance in enantioselective hydrolysis of ketoprofen ester. **J. Mol. Cat. B: Enzim.**, v.10, p. 523-529, 2000.

LÓPEZ, L. A.; PUJALTE, M. J.; BENLLOCH, S.; MATA, R. M.; ROSSELLÓ, M. R.; GARAY, E.; RODRÍGUEZ, V. F. *Thalassospira lucentensis* gen. nov., sp. nov., a new marine member of the alpha-Proteobacteria. **Inter. J. System. Evol. Microbiol.**, v. 52, p.1277-1283, 2002.

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

LORENZ, P.; LIEBERTON, K.; NIEHAUS, F.; ECK, J. Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. **Curr. Opin. Microbiol.**, v. 13, p. 572-577, 2002.

LORENZ, P.; SCHLEPER, C. Metagenome: a challenging source of enzyme discovery. **J. Mol. Catal., B Enzym.**, v.102, p. 13-19. 2002.

LOY, A.; BEISKER, W.; MEIER, H. Diversity of bacteria growing in natural mineral water after bottling. **Appl. Environ. Microbiol.**, v. 71, p.3624-3632, 2005.

MAGOT, M.; OLLIVIER, B.; PATEL, B.K.C., Microbiology of Petroleum Reservoirs. **Ant. van Leeuw.**, v. 77, p. 103-116, 2000.

MAGOT, M., Similar bacteria in remote oil fields. **Nature**, v.379, p.681, 1996.

MANTOVANI, S.M. **Aplicação de Triagem de Alto Desempenho na Investigação das Atividades Enzimáticas e Enantiosseletividades de Microrganismos Brasileiros**. 102 f. Dissertação (Mestrado em Química Analítica) – Instituto de Química, Universidade de Campinas, Campinas, 2007.

MCFARLAND, B. L.; BORON, D. J.; DEEVER, W.; MEYER, J. A.; JOHNSON, A. R.; ATLAS, R. M. Biocatalytic sulfur removal from fuels: applicability for producing low sulfur gasoline. **Crit. Rev. Microbiol.**, v. 24, p. 99-147, 1998.

MECKENSTOCK, R. U.; ANNWEILER, E.; MICHAELIS, W.; RICHNOW, H. H.; SCHINK, B. Anaerobic naphthalene degradation by a sulfate-reducing enrichment culture. **Appl. Environ. Microbiol.**, v. 66, p. 2743-2747, 2000.

MELO, I. S.; AZEVEDO, J. L. **Microbiologia e Ambiental**. Jaguariuna: Editora Embrapa, 1997.

MELO, I. S.; SILVA, C. M. M. S.; SCRAMIN, S.; SPESSOTO, A. **Biodegradação**. Jaguariuna: Editora. Embrapa Meio Ambiente, 2001.

MOREIRA, F. M. S.; SIQUEIRA, J. O. **Microbiologia e bioquímica do solo**. Lavras: Editora UFLA, 1994.

MORGAN, P.; R. J. WATKINSON. Microbiological methods for the cleanup of soil and ground water contaminated with halogenated organic compounds. **FEMS Microbiol. Rev.**, v. 63, p. 277-300, 1989.

MOSQUEDA, G.; RAMOS-GONÇALES, M. I.; RAMOS, R. L. Toluene metabolism by the solvent-tolerant *Pseudomonas putida* DOT-T1 strain, and its role in solvent impermeabilization. **Gene**, v. 238, p. 69-76, 1999.

NAKATSU, C. H.; CARMOSINI, N.; BALDWIN, B.; BEASLEY, F.; KOURTEV, P.; KONOPKA, A. Soil microbial community responses to additions of organic carbon substrates and heavy metals (Pb and Cr). **Appl. Environ. Microbiol.**, v. 71, p. 7679-7689, 2005.

NAZINA, T.N., IVANOVA, A.E., MITYUSHINA, L.L., BELYAEV, S.S., Thermophilic hydrocarbon-oxidizing bacteria in oil strata. **Microbiology** (English Translation), v. 62, p. 359-365, 1993.

NIE, Y.; XU, Y.; YANG, M.; MU, X. Q. A novel NADH-dependent carbonyl reductase with unusual stereoselectivity for (*R*)-specific reduction from an (*S*)-1-phenyl-1,2-ethanediol-producing micro-organism: purification and characterization **Lett. Appl. Microbiol.**, v. 44, p. 55-562, 2006.

NIES, D. H. Microbial heavy metal resistance. **Appl. Microbiol. Biotechnol.**, v. 51, p. 730-750, 1999

OHSHIRO, T.; IZUMI, Y. Microbial desulfurization of organic sulfur compounds in petroleum. **Biosci. Biotechnol. Biochem.**, v. 63, p. 1-9, 1999.

OKERENTUNGBA, P. O.; EZERONYE, O. U. Petroleum degrading potencies of single and mixed microbial cultures isolated from rivers and refinery effluent in Nigeria. **Afr. J. Biotechnol.**, v. 2, p. 288-292, 2003.

OKUDO, Y.; MAEDA, M.; KUDO, T. *Pseudomonas putida* CE2010 can degrade biphenyl by a mosaic pathway encoded by the *tod* operon and *cmtE*, which are identical to those of *P. putida* F1 except for a single base difference in the operator–promoter region of the *cmt* operon. **Microbiology**, v. 147, p. 31-41, 2001.

ORPHAN, V. J.; TAYLOR, L. T.; HAFENBRADL, D.; DELONG, E. F. Culture dependent and culture independent characterization of microbial assemblages associated with high-temperature petroleum reservoirs. **Appl. Environ. Microbiol.**, v. 66, p. 700-711, 2000.

OSBURNE, M. S.; GROSSMAN, T. H.; AUGUST, P. R.; MACNEIL, L. A. Tapping into microbial diversity for natural products drug discovery. **ASM News**, v.66, p. 122- 127, 2000.

PACE, N. R. New perspective on the natural microbial world: molecular microbial ecology. **ASM News**, v. 62, p. 463-470, 1996.

PARALES, R. E.; BRUCE, N. C.; SCHMID, A.; WACKETT, L. B. Biodegradation, biotransformation, and biocatalysis (B3). **Appl. Environ. Microbiol.**, v. 68, p. 4699 – 4709, 2002.

PARK, H. J.; JEON, J. H.; KANG, S. G.; LEE, J. H.; LEE, S. A.; KIM, H. K. Functional expression and refolding of new alkaline esterase, EM2L8 from deep-sea sediment metagenome. **Protein Expr. Purif.**, v.52, 340-347,2007.

PARKINSON, D.; COLEMAN, D. C. Methods for assessing soil microbial populations, activity and biomass – Microbial communities, activity and biomass. **Agr. Ecosystems. Environ.**, v. 34, p. 3-33, 1991.

PAUL, E. A.; CLARCK, F. E. **Soil microbiology and biochemistry**. San Diego: Academic Press, 1996. p. 340.

PÉREZ-JIMENES, J. R.; YONG, L. Y.; KERKOF, L. J. The molecular characterization of sulfate reducing bacteria in anaerobic hydrocarbon degrading consortia and pure cultures using the dissimilatory sulfite reductase (*dsrAB*) genes. **FEMS Microbiol. Ecol.**, v. 35, p. 145-150, 2001.

PIZA, F. F.; PRADO, P. I.; MANFIO, G. P. Investigation of bacterial diversity in Brazilian tropical estuarine sediments reveals high actinobacterial diversity. **Anton. van Leeuw.**, v. 86, p. 328, 2004.

PRITCHARD, P. H.; MUELLER, J.G.; LANTZ, S.E.; SANTAVY, D.L. The potential importance of biodiversity in environmental biotechnology applications: bioremediation of PAH-contaminated soils and sediments. In: ALLSOPP D, COLWELL, R. R.; HAWKSWORTH, D. L. (Ed.). **Microbial diversity and ecosystem function**. Wallingford UK: CABI, 1995. p. 161–82.

RADWAN, S.S., SORKHOH, N.A., FARDOUN, F., AL-ASAN, R.H. Soil management enhancing hydrocarbon biodegradation in the polluted Kuwaiti desert. **Appl. Microbiol. Biotechnol.**, v. 44, p. 265-270, 1995.

RESOLUÇÃO Nº 07/2007. **Programa de Revitalização, Urbanização e Recuperação Ambiental dos Canais do Cunha e do Fundão**, Rio de Janeiro: UFRJ, 2007.

REYMOND, J. L. **Enzyme assays: High-Throughput Screening**, genetic selection and fingerprint. Suíça: Wiley-VCH, 2006.

REYMOND, J-L. New High-Throughput Screening Assays for Biocatalysis. **Chimia**, v.55, p. 1049-1052, 2001.

ROBERTSON, D. E.; CHAPLIN, J. A.; DESANTIS, G.; PODAR, M.; MADDEN, M.; CHI, E.; RICHARDSON, T.; MILAN, A.; MILLER, M.; WEINER, D. P.; WONG, K.; MCQUAID, J.; FARWELL, B.; PRESTON, L. A.; TAN, X.; SNEAD, M. A.; KELLER, M.; MATHUR, E.; KRETZ, P. L.; BURK, M. J.; SHORT, J. M. Exploring nitrilase sequence space for enantioselective catalysis. **Appl. Environ. Microbiol.**, v.70, p. 2429-2436, 2004.

ROLIN, W. F. M.; HEAD, I. M.; LARTER, S. R. The microbiology of hydrocarbon degradation in subsurface petroleum reservoirs: perspectives and prospects. **Res. Microbiol.**, v. 154, p. 321-328, 2003.

RÖLING, W. F. M.; HEAD, I. M.; LARTER, S. R. The microbiology of hydrocarbon degradation in subsurface petroleum reservoirs: perspectives and prospects. **Res. Microbiol.**, v.154, p.3 21-328, 2003.

RONDON, M. R.; GOODMAN, R. M.; HANDELSMAN, J. The Earth's bounty: assessing and accessing soil microbial diversity. **Trends Biotechnol.**, v. 17, p. 403-409, 1999.

RONDON, M. R.; RAFFEL, S. J.; GOODMAN, R. M.; HANDELSMAN, J. Toward functional genomics in bacteria: Analysis of gene expression in *Escherichia coli* from a bacterial artificial chromosome library of *Bacillus cereus*. **Proc. Nat. Acad. Sci.**, v. 96, p. 6451-6455, 1999.

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

RUSSEL, N. J. Alteration in fatty acid chain length in *Micrococcus coprophilus* grown at different temperatures. **Biochem. Biophys. Acta**, v. 231, p. 254-256, 1971.

SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Mol. Biol. Evol.**, v.4, p. 425,1987.

SCHWARTZ, R.D.; LEATHEN, W.W. **Petroleum microbiology**. New York: Industrial Microbiology, 1976.

SEBAT, J. L.; COLWELL, F. S.; CRAWFORD, R. L. Metagenomic Profiling: Microarray Analysis of an Environmental Genomic Library. **Appl. Environ. Microbiol.**, v.69, p.4927-4934, 2003.

SHIZUYA, H.; BIRREN, B.; KIM, U. J.; MANCINO, V.; SLEPAK, T.; TACHIRI, Y.; SIMON, M. Cloning and stable maintenance of 300-kilobase-pair fragments

of human DNA in *Escherichia coli* using an F-factor-based vector. **Proc. Nat. Acad. Sci.**, v. 89, p. 8794-8797, 1992.

SMITS, T. H. M.; BALADA, S. B.; WITHOLT, B.; VAN BEILEN, J. B. Functional analysis of alkane hydroxylases from gram-negative and gram-positive bacteria. **J. Bacteriol.**, v.105 , p. 1733 –1742, 2002.

SPORMANN, A. M.; WIDDEL, F. Metabolism of alkylbenzenes, alkanes, and other hydrocarbons in anaerobic bacteria. **Biodegradation**, v. 11, p. 85–105, 2000.

STALEY, J. T.; KONOPKA, A. Measurement of *in situ* activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. **Ann. Rev. Microbiol.**, v.39, p. 321-46, 1985.

STEIN, J. L.; MARSH, T. L.; WU, K. Y.; SHIZUYA, H.; DELONG, E. F. Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a plancktonic marine archaeon. **J. Bacteriol.**, v. 178, p. 591-599, 1996.

SWAVING, J.; BONT, DE J.; WESTPHAL, A.; KOK, A. A novel type of pyridine nucleotide-disulfide oxidoreductase is essential for NAD⁺- and NADPH-dependent degradation of epoxyalkanes by *Xanthobacter* strain Py2. **J. Bacteriol.**, v. 178, p. 6644–6646 ,1996

THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. **Nucleic Acids Res.**, v. 22, n. 4673-4680, 1994.

TORSVIK, V.; GOKSOYR, J.; DAAE, F.L. High diversity in DNA of soil bacteria. **Appl. Environ. Microbiol.**, v. 56, p. 782-787, 1990.

TYSON, G. W.; CHAPMAN, J.; HUGENHOLTZ, P.; ALLEN, E. E.; RAM, R. J.; RICHARDSON, P. M.; SOLOVYEV, V. V.; RUBIN, E. M.; ROKHSAR, D. S.; BANFIELD, J. F. Community structure and metabolism through reconstruction of microbial genomes from the environment. **Nature**, v. 428, p. 37-43, 2004.

VAN BAILEN, J. B.; PANKE, S.; LUCHINNI, S.; FRANCHINI, A. G.; RÖTLISBERGER, M.; WITHOLT, B. Analysis of *Pseudomonas putida* alkane-degradation gene clusters and flanking insertion sequences: evolution and regulation of the *alk* genes. **Microbiology**, v.147, p.1621-1630, 2001.

VAN BAILEN, J. B.; SMITS, T. H. M.; WHYTE, E. L. G.; SCHORCHT, S.; RÖTLISBERGER, M.; PLAGGEMEIER, T.; ENGESSER, K. H.; WITHOLT, B. Alkane hydroxylases homologues in Gram positive strains. **Environ. Microbiol.**, v. 4, p. 676- 682, 2002.

VAN HELLEMOND, E. W.; JANSSEN, D. B.; FRAAIJE, M. W. Discovery of a thermostable Baeyer–Villiger monooxygenase by genome mining. **Appl. Environ. Microbiol.**, v. 90, p. 5832-5839, 2007.

VAN HEMME, J. D.; SINGH, A.; WARD, O. P. Recent advances in petroleum microbiology. **Microbiol. Mol. Biol. Rev.**, v.9, p. 503-549, 2003.

VASCONCELLOS, S. P. **Atividades Enzimáticas e de Biodegradação de Microrganismos do Petróleo da Bacia de Campos (Pampo Sul)**. n.f. 197, Tese (Doutorado em Química Analítica) – Instituto de Química, Universidade de Campinas, Campinas, 2006.

VAZOLLER, R. F. **Avaliação do ecossistema microbiano de um biodigestor anaeróbio de fluxo ascendente e manta de lodo, operando com vinhaça sob condições termofílicas**. 145 f. Tese (Doutorado em Engenharia Biotecnológica) – Universidade de São Paulo, São Carlos, 1995.

VENTER, J. C.; REMINGTON, K.; HEIDELBERG, J. F.; HALPERN, A. L.; JONATHAN RUSCH, D.; EISEN, J. A.; WU, D.; PAULSEN, I.; NELSON, K. E.; NELSON, W.; FOUTS, D. E.; LEVY, S.; KNAP, A. H.; LOMAS, M. W.; NEALSON, K.; WHITE, O.; PETERSON, J.; HOFFMAN, J.; PARSONS, R.; BADEN-TILLSON, H.; PFANNKUCH, C.; ROGERS, Y. -H.; SMITH, H. O. Environmental genome shotgun sequencing of the Sargasso Sea. **Science**, v. 304, p. 66-74, 2004.

VOLKMAN, J. K. Sterols and other triterpenoids: source specificity and evolution of biosynthetic pathways. **Org. Geochem.**, v. 36, p. 139-159, 2005.

WALKER, D. J.; COLWELL, R. R. Enumeration of petroleum-degrading microorganisms. **Appl. Environ. Microbiol.**, v. 31, p. 198-207, 1976.

WANG, C. Extreme tolerance to cadmium and high resistance to copper, nickel and zinc in different *Acidiphilium* strains **Appl. Environ. Microbiol.**, v. 72, p. 384-391, 2007.

WARD, D. M.; WELLER, R.; BATENSON, M. M. 16S rRNA sequences reveal numerous uncultivated microorganisms in a natural environment. **Nature**, v. 345, p. 63-65, 1990a.

WARD, D. M.; Weller, R.; Batenson, M. M. 16S rRNA sequences reveal uncultured inhabitants of well-studied thermal community. **FEMS Microbiol. Rev.**, v. 75, p. 105-116, 1990b.

WATKINS, L. M.; RODRIGUEZ, R.; SCHNEIDER, D.; BRODERICK, R.; CRUZ, M.; CHAMBERS, R.; RUCKMAN, E.; CODY, M.; MRACHKO, G.T. Purification and characterization of the aromatic desulfurinase 2-(2'-hydroxyphenyl) benzenesulfinate desulfinase. **Arch. Biochem. Biophys.**, v. 415, p. 14 -23, 2003.

WHYTE, L. G.; GREER, C. W.; INNIS, W. E. Assessment of the biodegradation potential of psychrotrophic microorganisms. **Can. J. Microbiol.**, v.42, p. 99–106, 1996.

WILD, J.; HRADECNA, Z.; SZYBALSKI, W. Conditionally amplifiable BACs: Switching from single-copy to high-copy vectors and genomic clones. **Genome Res.**, v. 12, p. 1434-44, 2002.

WSTKINSON, R. J.; MORGAN, P. Physiology of aliphatic hydrocarbon-degrading microorganisms. **Biodegradation**, v.1, p. 79-92, 1990.

YOUNG, A.; BLAKESLEY, R. Sequencing plasmids from single colonies with the dsDNA cycle sequencing system. **Focus**, v.13, p. 13,1991.

ZHOU, J.; BRUNS, M. A.; TIEDJE, J. M. DNA recovery from soils of diverse composition. **Appl. Environ. Microbiol.**, v.62, p. 316-22, 1996.

ZHU, Q.; HALFON, M. S. Vector-dependent gene expression driven by insulated p-element reporter vectors. **Fly Austin**, v.15 , p. 55-6, 2007

ZYLSTRA, G. J.; GIBSON, D. T. Aromatic hydrocarbon degradation: a molecular approach. **Genetic Engineering**, v. 13, p. 183-203, 1991.

ZYLSTRA, G. J.; GIBSON, D. T. Toluene degradation by *Pseudomonas putida* P1: nucleotide sequence of the *todC1C2BADE* genes and their expression in *Escherichia coli*. **J. Biol. Chem.**, v. 264, p. 14940-14946, 1989.

ANEXO A – SEQUÊNCIAS DE RDNA 16S OBTIDAS A PARTIR DE ISOLADOS DE ENRIQUECIMENTOS DE SEDIMENTOS DE MANGUE

>isob1

CAAGTCGAACGATGAAGCCAGCTTGCTGGGTGGATTAGTGGCGAACGGGTGAGTAACACGTGGGTGATCTGCCCTGC
ACTCTGGGATAAGCCTGGGAACTGGGTCTAATACCGGATATGACCTCGGGATGCATGTTCTGGGGTGGAAAGTTTTT
CGGTGCAGGATGAGCCCCGGCCTATCAGCTTGTGGTGGGGTAATGGCCTACCAAGGCGACGACGGGTAGCCGGCCT
GAGAGGGCGACCGGCCACACTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAA
TGGGCGCAAGCCTGATGCAGCGACGCCGCTGAGGGATGACGGCCTTCGGGTGTAAACCTCTTTCACCCATGACGAA
GCGCAAGTGACGGTAGTGGGAGAAGAAGCACCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGCAGCG
TTGTCCGGAATTACTGGGCGTAAAGAGCTCGTAGGCGGTTTGTGCGTCTGTGTGAAATCCCGCAGCTCAACTGCGG
GCTTGCAGGCGATACGGGCAGACTCGAGTACTGCAGGGGAGACTGGAATTCCTGGTGTAGCGGTGAAATGCGCAGATA
TCAGGAGGAACACCGGTGGCGAAGGCGGGTCTCTGGGCAGTAACTGACGCTGAGGAGCGAAAGCGTGGGTAGCGAACA
GGATTAGATACCTGGTAGTCCACGCCGTAAACGGTGGGCGCTAGGTGTGGGTTTCTTCCACGGGATCCGTGCCGTA
GCCAACGCATTAAGCGCCCCGCTGGGGAGTACGGCCGCAAGGCTAAAACCTCAAAGGAATTGACGGGGGCCCGCACAA
GCGGCGGAGCATGTGGATTAATTCGATGCAACGCGAAGAACCTTACCTGGGTTTGACATGTACCGGACGACTGCAGAG
ATGTGGTTTCCCTTGTGGCCGGTAGACAGGTGGTGCATGGCTGTGTCAGCTCGTGTGTCGTGAGATGTTGGGTTAAGTC
CCGCAACGAGCGCAACCCTTGTCTGTGTTGCCAGCAGTATGGTGGGGACTCGCAGGAGACTGCCGGGGTCAACTC
GGAGGAAGGTGGGACGACGTCAAGTCATCATGCCCCCTTATGTCCAGGGCTTCACACATGCTACAATGGTCGTACAG
AGGGCTGCGATACCGTGAGGTGGAGCGAATCCCTTANAGCCGGTCTCAGTTCGGATCGGGGTCTGCAACTCGACCCG
TGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCAACGCTGCGGTGAATACGTT

>isob2

ATGCAAGTCGAACGATGAAGCCAGCTTGCTGGGTGGATTAGTGGCGAACGGGTGAGTAACACGTGGGTGATCTGCC
TGCACTCTGGGATAAGCCTGGGAACTGGGTCTAATACCGGATATGACCTCGGGATGCATGTTCTGGGGTGGAAAGTT
TTTCGGTGCAGGATGAGCCCCGGCCTATCAGCTTGTGGTGGGGTAATGGCCTACCAAGGCGACGACGGGTAGCCGG
CCTGAGAGGGCGACCGGCCACACTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCA
CAATGGGCGCAAGCCTGATGCAGCGACGCCGCTGAGGGATGACGGCCTTCGGGTGTAAACCTCTTTCACCCATGAC
GAAGCGCAAGTGACGGTAGTGGGAGAAGAAGCACCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGCGA
GCGTTGTCCGGAATTACTGGGCGTAAAGAGCTCGTAGGCGGTTTGTGCGTCTGTGTGAAATCCCGCAGCTCAACTC
CGGGTTGACAGACTCAGCGGACAGCTCGAGTACTGCAGGGGAGACTGGAATTCCTGGTGTAGCGGTGAAATGCGCAG
ATATCAGGAGGAACACCGGTGGCGAAGGCGGGTCTCTGGGCAGTAACTGACGCTGAGGAGCGAAAGCGTGGGTAGCGA
ACAGGATTAGATACCTGGTAGTCCACGCCGTAAACGGTGGGCGCTAGGTGTGGGTTTCTTCCACGGGATCCGTGCC
GTAGCCAACGCATTAAGCGCCCCGCTGGGGAGTACGGCCGCAAGGCTAAAACCTCAAAGGAATTGACGGGGGCCCGCA
CAAGCGGCGGAGCATGTGGATTAATTCGATGCAACGCGAAGAACCTTACCTGGGTTTGACATGTACCGGACGACTGCA
GAGATGTGGTTTCCCTTGTGGCCGGTAGACAGGTGGTGCATGGCTGTGTCGTGAGATGTTGGGTTAA
GTCCCGCAACGAGCGCAACCCTTGTCTGTGTTGCCAGCAGTGTGGTGGGGACTCGCAGGAGACTGCCGGGGTCAA
CTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCCTTATGTCCAGGGCTTCACACATGCTACAATGGTCGGTA
CAGAGGGCTGCGATACCGTGAGGTGGAGCGAATCCCTTAAAGCCGGTCTCAGTTCGGATCGGGGTCTGCAACTCGACC
CCGTGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCAA

>isol6e

GAGCTAATACCGGATAATGTCTTCGGACCAAAGATTTATCGCCTTTGGATGAGCCCGCGTAAGATTAGCTAGTTGGTA
GGGTAAAGGCCCTACCAAGGCGACGATCTTTAGCTGGTCTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCC
AGACTCCTACGGGAGGCAGCATGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCAATGCCGCGTGAGTGATG
AAGGCCCTTCGGGTGCTAAAGCTCTTTTACCAGGGATGATAATGACAGTACCTGGAGAATAAGCTCCGGCTAACTCCGT
GCCAGCAGCCGCGGTAATACGGAGGGAGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCTAGGCGGTCTACC
AAGTCAGGGGTGAAATCCCGGGGCTCAACCCCGGAAGTGCCTTGAAACTATTAGACTAGAACTCTTGAGAGGTCAGT
GGAATTCGAGTGATAGAGGTGAAATTCGTAGATATTCCGAAGAACACCAAGTGGCGAAGGCGACTGACTGGACAAGTAT
TGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCTGGTAGTCCACGCCGTAAACGATGATAACTA
GCTGTCCGGGCACATGGTGTCTGGGTGGCGCAGCTAACGCATTAAGTTATCCGCCTGGGGAGTACGGTCGCAAGATTA
AAACTCAAAGGAATTGACGGGGCCTGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCCTTAC
CAGCCTTTGACATCCCGGTGCGGATTACCAGAGATGGTTTTCTTCAGTTCGGCTGGACCGGTGACAGGTGCTGCATGG
CTGTGCTCAGCTCGTGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTCATCCTTAGTTGCTACCATTT
AGTTGGGCACTCTAAGGAACTGCCGGTGATAAGCCGGAGGAAGGTGGGGATGACGTCAAGTCTCATGGCCCTTACA
GGCTGGGCTACACACGTGCTACAATGGCATCTACAGTGGGCACTATCACGCGAGTGTGAGCTAATCTCCAAAAGATG
TCTCAGTTCGGATTGTTCTCTGCAACTCGAGAGCATGAAGCGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGG

>isol9b

TGAGTAACGCGTGGGGACCTACCTCTTAGTGGGGGATAACGGTTGGAAACGACCGCTAATACCGCATACGCCCTTCGG
GGGAAAGATTTATCGCTAAGAGATGGACCCGCGTTGGATTAGATAGTTGGTGGGTGAGGTAACGGCTCACCAAGTCAGCGAT
CCATAGCTGGTTTGGAGGATGATCAGCCACACTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGGG

GAATATTGGACAATGGGGGCAACCCTGATCCAGCCATGCCCGGTGAGTGAAGAAGGCCTTCGGGTGTAAAGCTCTTT
CAGATGCGAAGATGATGACGGTAACATCAGAAGAAGCCCCGGCTAATTTCTGTGCCAGCAGCCCGGTAATACGAAAGG
GGTAGCGTGTTCGGATTACTGGGCGTAAAGGGCACGCAGGCGGTCTTGCCAGTCAGGGGTGAAAGCCCCGGGGCTC
AACC CGGAAGCTGCTGATCTGCAAGACTAGAGACTAGGAGAGGGTGGTGGAAATCCAGTGTAGAGGTGAAATT
CGTAGATATTGGGAGGAACACCAGAGGCGAAGGCGGCCACCTGGACTAGATCTGACGCTCAGGTGCGAAAGCGTGGGG
AGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGTAGTTGTTCGGGATTTCGATTTCGGTGAC
GCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGGTCGCAAGATTAAACTCAAAGGAATTGACGGGGGGCCGCA
CAAGCGGTGGAGCATGTGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATCCCTATCGCGATTTC
AGAGATG

>iso22c

GAGTGGCGAACGGGTGAGTAACACGTGGGTGATCTGCCCTGCACTTTGGGATAAGCCTGGGAAACTGGGTCTAATACC
GAATAGGATTCTGGACTGCATGGTCGGGGGTGGAAGCTTTTGGGTGTGGGATGGGCCCCGGGCTATCAGCTTGTT
GGTGGGGTGTGACCTACCAAGGCGACGACGGGTAGCCGGCTGAGAGGGTGTCCGGCCACACTGGGACTGAGATACG
GCCAGACTCCTACGGGAGGCGACAGTGGGGAATATTGCACAATGGGCGCAAGCCTGATGACGCGACGCCGCGTGAGG
GATGACGGCCTTCGGGTGTAAACCTCTTTGCCAGGGACGAAGCGCAAGTGACGGTACCTGGAGAAGAAGCACCGGC
CAACTACGTGCCAGCAGCCGCGTAATACGTAGGGTGCGAGCGTTGTCCGGAATTACTGGGCGTAAAGAGCTCGTAGG
TGGTTTGTGCGGTGTTCGTGAAAACCTCACAGCTTAACGTGTGGGCGTGCGGGCGATACGGGCGAGCTGGAGTACTGCA
GGGGAGACTGGAATTCCTGGTGTAGCGGTGGAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGGCGGGTCTCTG
GGCAGTAACTGACGCTGAGGAGCGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGG
TGGGTACTAGGTGTGGGTTCCTTCCTTGGGATCCGTGCCGTAGCTAACGCATTAAGTACCCCGCTGGGGAGTACGG
CCGCAAGGCTAAAACCTCAAAGGAATTGACGGGGGCCCGCACAAAGCGCGGAGCATGTGGATTAATTCGATGCAACGCG
AAGAACCTTACCTGGGTTTGACATGCACAGGACGCCGCGCAGAGATGTTCGGTTCCCTTGTGGCCTGTGTGCAGGTGGTG
CATGGCTGTGCTGACGCTGCTGCTGAGATGTTGGTTAAAGTCCCGCAACGAGCGCAACCCCTGTCTCATGTTGCCAG
CACGTTATGGTGTGACTCGTGAGAGACTGCCGGGGTCAACTCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGCC
CCTTATGTCCAGGACTTCACACATGCTACAATGGCCGGTACAAAGGGTTCGATGCCGTGAGGTGGAGCGAATCCTTT
CAAAGCCGTCTCAGTTCGGATCGGGGTCTGCAACTCGAACCCGTGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCA
ACGCTGCGGTGAA

>iso22f

CGGGATAGTGGCAGACGGGTGAGTAACGCGTGGGAATCTACCTTTTGGTACGGAATAGCTCCGGGAAACTGGAATTAA
TACCGTATGTGCTTTCGGGAGAAAGATTTATCGCCAAGAGATGAGCCCGCGTTTGATTAGGTAGTTGGTGGGTAAT
GGCCTACCAAGCCGACGATCAATAGCTGGTCTGAGAGGATGATCAGCCACATGGGACTGAGACACGGCCCAAACCTCC
TACGGGAGGCGACAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGAGTGATGAAGGCC
TAGGGTTGTAAAGCTCTTTCGCCGGTGAAGATAATGACTGTAGCCGGTAAAGAAGCCCCGGCTAACTTCGTGCCAGCA
GCCGCGTAATACGAAGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACATTTAAGTCAG
GGGTGAAATCCCGGAGCTCAACTCCGGAACGCTTTGATAGTGGGTGTCTCGAGTATGGAAGAGGTAAAGTGAATTG
CGAGTGTAGAGGTGAAATTCGTAGATATTTCGAGGAACACCAGTGGCGAAGGCGGCTTACTGGTCCATTACTGACGCT
GAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGATAGCTAGCCGTGCG
GTGTGCATGCATATCGGTGGCGCAGTTAACGCATTAAGCTATCCGCGTGGGGAGTACGGTCGCAAGATTAAACTCAA
AGGAATTGACGGGGGGCCCGCACAAAGCGGTGGAGCATGTGGTTAATTGCAAGCAACGCGCAGAACCTTACCAGCCCTT
GACATTTGGTGCTACCTCGAGAGATCGAGGGTTCCTTCGGGACGCCAGGACAGGTGCTGCATGGCTGTGCTCAGCTCG
TGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTCGCCCTTAGTTGCCAGCATTAGTTGGGCACTCTA
AGGGGACTGCCAGTGATAAGCTGGAGGAAGGTGGGGATGACGTCAAGTCCTCATGGCCCTTACGGGCTGGGCTACACA
CGTGCTACAATGGTGGTGACAGTGGGCAGCGAGCACGCGAGTGTGAGCTAATCTCCAAAAGCCATCTCAGTTCGGATT
GCACTCTGCAACTCGGGTGCATGAAGTTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGG
GCC

>iso27b

GGAGAGTGGCGCACGGGTGAGTAACGCGTGGGGACCTACCTCTTAGTGGGGGATAACGGTTGGAACGACCGCTAATA
CCGCATACGCCCTTCGGGGGAAAGATTTATCGCTAAGAGATGGACCCGCGTTGGATTAGATAGTTGGTGAGGTAACGG
CTCACCAAGTCAGCGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCAGACTCCTA
CGGGAGGCGACAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGAGTGAAGAAGGCCTTC
GGGTTGTAAAGCTCTTTCAGATGCGAAGATGATGACGGTAACATCAGAAGAAGCCCCGGCTAATTTCTGTGCCAGCAGC
CGCGGTAATACGAAAGGGGCAAGCGTTGTTTCGGATTACTGGGCGTAAAGGGCACGCAGGCGGTCTTGCCAGTCAGGG
GTGAAAGCCCCGGGCTCAACCCGGAACCTGCTGATAGTCTGAGGACTAGAGACTAGGAGAGGGTGGTGGAAATCCC
AGTGATAGAGGTGAAATTCGTAGATATTGGGAGGAACACCAGAGGCGAAGGCGGCCACCTGGACTAGATCTGACGCTCA
GGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAGTTGTTCGGG
ATTTTCGATTTCGGTGACGCAGCTAACGCATTAAGCACTCCGCGTGGGGAGTACGGTCGCAAGATTAAACTCAAAGGA
ATTGACGGGGGGCCCGCACAAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACA
TCCCTATCGCGAATTCAGAGATGGATATCATCAGTTCGGCTGGATAGGTGACAGGTGCTGCATGGCTGTGCTCAGCT
CGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTGTTCCCTAGTTGCCAGCATTAGTTGGGCACTC
TAGGGAGACTCGCGGTGACAAGCCGAGGAAGGCGGGGATGACGTCAAGTCCTCATGGCCCTTACGGGCTGGGCTACA
CACGTGCTACAAATGGTAACTACAGAGGGCAGCGACTTAGCGATAAGTAGCCAATCCAAAAAGTTATCTCAGTTTCG

>iso5

TGAGTAATGCCTAGGAATCTGCCTGGTAGTGGGGGATAACGTCCGGAAACGGGCGCTAATACCCGATACGTCCTACGG
GAGAAAGCAGGGGACCTTCGGGCGCTTGCCTATCAGATGAGCCTAGGTTCGGATTAGCTAGTTGGTGAGGTAATGGCTC

ACCAAGGCGACGATCCGTAACCTGGTCTGAGAGGATGATCAGTCACACTGGAAGTGAAGACACGGTCCAGACTCCTACGG
GAGGCAGCAGTGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGAAGGTCTTCGGA
TTGTAAAGCACTTTAAGTTGGGAGGAAGGACAGTAAGTTAATACCTTGCTGTTTTGACGTTACCGACAGAATAAGCAC
CGGCTAACCTTCGTGCCAGCAGCCGCGGTAATACGAAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCGCG
TAGGTGGTTTCGTAAAGTTGGATGTGAAAGCCCCGGGCTCAACCTGGGAAGTGCATCCAAAAGTGGCGAGCTAGAGTAC
GGTAGAGGGTGGTGAATTTCTGTGTAGCGGTGAAATGCGTAGATATAGGAAGGAACACCAGTGGCGAAGGCGACCA
CCTGGACTGATACTGACACTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAA
ACGATGTCAACTAGCCGTTGGAGTCCTTGAGATTTTAGTGGCGCAGCTAACGCATTAAGTTGACCGCCTGGGGAGTAC
GGCCGCAAGGTTAAACTCAAATGAATTGACGGGGGGCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGC
GAAGAACCCTTACCTGGCCTTGACATGCTGAGAACTTTCCAGAGATGGATTGGTGCCTTCGGGAAGTACAGACACAGGTG
CTGCATGGCTGTGCTAGCTAGCTGCTGCTGAGATGTTGGGTAAAGTCCCGTAACGAGCGCAACCCTTGTCCTTAGTTAC
CAGCACGTTATGGTGGGCACTCTAAGGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCATCAT
GGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGTTCGGTACAAAGGGTTGCCAAGCCGCGAGGTGGAGCTAATCC
CATAAAACCGATCGTAGTCCGGATCGCAGTCTGCAACTCGACTGCGTGAAGTCGGAATCGCTAGTAATCGTGAATCA

>iso6

TAGGAATCTGCCTAGTAGTGGGGGACAACGTTTCGAAAGGAACGCTAATACCGCATACGTCCTACGGGAGAAAGTGGG
GGATCTTCGGACCTCACGCTATTAGATGAGCCTAGGTCGGATTAGCTAGTTGGTAGGGTAAAGGCCCTACCAAGCGGAC
GATCCGTAACCTGGTCTGAGAGGATGATCAGTCACACTGGAAGTGAAGACACGGTCCAGACTCCTACGGGAGGCGAGCAGT
GGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGAAGGCCCTTCGGGTTCGTAAAGCAC
TTTAAGTTGGGAGGAAGGCTCATAGCGAATACCTGTGAGTTTTCGACGTTACCAACAGAATAAGCACCGGCTAACTTC
GTGCCAGCAGCCGCGGTAATACGAAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGGTGGCTTG
ATAAGTTGGATGTGAAATCCCCGGGCTCAACCTGGGAAGTGCATCCAAAAGTGTCTGGCTAGAGTGGGTAGAGGGTA
GTGGAATTTCCAGTGTAGCGGTGAAATGCGTAGATATTGGAAGGAACACCAGTGGCGAAGGCGACTACCTGGACTGAC
TCTGACTCTGATGTGCGAAAGCGTGGGGATCTAACAGGATTAGATACCCTGGTAGTCCACTCCGTAAACGATGTCAAC
TAGCCGTTGGGATCCTTGAGATCTTAGTGGCGCAGCTAACGCATTAAGTTGACCGCCTGGGGAGTACGGCCGCAAGGT
TAAAACCTCAAATGAATTGACGGGGGGCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTT
ACCTGGCCTTGACATGCTGAGAACTTTCCAGAGATGGATTGGTGCCTTCGGGAAGTACAGACACAGGTGCTGCATGGCT
GTCGTGAGTCTGCTGCTGAGATGTTGGGTAAAGTCCCGTAACGAGCGCAACCCTTGTCCTTAGTTACCAGCACGTTA
TGGTGGGCACTCTAAGGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACG
GCCAGGGCTACACACGTGCTACAATGGTTCGGTACAAAGGGTTGCCAAGCCGCGAGGTGGAGCTAATCCCATAAAACC
GATCGTAGTCCGGATCGCAGTCTTGAACTCGACTGCGTGAAGTCGGAATCG

>iso9

CGGGTGAGTAATGCCTAGGAATCTGCCTGGTAGTGGGGGATAACGTTCCGAAAGGAACGCTAATACCGCATACGTCCT
ACGGGAGAAAGCAGGGGACCTTCGGGCCCTTGCGCTATCAGATGAGCCTAGGTTCGGATTAGCTAGTTGGTGAAGTAATG
GCTAACCAAGGCGACGATCCGTAACCTGGTCTGAGAGGATGATCAGTCACACTGGAAGTGAAGACACGGTCCAGACTCCT
ACGGGAGGCGAGCAGTGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGAAGGTCTT
CGGATTGTAAAGCACTTTAAGTTGGGAGGAAGGGCTGTAGGCTAATATCCTGCGGTTTTGACGTTACCAACAGAATAA
GCACCGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCG
CGCGTAGGTGGTTTCGTAAAGTTGGATGTGAAAGCCCCGGGCTCAACCTGGGAAGTGCATCCAAAAGTGGCGAGCTAGA
GTACGGTAGAGGGTAGTGAATTTCTGTGTAGCGGTGAAATGCGTAGATATAGGAAGGAACACCAGTGGCGAAGGCG
ACTACCTGGACTGATACTGACACTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCC
GTAAACGATGTCAACTAGCCGTTGGAATCCTTGAGATTTTAGTGGCGCAGCTAACGCATTAAGTTGACCGCCTGGGGA
GTACGGCCGCAAGGTTAAAACCTCAAATGAATTGACGGGGGGCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGC
AACGCGAAGAACCTTACCTGGCCTTGACATGCTGAGAACTTTCCAGAGATGGAATGGTGCCTTCGGGAGCTCAGACCC
CGGTGCTGCATGGCTGCTGCTCAGCTCGTGTGCTGAGATGTTGGGTAAAGTCCCGTAACGAGCGCAACCCTTGTCCTTA
GTTACCAGCACGTTATGGTGGGCACTCTAAGGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTC
ATCATGGCCCTTACGGCCAGGTCTACACACGTGCTACAATGGTTCGGTACAAAGGGTTGCCAAGCCGCGAGGTGGAGCT
AATCCCATAAAACCGATCGTAGTCCGGATCGCAGTCTGCAACTCGACTGCGTGAAGTCGGAATCGCTAGTAATCGT

>iso16b

CGGGAAGTGGAGAGTGGCGCACGGGTGAGTAACGCGTGGGGACCTACCTCTTAGTGGGGGATAACGGTTGGAACGAC
CGCTAATACCGCATACGCCCTTCGGGGGAAAGATTTATCGCTAAGAGATGGACCCGCGTTGGATTAGATAGTTGGTGA
GGTAACGGCTCACCAAGTCAGCGATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCA
GACTCCTACGGGAGGCGAGCAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGAAGTGAAGA
AGGCCTTCGGGTTGTAAAGCTCTTTTCAGATGCGAAGATGATGACGGTAACATCAGAAGAAGCCCGGCTAATTTTCGTG
CCAGCAGCCGCGGTAATACGAAAGGGGCAAGCGTTGTTTCGGATTACTGGGCGTAAAGGGCACGCGAGGCGGCTTGCC
AGTCAGGGGTGAAAGCCCGGGGCTCAACCCCGGAAGTGCCTCTGATACTGCAGGACTAGAGACTAGGAGAGGGTGGTG
GAATTTCCAGTGTAGAGGTGAAATTCGTAGATATTGGGAGGAACACCAGAGGCGAAGGCCGCCACCTGGACTAGATCT
GACGCTCAGGTTTCGAAAGCGTGGGGAGCATACAAGGATTAGATACCCTGGTAGTCCACTCCGTAAACGATGAGTGCTA
GTTTCGCGGATTTTCGATTTTCGGTGACGCGAGCTAACGCATTAAGCACTCCGCTGGGGAGTACGGTCGCAAGATTAAAA
CTCAAGGAATTGACGGGGGGCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCCTTACCAA
CCC

>iso3

GGCGCACGGGTGCGTAACGCGTGGGAATCTGCCCTTCGGTTCGGAATAACTCAGGGAACTTGAAGCTAATACCGGATA
ATGACTTCGGTCCAAAGATTTATTGCCGAGGGATGAGCCCGGTAGGATTAGCTAGTTGGTGGGGTAATGGCTACCA

AGGCGACGATCCTTAGCTGGTCTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGG
CAGCAGTGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCAATGCCGCGTGAGTGATGAAGGCCTTAGGGTTGT
AAAGCTCTTTTACCAGGGATGATAATGACAGTACCTGGAGAATAAGCTCCGGCTAACTCCGTGCCAGCAGCCGCGGTA
ATACGGAGGGAGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGTTACTCAAGTCAGAGGTGAAAG
CCCGGGGCTCAACCCCGGAAGTGCCTTTGAAACTAGGTGACTAGAATCTTGAGAGAGGTGAGTGAATTCAGGTGTAG
AGGTGAAATTCGTAGATATTCGGAAGAACACCAGTGGCGAAGGCGACTGACTGGACAAGTATTGACGCTGAGGTGCGA
AAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGATAACTAGCTGTCCGGGCACATG
GTGTTTGGGTGGCGCAGCTAACGCATTAAGTTATCCGCTGGGGAGTACGGTCGCAAGATTAAACTCAAAGGAATTG
ACGGGGGCTGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAGCGTTTGACATCCT
G

>iso01

GTGGGGGATAACTCGGGGAACTCGAGCTAATACCGCATACTGCTCCTACGGGAGAAAGCGGGGGATCTTCGGACCTCGC
GCTACCAGATGAGCCTAGGTCGGATTAGCTAGTTGGTGAGGTAAAGGCTCACCAAGGCGACGATCCGTAGCTGGTCTG
AGAGGATGATCAGCCACACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGACAAT
GGGCGAAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGAAGGTCTTCGGATTGTAAAGCACTTTAAGTTGGGAGGAAG
GGCAGTAAGTTAATACCTTGCTGTTTTGACGTTACCGACAGAATAAGCACCGGCTAACTTCGTGCCAGCAGCCGCGGT
AATACGAAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGGTGGTTTGATAAGTTGGATGTGAAA
GCCCCGGGCTCAACCTGGGAATTGCATCCAAAAGTGTCTGACTAGAGTATGGCAGAGGGTGGTGAATTTCTGTGTGTA
GCGGTGAAATGCGTAGATATAGGAAGGAACACCAGTGGCGAAGGCGACCACCTGGGCTAATACTGACACTGAGGTGCG
AAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGTGCGACTAGCCGTTGGGATCCTT
GAGATCTTAGTGGCGCAGCTAACGCATTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTAAAACTCAAATGAATT
GACGGGGGCCCGCACAAAGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGCCTTGACATGCA
GAGAACCTTTCCAGAGATGGATTGGTGCCCTTCGGGAACCTGCACACAGGTGCTGCATGGCTGTCGTGCTGTCGT
GAGATGTTGGGTAAAGTCCCGTAACGAGCGCAACCCCTTGCTCTAGTTACCAGCACGTTAAGGTGGGCACTCTAAGGA
GACTGCCGTTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACGGCCTGGGCTACACACGTG
CTACAATGTTGCGGTACAAAGGGTTGCCAAGCCGCGAGGTGGAGCTAATCCCATAAACCGATCGTAGTCCGGATCGCA
GTCTGCAACTCGACTGCGTGAAGTCGGAATCGCTAGTAATCGTGAA

>iso22a

ACGGGTAGCCGGCCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTG
GGGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGGGGGATGACGGTCTTCGGATTGTAAACCCCT
TTCAGTAGGGACGAAGCGCAAGTGACGGTACCTGCAGAAGAAGCACCGGCTAACTACGTGCCAGCAGCCGCGGTAATA
CGTAGGGTGCAGCGTTGTCCGGAATTACTGGGCGTAAAGAGCTCGTAGGCGGTTTGTACGTCGTCTGTGAAATCCC
TCGGCTTAACCGGGGGCGTGCAGGCGATACGGGCAGACTTGAGTACTACAGGGGAGACTGGAATTCCTGGTGTAGCGG
TGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGGCGGGTCTCTGGGTAGTAAGTACGCTGAGGAGCGAAAG
CATGGGTAGCGAACAGGATTAGATACCCTGGTAGTCCATGCCGTAAACGGTGGGCGCTAGGTGTGGGGTCCCTTCCACG
GACTCCGTGCCGTAGCTAACGCATTAAGCGCCCCGCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTGAC
GGGGGCCCGCACAAAGCGGCGGAGCATGTGGATTAATTCGATGCAACGCGAAGAACCTTACCTAGGCTTGACATATACA
GGACGACGCGAGAGATGTGTTTTCCCTTGTGGCTTGTATACAGGTGGTGCATGTTTGTGTCGTAGCTCGTGTGAGAGA
TGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTGTCTCATGTTGCCAGCACGTAATGGTGGGGACTCGTGAGAGACT
GCCGGGTCAACTCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCCTATGTCTAGGGCTTCACACATGCTAC
AATGGCTAGTACAGAGGGCTGCGATACCGCAAGGTGGAGCGAATCCCTTAAAGCTGGTCTCAGTTCCGATTGAGGTCT
GCAACTCGACCCCATGAAGTCGGAGTCGCTAGTAATCGAAGATCAGCAACGCTGCGGTGAATACGTTCCC

>iso22g

CGGGTGAGTAACACGTGGGTAACTTGCCCTGCACCTTCGGGATAAGCCTGGGAACTGGGTCTAATACCGGATATGAAC
TCCTGCCGCATGGTGGGGTTTGGAAAGTTTTTCGGTGCAGGATGAGCCCGCGGCTATCAGCTTGTGGTGGGGTAAT
GGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCAGACTCC
TACGGGAGGCGAGCTAGTGGGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCTGGGGATGACGGTCT
TCGGATTGTAAACCCCTTTCAGTAGGGACGAAGCGCAAGTGACGGTACCTGCAGAAGAAGCACCGGCTAACTACGTGC
CAGCAGCCGCGGTAATACGTAGGGTGCAGCGTTGTCCGGAATTACTGGGCGTAAAGAGCTCGTAGGCGGTTTGTAC
GTCGTCTGTGAAATCCCTCGGCTTAACCGGGGGCGTGCAGGCGATACGGGCAGACTTGAGTACTACAGGGGAGACTGG
AATTCCTGGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGGCGGGTCTCTGGGTAGTAAGT
ACGCTGAGGAGCGAAAGCATGGGTAGCGAACAGGATTAGATACCCTGGTAGTCCATGCCGTATACGGTGGGCGCTAGG
TGTGGGGTCCCTCCACGGACTCCGTGCCGTAGCTAACGCATTAAGCGCCCCGCTGGGGAGTACGGCCGCAAGGCTAA
AACTCAAAGGAATTGACGGGGGCCCGCACAAAGCGCGGAGCATGTGGATTAAATTCGATGCAACGCGAAGAACCTTACC
TAGGCTTGACATATACAGGACGACGGCAGAGATGTGTTTTCCCTTGTGGCTTGTATACAGGTGGTGCATGGTTGTGCT
CAGCTCGTGTGAGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTGTCTCATGTTGCCAGCACGTAATGGTG
GGGACTCGTGAGAGACTGCCGGGTCAACTCGGAGGAAGGTGGAGATGACGTCAAATCATCATGACCCTTATGTCTAG
GGCTTCACACATGCTACAAATGGCTAGTACAGAGGGCTGAGATACCGCGAGGTGAAGCGAATCCCTTAAAGCTGGTCT
CAGTTCCGATTGGGGT

>iso29a

GGTTGAGTAACACTTGGGTAACTTGCCCTGCACCTTCGGGATAAGCCTGGGAACTGGGTCTAATACCGGATATGAAC
CCTGCCGCATGGTGGGGTTTGGAAAGTTTTTCGGTGCAGGATGAGCCCGCGGCTATCAGCTTGTGGTGGGGTAATG
GCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCAGACTCCT
ACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGGGGGATGACGGTCTT

CGGATTGTAAACCCCTTTTCAGTAGGGACGAAGCGCAAGTGACGGTACCTGCAGAAGAAGCACCGGCTAACTACGTGCC
AGCAGCCCGCGGTAATACGTAGGGTGCGAGCGTTGTCCGGAATTACTGGGCGTAAAGAGCTCGTAGGCGGTTTGTACAG
TCGTCTGTGAAATCCCTCGGCTTAACCGGGGGCGTGCAGGCGATACGGGCAGACTTGAGTACTACAGGGGAGACTGGA
ATTCTGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGGCGGGTCTCTGGGTAGTAACTGA
CGCTGAGGAGCGAAAGCATGGGTAGCGAACAGGATTAGATACCCTGGTAGTCCATGCCGTAAACGGTGGGCGCTAGGT
GTGGGGTCTTCCACGGACTCCGTGCCGTAGCTAACGCATTAAGCGCCCCGCTGGGGAGTACGGCCGCAAGGCTAAA
ACTCAAAGGAATTGACGGGGGCCCCGCACAAGCGGCGGAGCATGTGGATTAATTCGATGCAACGCGAAGAACCTTACCT
AGGCTTGACATATACAGGACGACGGCAGAGATGTCGTTTCCCTTGTGGCTTGTATACAGGTGGTGCATGGTTGTCGTC
AGCTCGTGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCCTGTCTCATGTTGCCAGCACGTAATGGTGG
GGACTCGTGAGAGACTGCCGGGGTCAACTCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATG

>iso5

TGAGTAATGCCTAGGAATCTGCCTGGTAGTGGGGGATAACGTCCGGAACGGGCGCTAATACCGCATACGTCTTACGG
GAGAAAGCAGGGGACCTTCGGGCTTGCCTATCAGATGAGCCTAGGTCGGATTAGCTAGTTGGTGAGGTAATGGCTC
ACCAAGCGACGATCCGTAACCTGGTCTGAGAGGATGATCAGTCACACTGGAAGTACGACACGGTCCAGACTCCTACGGG
AGGCAGCAGTGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGAAGGTCTTCGGAT
TGTAAGCACTTTAAGTTGGGAGGAAGGACAGTAAGTTAATACCTTGCTGTTTACGTTACCGACAGAATAAGCACCC
GGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCGCT
AGGTGGTTCGTTAAGTTGGATGTGAAAGCCCCGGGCTCAACCTGGGAAGTGCATCCAAAAGTGGCGAGCTAGAGTACG
GTAGAGGGTGGTGGAATTTCTGTGTAGCGGTGAAATGCGTAGATATAGGAAGGAACACCAGTGGCGAAGGCGACCAC
CTGGACTGATACTGACACTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAA
CGATGTCAACTAGCCGTTGGAGTCTTGGAGATTTAGTGGCGCAGCTAACGCATTAAGTTGACCGCCTGGGGAGTACG
GCCGCAAGGTTAAACTCAAATGAATTGACGGGGGCCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCG
AAGAACCTTACCTGGCCTTGACATGCTGAGAAGTTCCAGAGATGGATTGGTGCCCTTCGGGAAGTACAGACACAGGTGC
TGCATGGCTGTGCTCAGCTCGTGTGAGATGTTGGGTAAAGTCCCGTAACGAGCGCAACCCCTGTCTTAGTTACC
AGCACGTTATGGTGGGCACTCTAAGGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCATCATG
GCCCTTACGGCCAGGGCTACACACGTGCTACAATGGTTCGTTACAAAGGGTTGCCAAGCCGCGAGGTGGAGCTAATCCC
ATAAAACCGATCGTAGTCCGGATCGCAGTCTGCAACTCGACTGCGTGAAGTCGGAATCGCTAGTAATCGTGAATCA

>iso7

CCGGGAACGTATTACCGTGACATTCTGATTACGATTACTAGCGATTCCGACTTCACGCAGTCGAGTTGCAGACTGCG
GATCCGGACTACGATCGGTTTTATGGGATTAGCTCCACCTCGCGGCTTGGCAACCCTTTGTACCGACCATTGTAGCAC
GTGTGTAGCCCTGGCCGTAAAGGGCCATGATGACTTGACGTATCCCCACCTTCCTCCGGTTTGTACCGGCAGTCTCC
TTAGAGTGCCCAACATAACGTGCTGGTAACTAAGGACAAGGGTTGCGCTCGTTACGGGACTTAACCCAACATCTCACG
ACACGAGCTGACGACAGCCATGCAGCACCTGTGTCTGAGCTCCCGAAGGCACCAATCCATCTCTGGAAAGTTCTCAGC
ATGTCAAGGCCAGGTAAGGTTCTTCGCGTTGCTTCAATTAACACCATGCTCCACCGCTTGTGCGGGCCCCCGTCAA
TTCAATTTAGTTTTTAACCTTGCGGCCGTACTCCCGAGGCGGTCAACTTAATCGGTTAGCTGCGCCACTAAAATCTCAA
GGATTCCAACGGCTAGTTGACATCGTTTTACGGCGTGGACTACCAGGGTATCTAATCCTGTTTGTCTCCACGCTTTTCG
CACCTCAGTGTCAGTATCAGTCCAGGTAGTCGCCTTCGCCACTGGTGTTCCTTCCTATATCTACGCATTTTACCGCTA
CACAGGAAATTCCACTACCTCTACCGTACTCTAGCTCGCCAGTTTTTGGATGCAGTTCCAGGTTGAGCCCGGGGCTT
TCACATCCAACCTTAACGAACCACCTACGCGCGCTTTACGCCAGTAATCCGATTAACGCTTGCACCCCTTCGTATTAC
CGCGGCTGCTGGCACGAAGTTAGCCGGTGCTTATTCTGTTGGTAACGTCAAACCGCAGGATATTAGCCTACAGCCCT
TCCTCCCAACTTAAAGTGCTTTACAATCCGAAGACCTTCTTACACACGCGGCATGGCTGGATCAGGCTTTTCGCCAT
TGTCCAATATTCCCACTGCTGCCTCCCGTAGGAGTCTGGACCGTGTCTCAGTTCCAGTGTGACTGATCATCCTCTCA
GACCAGTTACGGATCGTCGCTTGGTGAGCCATTACCTCACCAACTAGCTAATCCGACCTAGGCTCATCTGATAGCGC
AAGGCCCGAAGGTCCCTGCTTTCTCCCGTAGGACGTATGCGGTATTAGCGTTCCCTTTCGGAACGTTATCCCCACT

>iso8

AGGAATCTGCCTGGTAGTGGGGGATAACGTTCCGAAAGGAACGCTAATACCGCATACGTCTTACGGGAGAAAGCAGGG
GACCTTCGGGCCCTTGCGCTATCAGATGAGCCTAGGTTCGGATTAGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACG
ATCCGTAACCTGGTCTGAGAGGATGATCAGTCACACTGGAAGTACGACACGGTCCAGACTCCTACGGGAGGCAGCAGTG
GGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGAAGGTCTTCGGATTGTAAGCACT
TTAAGTTGGGAGGAAGGGCTGTAGGCTAATATCCTGCGGTTTTGACGTTACCAACAGAATAAGCACCGGCTAACTTCG
TGCCAGCAGCCGCGGTAATACGAAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGGTGGTTCGT
TAAGTTGGATGTGAAAGCCCCGGGCTCAACCTGGGAAGTGCATCCAAAAGTGGCGAGCTAGAGTACGGTAGAGGGTAG
TGGAATTTCTGTGTAGCGGTGAAATGCGTAGATATAGGAAGGAACACCAGTGGCGAAGGCGACTACCTGGACTGATA
CTGACACTGAGGTGCGAAAGCGTGGGGAGCAAAACAGGATTAGATACCCTGGTAGTCCACGCGCTAAACGATGCTCAACT
AGCCGTTGGAATCCTTGAGATTTTAGTGGCGCAGCTAACGCATTAAGTTGACCGCCTGGGGAGTACGGCCGCAAGGTT
AAAACCTCAAATGAATTGACGGGGGCCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCCTTA
CCTGGCCTTGACATGCTGAGAACTTTCCAGAGATGGATTGGTGCCTTCGGGAGCTCAGACACAGGTGCTGCATGGCTG
TCGTACAGCTCGTGTGAGATGTTGGGTAAAGTCCCGTAACGAGCGCAACCCCTGTCTTAGTTACCAGCACGTTAT
GGTGGGCACTCTAAGGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACGG
CCAGGGCTACACACGTGCTAATGGTTCGTTACAAAGGGTTGCCAAGCCGCGAGGTGGAGCTAATCCCATAAACCGAT
CGTAGTCCGGATCGCAGTCTGCAACTCGACTGCGTGAAGTCGGAATCGCTAGTAATCGTGAATCAGAATGTCACGGTG
AATACGTTCCCG

>iso10

AGCGATTCCAGCTTCATGTAGGCGAGTTGCAGCCTACAATCCGAACTGAGAACGGTTTTATGAGATTAGCTCCACCTC
GCGGTCTTCGAGCTCTTTGTACCGTCCATTGTAGCACGTGTGTAGCCCAGGTCATAAGGGGCATGATGATTTGACGTC
ATCCCCACCTTCTCCGGTTGTACCCGGCAGTCACCTTAGAGTGCCCACTTAATGATGGCAACTAAGATCAAGGGTT
GCGCTCGTTGCGGGACTTAACCAACATCTCACGACAGAGTGCAGACAACCATGCACCACCTGTCACTCTCTCCCG
AAGGAGAAGCCCTATCTCTAGGGTTTTTCAGAGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAATTA
ACATGCTCCACCGCTTGTGCGGGCCCCGTCATTTCTTTGAGTTTCAGCCTTGCGGCCGTACTCCCCAGGCGGAGTGC
TTAATGCGTTAACTTCAGCACTAAAGGGCGGAAACCTCTAACACTTAGCACTCATCGTTTACGGCGTGGACTACCAG
GGTATCTAATCCTGTTGCTCCCCACGCTTTCGCGCCTCAGTGTACAGTTACAGACCAGAAAGTCGCTTCGCCACTGGT
GTTCTCTCATATCTCTACGCATTTACCGCTACACATGGAATTCACCTTTCTCTTCTGCACTCAAGTCTCCAGTTTC
CAATGACCTCCACGGTTGAGCCGTGGGCTTTCACATCAGACTTAAGAAACCACCTGCGCGCGCTTTACGCCCAATAA
TTCCGGATAACGCTTGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTGGCTTCTGGTTAGGTACCG
TCAAGGTGCCAGCTTATTCAACTAGCACTTGTCTTCCCTAACAAACAGAGTTTACGACCCGAAAGCCTTCATCACTC
ACGCGGCGTTGCTCCGTGAGACTTTTCGTCCATTGCGGAAGATTCCCTACTGCTGCCTCCCGTAGGAGTCTGGGCCGTG
TCTCAGTCCAGTGTGGCCGATCACCTCTCAGGTGCGCTACGCATCGTTGCCTTGGTGAGCCGTTACCTCACC
AGCTAATGCGACGCGGGTCCATCCATAAGTGACAGCCGAAGCCGCTTTCAATTTGAAACCATGCGGTTCAAATGTT
ATCCGGTATTAGCCCCGGTTTCCCGGAGTTATCCAGTCTTATGGGCAGGTTACCCACGTGTTACTCACC

>iso11

CTGCCATAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATAACATTTTGAACCGCATGGTTCGAAATTG
AAAGGCGGCTTCGGCTGTCACTTATGGATGGACCCGCGTCGATTAGCTAGTTGGTGAGGTAAACGGCTCACCAAGGCA
ACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCGAGCA
GTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGATGAAGGCTTTCGGGTTCGTA
TCTGTTGTTAGGGAAGAACAAGTGCTAGTTGAATAAGCTGGCACCTTGACGGTACCTAACAGAAAGCCACGGCTAAC
TACGTGCCAGCAGCCGCGTAAATACGTAGGTGGCAAGCGTTATCCGGAATTATTGGGCGTAAAGCGCGCGCAGGTGGT
TTCTTAAGTCTGATGTGAAAGCCACGGCTCAACCGTGGAGGGTCAATTGGAAGTGGGAGACTTGAGTGCAGAAGAGG
AAAGTGAATTCATGTGTAGCGGTGAAATGCGTAGAGATATGGAGGAACACCAAGTGGCGAAGGCGACTTTCTGGTCT
GTAACGTACACTGAGGCGCGAAAGCGTGGGAGCAACAGGATTAGATACCTGGTAGTCCACGCCGTAAACGATGAG
TGCTAAGTGTTAGAGGGTTTCCGCCCTTTAGTGCTGAAGTTAACGCATTAAGCACTCCGCCGTGGGGAGTACGCCCGCA
AGGCTGAAACTCAAAGGAATTGACGGGGGGCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAA
CCTTACAGGTCCTTGACATCCTCTGAAAACCTAGAGTAGGGCTTCTCCTTCGGGAGCAGAGTGACATGTTGGTGCAT
GGTTGTCTGCTCAGCTCGTGTGCTGAGATGTTGGNGTTAAGTCCCGCAACGAGCGCAACCTTGATCTAGTTGCCATCA
TTAAGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCCTT
ATGACCTGGGCTACACACGTGCTACAATGGACGGTACAAGAGCTGCAAGACCGCGAGGTGGAGCTAATCTCATAAAA
CCGTTCTCAGTTTCGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCG
CGGTGAATACGTTCCCGG

>iso21b

CGGGTGAGTAACACGTGAGCAACCTGCCCCGACTCTGGGATAAGCGCTGGAAACGGCGTCTAATACTGGATACGAGA
CGTGATCGCATGGTCAACGTCTGGAAAGATTTTTCGGTCTGGGATGGGCTCGCGGCCTATCAGCTTGTGGTGAGGTA
ATGGCTCACCAAGGCGTCGACGGGTAGCCGCGCTGAGAGGGTGACCGGCCACACTGGGACTGAGACACGGCCAGACT
CCTACGGGAGGCGAGTGGGGAATATTGCACAATGGGCGGAAGCCTGATGCAGCAACGCCGCGTGAGGGATGACGGC
CTTCGGGTTGTAAACCTCTTTAGCAGGGAAGAAGCGAGAGTGACGGTACCTGCAGAAAAAGCGCCGGCTAACTACGT
GCCAGCAGCCGCGTAAATACGTAGGGCGCAAGCGTTATCCGGAATTATTGGGCGTAAAGAGCTCGTAGGCCGTCTGTC
GCGTCTGCTGTGAAAACCCGAGGCTCAACCTCGGGCCTGCAGTGGGTACGGGCAGACTAGAGTGCGGTAGGGGAGATT
GGAATTCCTGGTGTAGCGGTGGAATGCGCAGATATCAGGAGGAACACCGATGGCGAAGGCAGATCTCTGGGCCGTAAC
TGACGCTGAGGAGCGAAAGGTGGGGAGCAACAGGCTTAGATACCTGGTAGTCCACCCCGTAAACGTTGGGAACTA
GTTGTGGGGTCTTTCCACGGATTCCGTGACGCAGCTAACGCATTAAGTTCCCGCCTGGGGAGTACGGCCGCAAGGC
TAAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGCGGAGCATGCGGATTAATTCGATGCAACGCGAAGAACCTT
ACCAAGGCTTGACATACAGAGAACGCTGCAGAAATGTAGAATCTTTGGACACTCGTGAACAGGTGGTGCATGTTG
TCGTGAGCTCGTGTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTCGTTCTATGTTGCCAGCACGTAAT
GGTGGGAACCTATGGGATACTGCCGGGTCAACTCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCCTATGT
CTTGGGCTTCACGCATGCTACAATGGCCGGTACAAAGGCTGCAATACCGTGAGGTGGAGCGAATCCCAAAAAGCCGG
TCCCAGTTTCGATTGAGGTCTGCAACTCGACCTCATGAAGTTCGGAGTGCCTAGTAATCGCAGATCA

>iso26

ACGCATGGAACGTGTGCTAATACCGTATGAGCCCTTCGGGGGAAAGATTATCGGGAAAGGATGAGCCCGCTTGGAT
TAGCTAGTTGGTGGGGTAAAGGCCTACCAAGGCGACGATCCATAGCTGGTCTGAGAGGATGATCAGCCACATTGGGAC
TGAGACACGGCCCCAACTCCTACGGGAGGCGAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGC
CGCGTGAGTGATGAAGGCCCTTAGGGTTGTAAAGCTCTTTACCGGTGAAGATAATGACGGTAACCGGAGAAGAGCCC
CGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGGCTAGCGTTGTTCCGGAATTACTGGGCGTAAAGCGCACG
TAGGCGGACATTTAAGTCAGGGGTGAAATCCAGAGCTCAACTCTGGAAGTGCCTTTGATACTGGGTGTCTTGAGTAT
GGAAGAGGTAAGTGGAAATCCGAGTGTAGAGGTGAAATTCGTAGATATTTCGAGGAACACCAAGTGGCGAAGGCGCTT
ACTGGTCCATTACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAAACAGGATTAGATAACCTGGTAGTCCACGCGGTAA
ACGATGAATGTTAGCCGTGGGCGAGTATACTGTTCCGTGGCGCAGCTAACGCATTAACATTCGCCCTGGGGAGTACG
GTCGCAAGATTAAACTCAAAGGAATTGACGGGGGGCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGC
GCAGAACCTTACCAGCCCTTGACATC

>isol6c
GAATACTCAGAGAAATTGAGCTAATACGGATAATGTCTTCGGACCAAAGATTTATCGCCTTTGGATGAGCCCCGCGTAA
GATTAGCTAGTTGGTAGGGTAAAGGCCTACCAAGGCGACGATCTTTAGCTGGTCTGAGAGGATGATCAGCCACACTGG
GACTGAGACACGGCCAGACTCCTACGGGAGGCGACGAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCAGTCAA
TGCCGCGTGAGTGATGAAGGCCTTCGGGTGCTAAAGCTCTTTTACCAGGGATGATAATGACAGTACCTGGAGATAAG
CTCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGAGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGCGC
GCGTAGGCGGTCTACCAAGTCAGGGGTGAAATCCCAGGGCTCAACCCCGGAACTGCCCTTGAAACTATTAGACTAGAA
TCTTGGAGAGGTGAGTGGAAATCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAAGAACACCAGTGGCGAAGCGCA
CTGACTGGACAAGTATTGACGCTGAGGTGCGAAAGCGTGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCG
TAAACGATGATAACTAGCTGTCCGGGCACATGGTGCTTGGGTGGCGCAGCTAACGCATTAAGTTATCCGCCTGGGGAG
TACGGTCGCAAGATTAAAACTCAAAGGAATTGACGGGGCCTGCACAAGCGGTGGAGCATGTGGTTTAATTTCGAAGCA
ACGCGCAGAACCTTACCAGCCTTTGACATCCCGGT

>isol1
CTGCCCATAAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATAACATTTTGAACCGCATGGTTCGAAAATTG
AAAGGCGGCTTCGGCTGTCACTTATGGATGGACCCGCTCGCATTAGCTAGTTGGTGAGGTAAACGGCTCACCAAGGCA
ACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCA
GTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGATGAAGGCTTTCGGGTGCTAAAAAC
TCTGTTGTTAGGGAAGAACAAGTGCTAGTTGAATAAGCTGGCACCTTGACGGTACCTAACCGAAGGCCACGGCTAAC
TACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAATTATTGGGCGTAAAGCGCGCAGGTGGT
TTCTTAAGTCTGATGTGAAAGCCACGGCTCAACCGTGGAGGGTCATTGGAAGCTGGGAGACTTGAGTGCAGAAGAGG
AAAGTGGAATTCATGTGTAGCGGTGAAATGCGTAGAGATATGGAGGAACACCAGTGGCGAAGGCGACTTTCGGTCT
GTAAGTACACTGAGGCGGAAAGCGTGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAG
TGCTAAGTGTTAGAGGGTTTCCGCCCTTTAGTGCTGAAGTTAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGCA
AGGCTGAAACTCAAAGGAATTGACGGGGGCCGCAACAGCGGTGGAGCATGTGGTTTAATTTCGAAGCAACGCGAAGAA
CCTTACCAGGTCTTGACATCCTCTGAAAACCTAGAGATAGGGCTTCTCCTTCGGGAGCAGAGTGACAGGTGGTGAT
GGTTGTCTGTCAGCTCGTGTCTGAGATGTTGNGTTAAGTCCCGCAACGAGCGCAACCCCTTGATCTTAGTTGCCATCA
TTAAGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTT
ATGACCTGGGCTACACACGTGCTACAATGGACGGTACAAAGAGCTGCAAGACCGCGAGGTGGAGCTAATCTCATAAAA
CCGTTCTCAGTTTCGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCG
CGGTGAATACGTTCCCGG

>iso20a
CTGCACTTTGGGATAAGCCTGGGAAACTGGGTCTAATACCGAATATGACTCCCGACTGCATGGTCTGGGGTGGAAAGC
TTTTGCGGTGTGGGATGGGCCCCGCGCCTATCAGCTTGTGGTGGGGTGATGGCCTACCAAGGCGACGACGGGTAGCC
GGCCTGAGAGGGTGTCCGGCCACACTGGGACTGAGATACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTG
CACAATGGGCGCAAGCCTGATGCAGCGACGCCGCTGAGGGATGACGGCCTTCGGGTTGTAAACCTCTTTCGCCAGGG
ACGAAGCGCAAGTGACGGTACCTGGAGAAGAAGCACCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGC
GAGCGTTGTCCGAATTACTGGGCGTAAAGAGCTCGTAGGTGGTTTGTTCGCGTTGTTCTGTGAAAACCTCACAGCTTAA
TGTGGGCGTGCGGGCGATACGGGCAGACTGGAGTACTGCAGGGGAGACTGGAATTCCTGGTGTAGCGGTGGAATGCGC
AGATATCAGGAGGAACACCGGTGGCGAAGCGGGTCTCTGGGCAGTAACCTGACGCTGAGGAGCGAAAGCGTGGGGAGC
GAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGGTACTAGGTGTGGGTTTCCTTTCCTTGGGATCCGTG
CCGTAGCTAACGCATTAAGTACCCCGCTGGGGAGTACGGCCGAAGGCTAAAACTCAAAGGAATTGACGGGGGCCCG
CACAAGCGCGGAGCATGTGGATTAATTTCGATGCAACGCGAAGAACCTTAC

>iso21c
CCTGCACTTTGGGATAAGCCTGGGAAACTGGGTCTAATACCGAATAGGACTCCGGACTGCATGGTCTGGGGTGGAAAG
CTTTTGCGGTGTGGGATGGGCCCCGCGCCTATCAGCTTGTGGTGGGGTGATGGCCTACCAAGGCGACGACGGGTAGCC
CGGCTGAGAGGGTGTCCGGCCACACTGGGACTGAGATACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTG
GCACAATGGGCGCAAGCCTGATGCAGCGACGCCGCTGAGGGATGACGGCCTTCGGGTTGTAAACCTCTTTCGCCAGG
GACGAAGCGCAAGTGACGGTACCTGGAGAAGAAGCACCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGC
CGAGCGTTGTCCGAATTACTGGGCGTAAAGAGCTCGTAGGTGGTTTGTTCGCGTTGTTCTGTGAAAACCTCACAGCTTAA
CTGTGGGCGTGCGGGCGATACGGGCAGACTGGAGTACTGCAGGGGAGACTGGAATTCCTGGTGTAGCGGTGGAATGCG
CAGATATCAGGAGGAACACCGGTGGCGAAGCGGGTCTCTGGGCAGTAACCTGACGCTGAGGAGCGAAAGCGTGGGGAG
CGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGGTACTAGGTGTGGGTTTCCTTTCCTTGGGATCCGT
GCCGTAGCTAACGCATTAAGTACCCCGCTGGGGAGTACGGCCGAAGGCTAAAACTCAAAGGAATTGACGGGGGCCCG
GCACAAGCGCGGAGCATGTGGATTAATTTCGATGCAACGCGAAGAACCTTACCTGGGT

>iso22b
TGCACTTTGGGATAAGCCTGGGAAACTGGGTCTAATACCGAATATGACTCCGGACTGCATGGTCTGGGGTGGAAAGCT
TTTTGCGGTGTGGGATGGGCCCCGCGCCTATCAGCTTGTGGTGGGGTGATGGCCTACCAAGGCGACGACGGGTAGCCG
GCCTGAGAGGGTGTCCGGCCACACTGGGACTGAGATACGGCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGC
ACAATGGGCGCAAGCCTGATGCAGCGACGCCGCTGAGGGATGACGGCCTTCGGGTTGTAAACCTCTTTCGCCAGGGA
CGAAGCGCAAGTGACGGTACCTGGAGAAGAAGCACCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGC
AGCGTTGTCCGAATTACTGGGCGTAAAGAGCTCGTAGGTGGTTTGTTCGCGTTGTTCTGTGAAAACCTCACAGCTTAA
GTGGGCGTGCGGGCGATACGGGCAGACTGGAGTACTGCAGGGGAGACTGGAATTCCTGGTGTAGCGGTGGAATGCGCA
GATATCAGGAGGAACACCGGTGGCGAAGCGGGTCTCTGGGCAGTAACCTGACGCTGAGGAGCGAAAGCGTGGGGAGCG
AACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGGTACTAGGTGTGGGTTTCCTTTCCTTGGGATCCGTG

CGTAGCTAACGCATTAAGTACCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACCTCAAAGGAATTGACGGGGGCCCCG
ACAAGCGCGCGGAGCATGTGGATTAATTTCGATGCAACGCGAAGAACCTTACCTGGGTTTGACAT

>iso22i

CCTGCACTTCGGGATAAGCCTGGGAAACTGGGTCTAATAGCGGATATGAACTCCTGCCGCATGGTGGGGGTTGGAAAG
TTTTCCGGTGCAGGATGAGCCCGCGGCTATCAGCTTGTGGTGGGGTAATGGCCTACCAAGGCGACGACGGGTAGCC
GGCCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCAGACTCCTATTGGAGGCAGCAGTGGGGAATATTG
CCCAATGGGCGAAAGCCTGATGCAGCGACGCCGCTGGGGGATGACGGTCTTCGGATTGTAAAACCCCTTTTCAGTAGG
GACGAAGCGCAAGTGACGGTACCTGCAGAAGAAGCACC GGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTG
CGAGCGTTGTCCGGAATTACTGGGCGTAAAGAGCTCGTAGCGGTTTGTACGTCTGTGAAATCCCTCGGCTTAA
CCGGGGGCGTGCAGGCGATACGGGCAGACTTGAGTACTACAGGGGAGACTTGAATTCCTGGTGTAGCGGTGAAATGCG
CAGATATCAGGAGGAACACCGGTGGCGAAGCGGGTCTCTGGGTAGTAAGTACGTGAGGAGCGAAAGCATGGGTAG
CGAACAGGATTAGATACCCTGGTAGTCCATGCCGTAAACGGTGGGCGCTAGGTGTGGGGTCTTCCACGGACTCCGTG
CCGTAGCTAACGCATTAAGCGCCCCGCTGGGGAGTACGGCCGCAAGGCTAAAACCTCAAAGGAATTGACGGGGGCCCCG
CACAAGCGCGGAGCATGTGGATTAATTCGATGCAACGCGAAGAACCTTACCTAGGCTTGA

>iso21d

AGCCTAGGTAAGGTTCTTCGCGTTGCATCGAATTAATCCACATGCTCCGCCGCTTGTGCGGGCCCCCGTCAATTCCTT
TGAGTTTTAGCCTTGCGGCCGTACTCCCCAGGCGGGGCGCTTAATGCGTTAGCTACGGCACGGAGTCCGTGGAAGGAC
CCCACACCTAGCGCCACCGTTTACGGCATGGACTACCAGGGTATCTAATCCTGTTTCGTACCCATGCTTTTCGCTCCT
CAGCGTCAGTTACTACCCAGAGACCCGCCCTTCGCCACCGGTGTTCTCCTGATATCGGCGCATTTACCCGCTACACCA
GGAATTCAGTCTCCCTGTAGTACTCAAGTCTGCCCGTATCGCCTGCACGCCCCCGGTTAAGCCGAGGGATTTCACA
GACGACGTGACAAACCGCCTACGAGCTCTTTACGCCAGTAATTCGGACAACGCTCGCACCTACGTATTACCGCGG
CTGTGGCAGCTAGTTAGCCGGTGCTTCTCTGCAAGTACCGTCACTTGCGCTTCGTCCCTACTGAAAGGGGTTTACA
ATCCGAAGACCGTCATCCCCACGCGCGCTCGCTGCATCAGGCTTTTCGCCATTGTGCAATATTCCCACTGCTGCCT
CCCGTAGGAGTCTGGGCCGTGTCTCAGTCCAGTGTGGCCGATACCCTCTCAGGCCGGCTACCCGTCGTGCCTTGG
TAGGCCATTACCCACCAACAAGCTGATAGCCCGGGCTCATCCTGCACCGAAAACTTTCCAACCCCCACCATGCG
GCAGGAGTTCATATCCGGTATTAGACCCAGTTTCCAGGCTTATCCCGAAGTGCAGGGCAG

>isoa2

ACGAGCGCGAAGCGGTGAGTAACACGTGGGTGATCTGCCCTGCACTTCGGGATAAGCCTGGGAAACTGGGTCTAATA
CCGGATATGACCTCCTATCGCATGGTGGGTGGTGGAAAGATTATCGGTGCAGGATGGGCCCCGCGGCTATCAGCTTG
TTGGTGGGGTAATGGCCTACCAAGGCGACGACGGGTAGCCGACCTGAGAGGGTGACCGGCCACACTGGGACTGAGACA
CGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGA
GGGATGACGGCCTTCGGGTTGTAAACCTCTTTCAGCAGGGACGAAGCGCAAGTGACGGTACCTGCAGAAGAAGCACCG
GCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGAAGCGTTGTCCGGAATTACTGGGCGTAAAGAGTTTCGTA
GGCGGTTTTGTGCGCTCGTTTGTGAAAACCGCAGCTCAACTGCTGGCTTGCAAGCGATACGGGCAGACTTGAGTACTG
CAGGGGAGACTGGAATTCCTGGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGGCGGGTCTC
TGGGCAGTAAGTACGCTGAGGAACGAAAGCGTGGGTAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAAC
GGTGGGCGTAGGTGTGGGTTTCCTTCCACGGAATCCGTGCCGTAGCTAACGCATTAAGCGCCCCGCTGGGGAGTACG
GCCGCAAGGCTAAAACCTCAAAGGAATTGACGGGGGCCCCGACAAGCGGCGGAGCATGTGGATTAATTCGATGCAACGC
GAAGAACCTTACCTGGGTTTGACATATACCGGAAAGCTGCAGAGATGTGGCCCCCTTGTGGTGGTATACAGGTGGT
GCATGGCTGTGCTCAGCTCGTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCTTATGTTGCCA
GCACGTTATGGTGGGACTCGTAAGAGACTGCCGGGGTCAACTCGGAGGAAGGTGGGACGACGTCAAGTCATCATGC
CCCTTATGTCCAGGGCTTCACACATGCTACAATGGCCAGTACAGAGGGCTGCGAGACCGTGAGGTGGAGCGAATCCCT
TAAAGCTGGTCTCAGTTCGGATCGGGTCTGCAACTCGACCCCGTGAAGTCGGAGTGCCTAGTAATCGCAGATCAGCA
ACGCTGCGGTGAATACGTTCCCG

>iso22f

CGGGATAGTGGCAGACGGGTGAGTAACGCGTGGGAATCTACCTTTTGGTACGGAATAGCTCCGGGAAACTGGAATTAA
TACCGTATGTGTCCTTCGGGAGAAAGATTATCGCCAAGAGATGAGCCCGCGTTTGATTAGGTAGTTGGTGGGGTAAT
GGCCTACCAAGCCGACGATCAATAGCTGGTCTGAGAGGATGATCAGCCACATTGGGACTGAGACACGGCCCAAACCTCC
TACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGAGTGATGAAGGCC
TAGGGTTGTAAAGCTCTTTCGCCGCGTGAAGATAATGACTGTAGCCGGTAAAGAAGCCCCGGCTAACTTCGTGCCAGCA
GCCGCGGTAATACGAAGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACATTTAAGTCAG
GGGTGAAATCCCGGAGCTCAACTCCGGAAGTGCCTTTGATACTGGGTGTCTCGAGTATGGAAGAGGTAAGTGAATTG
CGAGTGTAGAGGTGAAATTCGTAGATATTTCGAGGAACACCAAGTGGCGAAGCGGCTTACTGGTCCATTACTGACGCT
GAGGTGCGAAAGCGTGGGGAGCAAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGATAGCTAGCCGTG
GTGTGCATGCATATCGGTGGCGCAGTTAACGCATTAAGCTATCCGCTGGGGAGTACGGTCGCAAGATTAAAACCTCAA
AGGAATTGACGGGGGCCCCGACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGCAGAACCTTACCAGCCCTT
GACATTTGGTGCTACCTCGAGAGATCGAGGGTTCCTTCGGGACGCCAGGACAGGTGCTGCATGGCTGTGCTCAGCTCG
TGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTCGCCCTTAGTTGCCAGCATTACGTTGGGCACTCTA
AGGGCACTGCCAGTGATAAGCTGGAGGAAGGTGGGATGAGCTCAAGTCCTCATGGCCCTACGGGCTGGGCTACACA
GTGCTACAAATGGTGGTGACAGTGGGCGAGCGACGCGAGTGTGAGCTAATCTCCAAAAGCCATCTCAGTTCGGATTG
CACTCTGCAACTCGGGTGCATGAAGTTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGG
CC

>iso22j
ACACATGCAAGTCGAGCGCCCCCTCGGGGGAGCGGCAGACGGGTGAGTAATGCATGGGAATCTACCTGGCTCTACGCG
AATAACTCAGGGAAACTTGAACATAACCGTATACGTCCCTCGGGAGAAAGATTTATCGGAGTTAGATGAGCCCATGT
CTGATTAGCTAGTTGGTGAGGTAAAGGCTACCAAGGCGACGATCAGTAGCTGGTCTGAGAGGATGATCAGCCACACT
GGGACTGAGACACGCCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGC
CATGCCCGCTGAGTGATGAAGGCCCTAGGGTTGTAAAGCTCTTTCACCGGTGAAGATAATGACGGTAACCGGAGAAGA
AGCCCCGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCGTAAAGC
GCGCGTAGGCGGATCGGTGAGTTAGGGGTGAAATCCCGGGGCTCAACCCCGGAAGTGCCTCTAATACTGCCGATCTAG
AGTTCGAGAGAGGTGAGTGAATTCGAGTGTAGAGGTGAAATTCGTAGATATTTCGGAGGAACACAGTGGCGAAGGC
GGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATAACCTGGTAGTCCACGC
CGTAAACGATGGAAGCTAGCCGTGCGCAGGTTTACCTGTGCGGTGGCGCAGTTAACGCATTAAGCTTCCCGCTGGGGA
GTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAAGCGGTGGAGCATGTGGTTTAATTCGAAGC
AACGCGCAGAACCTTACCAGCCCTTGACATCTGGGTGCGGATACC

>iso27c
GTGGGGACCTACCTCTTAGTGGGGGATAACGGTTGGAAACGACCGCTAAACCGCATACGCCCTTCGGGGGAAAGATTA
TCGCTAAGAGATGGACCGGTTGGATTAGATAGTTGGTGAGGTACGGCTACCAAGTCAGCGATCCATAGCTGGTT
TGAGAGGATGATCAGCCACACTGGGACTGAGACACGCCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGACA
ATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGAGTGAAGAAGGCCCTTCGGGTTGTAAAGCTCTTTCAGATGCGAAGA
TGATGACGTTAACATCAGAAAGAGCCCCGGCTAATTTTCGTGCCAGCAGCCGCGGTAATACGAAAGGGGCAAGCGTTGT
CGGATTTACTGGGCGTAAAGGGCACGCAGGCGGTCTTGCCAGTCAGGGGTGAAAGCCCGGGGCTCAACCCCGGAAGTGC
CTCTGATACTGCAGGACTAGAGACTAGGAGAGGGTGGTGAATTCACAGTGTAGAGGTGAAATTCGTAGATATTGGGA
GGAACACCAGAGGCGAAGGCGGCCACCTGGACTAGATCTGACGCTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTA
GATACCTTGGTAGTCCACGCGGTAAACGATGAGTGCTAGTTGTGCGGATTTTCGATTTTCGGTGACGCAGCTAACGCATT
AAGCACTCCGCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAAGCGGTGGAGCA
TGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCTTGACATCCCTATC

>iso19a
GGGGGTACCGACTGGCGAACTGGTGAGTACACGTGGGTAACTCTGCCCTGCACTTCGGGATAAGCCTGGGAAACTGGGT
CTAATACCGGATATGAACTCCTGCCGCATGGTGGGGGTTGGAAAGTTTTTCGGTGACAGGATGAGCCCGCGCCTATCA
GCTTGTGGTGGGGTAATGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTGATCGGCCACACTGGGACTG
AGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCG
CGTGGGGGATGACGGTCTTCGGATTGTAAACCCCTTTTCAGTAGGGACGAAGCGCAAGTGACGGTACCTGCAGAAGAAG
CACC GGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGCGAGCGTTGTCCGGAATTACTGGGCGTAAAGAGC
TCGTAGGCGGTTTGTACGTCGTCTGTGAAATCCCTCGGCTTAACCGGGGGCGTGACGGCGATACGGGCAGACTTGAG
TACTACAGGGGAGACTGGAATTCCTGGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGGCGG
GTCTCTGGGTAGTAACGTGACGCTGAGGAGCGAAAGCATGGGTAGCGAACAGGATTAGATACCTTGGTAGTCCATGCCG
TAAACGGTGGGCGCTAGGTGTGGGGTCCTTCCACGACTCCGTGCCGTAGCTAACGCATTAAGCGCCCCGCTGGGGA
GTACGGCCGCAAGGCTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAAGCGGCGAGCATGTGGATTAATTCGATGC
AACGCGAAGAACCTTACCTAGGCTTGA

>iso19e
CTCTTAGTGGGGGATAACGGTTGGAAACGACCGCTAATACCGCATACGCCCTTCGGGGGAAAGATTTATCGCTAAGAG
ATGGACCCCGGTTGGATTAGATAGTTGGTGAGGTAAACGGCTACCAAGTCAGCGATCCATAGCTGGTTTGAGAGGATG
ATCAGCCACACTGGGACTGAGACACGCCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGGGCAA
CCCTGATCCAGCCATGCCGCGTGAGTGAAGAAGGCCTTCGGGTTGTAAAGCTCTTTCAGATGCGAAGATGATGACGGT
AACATCAGAAGAAGCCCCGCTAATTTTCGTGCCAGCAGCCGCGGTAATACGAAAGGGGCTAGCGTTGTTTCGGATTTAC
TGGGCGTAAAGGGCACGCAGGCGGTCTTGCCAGTCAGGGGTGAAAGCCCGGGGCTCAACCCCGGAAGTGCCTCTGATA
CTGCAAGACTAGAGACTAGGAGAGGGTGGTGAATTCACAGTGTAGAGGTGAAATTCGTAGATATTGGGAGGAACACC
AGAGGCGAAGGCGGCCACCTGGACTAGATCTGACGCTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCTT
GGTAGTCCACGCCGTAAACGATGAGTGCTAGTTGTGCGGATTTTCGATTTTCGGTGACGCAGCTAACGCATTAAGCACTC
CGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAAGCGGTGGAGCATGTGGTTT
AATTCGAAGCAACGCGCAGAACCTTACCAACCTT

ANEXO B - SEQUÊNCIAS DE RDNA 16S AMPLIFICADAS A PARTIR DE AMOSTRAS DE DNA EXTRAÍDOS DE SEDIMENTOS DE MANGUE CONTAMINADO COM PETRÓLEO

>P3D12

GAGCCCCGCTACCATTAGCTAGTTGGTGGGGTAATGGCCCCACCAAGGCCATGATGGTTAGCTGGTCTGAGAGGATGGC
CAGCCACACTGGAAGTACGATACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGCACAATGGGGGCAACC
CTGACGCAGCAACGCCGCTGAGTGATGAAGGCCTTCGGGTCTGTAAGCTCTGTCAGATGGGAAGTGACGGTACCGTC
AAAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTGTTCCGAATCATTGGGCG
TAAAGAGCGTG TAGGCGGTGGGCAAGTCAGATGTGAAATCCCTTGGCTTAACCAAGGAAGTGCATTTGAAACTGGCC
6GGCTTGAGTACGGGAGGGGGAAGTGAATTCCCGGTGTAGAGGTGAAATTGTAGAT

>P3F12

GAGCCCCGCTACCATTAGCTAGTTGGTGGGGTAATGGCCCCACCAAGGCCATGATGGTTAGCTGGTCTGAGAGGATGGC
CAGCCACACTGGAAGTACGATACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGCACAATGGGGGCAACC
CTGACGCAGCAACGCCGCTGAGTGATGAAGGCCTTCGGGTCTGTAAGCTCTGTCAGATGGGAAGTGACGGTACCGTC
AAAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTGTTCCGAATCATTGGGCG
TAAAGAGCGTG TAGGCGGTGGGCAAGTCAGATGTGAAATCCCTTGGCTTAACCAAGGAAGTGCATTTGAAACTGGCC
GGCTTGAGTACGGGAGGGGGAAGTGAATTCCCGGTGTAGAGGTGAAATTGTAGAT

>p2g5

GAGCCCCGCTCCTATCAGCTAGTTGGTGGGGTAACGGCCCCACCAAGGCTATGACGGGTAGCTGGTCTGAGAGGATGGT
CAGCCACACTGGGACTGAGACACTGCCCAGACTCCTACGGGAGGCTGCAGTCGAGAATCTTTTGCAATGCCCCGACAGG
GTGACAATGCGACGCCGCTGAGTGATGAAGGCCTTCGGGTCTGTAAGCTCTGTCAGATGGGAAGTGACGGTACCGTC
AAAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTGTTCCGAATCATTGGGCG
TAAAGAGCGTG TAGGCGGTGGGCAAGTCAGATGTGAAATCCCTTGGCTTAACCAAGGAAGTGCATTTGAAACTGCCC
GGCTTGAGTACGGGAGGGGGAAGTGAATTCCCGGTGTAGAGGTGAAATTGTAGAT

>P1G01

GAGCCCTTGTACTATTAGCTAGTTGGTGAGATAACAGCTCACCAAGGCGACGATGGTTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGTAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATATTGCACAATGGGGGAAGCC
CTGACGCAGTGACGCCGCTGAGTGATGAAGGCTTTTCAGTCTGTAAGCTCTGTGAGGTGGAAAGTGACGGTACCACC
AGAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGTGTTGTCCGAATAATTGGGCG
TACAGGGCGTG TAGGCGGTAGGCAAGTCAGATGTGAAATCCCTTGGCTTAACCAAGGAAGTGCATTTGAAACTGCTT
GACTAGAGTACAGGATGGGAAAGTGGATTCCCGGTGTAGAGGTGAAATTATTGAT

>P2E03

GAGCCCCGTGTACCATTAGCTTGTGGTAGGGTAATGGCCTACCAAGGCAACGATGGTTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAAGTACGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGCACAATGGGGGAAACC
CTGACGCAGCGACGCCGCTGAGTGATGAAGGCCTTCGGGTCTGTAAGCTCTGTGAGGTGGGAAGTGACGGTACCACC
AAAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCGAGCGTTGTTCCGAATTATTGGGCG
TAAAGAGCGTG TGGGCGGTTGAACAAGTCAGATGTGAAATCCCTCGGCTCAACCGAGGAAGCGCATTCGAAACTGTTT
AGCTTGAGTACGGGAGAGGGGAGCGGAATTCCCGGTGTAGAGGTGAAATTGTAGAT

>P3E11

GACCCGGTACCATTAGCTTGTGGTAGGGTAATGGCTTACCANAGGAACGAAGTTTAGCTGGCCTGAGAGGATGATCA
GCCACACTGGAAGTACGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGCACAAGGGGAAACACA
ACAACAGCGACGCCGCTGAGTGATGAAGGCCTTCGGGTCTGTAAGCTCTGTGAGGTGGGAAGTGCCGGTACCACAA
AGGAAGCACC GGCTAACTCCGTCCCAGCAGCCGCGGTAATACGGAGGGTGCGAGCGTTGTTCCGAATTATTGGGCGTA
AAGAGCTTGT TAGCGGTTTGCAAGTCAGATGTGAAAGCCCTCGGCTCAACCGAGGAAGTGCATTTGAAACTGCCGGAC
TAGAGTACGGGAGAGGAGAGCGGAATTCCCGGTGTAGAGGTGAAATTGTAGAT

>P2C03

GAGCCCCGCTCCCATTAGCTTGTGGTGGGGTAATGGCCTACCAAGGCTACGATGGGTAGCTGGTCTGAGAGGATGAT
CAGACACACTGGAAGTACACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGCACAATGGGGGAAACC
CTGACGCAGCAACGCCGCTGAGTGATGAAGGCCTTCGGGTCTGTAAGCTCTGTCAAGTGGGAAGTGACGGTACCACC
GAAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTATTCCGATTTACTGGGCG
TAAAAGCGCGCTAGCGGCCGTTTAAAGTCAGATGTGAAAGCCCCGGGCTCAACCCCGGAAGTGCATTTGATACTATT
GGCTTGAGTATGGGAGAGGATAGTGAATTCCCTGGTGTAGAGGTGAAATTGTAGAT

>p2e11

GAGCCCCGCTCCCATTAGCTTGTGGTGGGGTAATGGCCTACCAAGGCTACGATGGGTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAAGTACACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGCACAATGGGGGAAACC
CTGACGCAGCAACGCCGCTGAGTGATGAAGGCCTTCGGGTCTGTAAGCTCTGTCAAGTGGGAAGTGACGGTACCACC
GAAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTATTCCGATTTACTGGGCG
GAAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTATTCCGATTTACTGGGCG

TAAAGCGCGCTAGGCGGCCGTTTAAAGTCAGATGTGAAAGCCCGGGGCTCAACCCCGGAAGTGCATTTGATACTATTC
GGCTTGAGTATGGGAGAGGGAAGTGAATTTCCTGGTGTAGAGGTGAAATTGTAGAT

>p1g12

GAGCCCGCGTCCCATTAGCTTGTTGGTGGGGTAATGGCCTACCAAGGCGACGATGGGTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAAGTACACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGAGTGATGAAGGCTTTCCGGTTCGTAAAGCTCTGTCAAGTGGGAAGTGACGGTACCACT
GAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTATTTCGGATTTACTGGGCG
TAAAGCGCGCTAGGCGGCCGTTTAAAGTCAGATGTGAAAGCCCGGGGCTCAACCCCGGAAGTGCATTTGATACTATTC
GGCTTGAGTATGGGAGAGGGAAGTGAATTTCCTGGTGTAGAGGTGAAATTGTAGAT

>p2b2

GAGCCCGCGTCCCATTAGCTTGTTGGTGGGGTAATGGTCTACCAAGGCAACGATGGGTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGTAACTGACACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGAGTGATGAAGGCTTTCCGGTTCGTAAAGCTCTGTCAAGTGGGAAGTGACGGTACCACT
AAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTATTTCGGATTTACTGGGCG
TAAAGCGCGCTAGGCGGCCGTTTAAAGTCAGATGTGAAAGCCCGGGGCTCAACCCGGAAGTGCATTTGATACTATTC
GGCTTGAGTATGGGAGAGGGAAGTGAATTTCCTGGTGTAGAGGTGAAATTGTAGAT

>p3h8

GAGCCCGCGTCCCATTAGCTTGTTGGTGGGGTAATGGTCTACCAAGGCAACGATGGGTAGCTGGTCTGAAAGGATGAT
CAGCCACACTGGAAGTACACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGAGTGATGAAGGCTTTCCGGTTCGTAAAGCTCTGTCAAGTGGGAAGTGACGGTACCACT
AAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTATTTCGGATTTACTGGGCG
TAAAGCGCGCTAGGCGGCCGTTTAAAGTCAGATGTGAAAGCCCGGGGCTCAACCCGGAAGTGCATTTGATACTATTC
GGCTTGAGTATGGGAGAGGGAAGTGAATTTCCTGGTGTAGAGGTGAAATTGTAGAT

>P3C12

GAGCCCGCGTACCATTAGCTTGTTGGTAGGGTAATGGCCTACCAAGGCTTCGATGGTTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAAGTACACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGAGTGATGAAGGCTTTCCGGTTCGTAAAGCTCTGTCAAGTGGGAAGTGACGTTACCACT
GAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTATTTCGGAATTATTGGGCG
TAAAGGGCGCGTAGGCGGTTCTTACGTCAGATGTGAAAGCCCGGGGCTTAACCCCGGAAGTGCATTTGAAACTAAGG
GACTTGAGTATGGGAGAGGGAAGTGAATTTCCTGGTGTAGAGGTGAAATTGTAGAT

>p3e7

GAGCCCGCGTACCATTAGCTTGTTGGTAGGGTAATGGCCTACCAAGGCTTCGATGGTTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAAGTACACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGAGTGATGAAGGCTTTCCGGTTCGTAAAGCTCTGTCAAGTGGGAAGTGACGGTACCACT
GAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTATTTCGGAATTATTGGGCG
TAAAGGGCGCGTAGGCGGTCTCTTAAAGTCAGATGTGAAAGCCCGGGGCTTAACCCCGGAAGTGCATTTGAAACTAAGG
GACTTGAGTATGGGAGAGGGAAGTGAATTTCCTGGTGTAGAGGTGAAATTGTAGAT

>p3b6

GAGCCCGCGTACCATTAGCTTGTTGGTAGGGTAATGGCCTACCAAGGCTTCGATGGTTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGTAACTGACACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGAGTGATGAAGGCTTTCCGGTTCGTAAAGCTCTGTCAAGTGGGAAGTGACGGTACCACT
GAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTATTTCGGAATTATTGGGCG
TAAAGGGCGCGTAGGCGGTCTTTAAAGTCAGATGTGAAAGCCCGGGGCTTAACCCCGGAAGTGCATTTGAAACTAAGG
GACTTGAGTATGGGAGAGGGAAGTGAATTTCCTGGTGTAGAGGTGAAATTGTAGAT

>P2E04

GAGCCCGCGTACCATTAGCTTGTTGGTAGGGTAACGGCCTGCCAAGGCTGCGATGGTTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAAGTACACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGAGTGATGAAGGCCCTCCGGTTCGTAAAGCTCTGTCAAGTGGGAAGTGACGGTACCACT
GAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTATTTCGGAATTATTGGGCG
TAAAGGGCGCGTAGGCGGTTCTTAAAGTCAGATGTGAAAGCCCGGGGCTTAACCCCGGAAGTGCATTTGAAACTAAGG
AACTTGAGTATGGGAGAGGGAAGTGAATTTCCTGGTGTAGAGGTGAAATTGTAGAT

>P3C03

GAGCCCGCGTCCCATTAGCTTGTTGGTAGGGTAATGGCCTACCAAGGCTACGATGGGTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAAGTACACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGAGTGATGAAGGCTTTCCGGTTCGTAAAGCTCTGTCAAGTGGGAAGTGACGGTACCACT
GAAGGAAGCACCGGCTAACTCCGTGCCAAAAGCCGAGGTAATACGGGGGGTGCAAACGTTATTTCGGAATTACTGGGCG
TAAAGGGCGCGCAGGCGTCCCTTAAAGTCAGGCGTGAAAGCCCGGGGCTAAACCCCGGAAGTGCCTCCGAAACTAAGG
GACTCGAGTACGGGAGAGGGAAGTGAATTTCCTGTTGTAGAGGTGAAATTGTAGAT

>p1g9
GAGTCCGCGTGCCATTAGCTTGTTGGTGGGGTAATGGCCTACCAAGGCGACGATGGCTAGCTGGTCTGAGAGGATGGC
CAGCCACACTGGAACGTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGCAACGCCGCTGAGTGATGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGTGGGAAAGTGACGGTACCACT
GGAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTATTTCGGAATTACTGGGCG
TAAAGGGCGCGTAGGCGGCTCCTTAAGTCAGGCGTGAAAGCCCGGGGCTCAACCCCGGAAGTGC GTTTGAAACTAGGG
AGCTTGAGTACGGGAGAGGGAAGTGAATTCTTGGTGTAGAGGTGAAATTGTAGAT

>P1H06
GGGCCCCGCGTACCATTAGCTTGTTGGTGGGGTAACGGCCTACCAAGGCGACGATGGTTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAACGTGGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAAGC
CTGACGCAGCAACGCCGCTGAGTGATGAAGGATTTTCGGTTCGTAAAGCTCTGTGAGTGGGAAAGTGACGGTACCACT
GAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTATTTCGGAATTTACTGGGCG
TAAAGGGCGCGCAGGCGGCTGATAAGTCAGATGTGAAAGCCCGGGGCTCAACCCCGGAAGTGCATTTGAAACTGTCA
GGCTTGAGTATGGTAGAGGGAAGTGAATTCCAGGTGTAGCGGTGAAATGGTAGAT

>p1a3
GGGCCCCGCGCACCATTTAGCTTGTTGGTGAAGTAACGGCTCACCAGGGCTACGATGGTTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAACGTGAAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGCAACGCCGCTGAGTGAAGAAGGCTTTTCGGGTCGTAAAGCTCTGTCAAGTGGGAAATGATGGTACCACT
GGAGGAAGAACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTTCAAGCGTTGTTCGGAATTACTGGGCG
TAAAGGGCGTG CAGGCGGGTTGGTAAGTCAGATGTGAAAGCCCTGGGCTCAACCCAGGAAGTGCATTTGAAACTGCCA
TCCTTGAGTATGGGAGAGGAGAGTGAATTCCCAATGTAGAGGTGAAATTGTAGAT

>P1E06
GGGCCCCGCGCACCATTTAGCTTGTTGGTGAAGTAACGGCTCACCAGGGCTACGATGGTTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAACGTGAAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGCAACGCCGCTGAGTGAAGAAGGCTTTTCGGGTCGTAAAGCTCTGTCAAGTGGGAAATGATGGTACCACT
GGAGGAAGAACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTTCAAGCGTTGTTCGGAATTACTGGGCG
TAAAGGGCGTG CAGGCGGGTTGGTAAGTCAGATGTGAAAGCCCTGGGCTCAACCCAGGAAGTGCATTTGAAACTGCCA
TCCTTGAGTATGGGAGAGGAGAGTGAATTCCCAATGTAGAGGTGAAATTGTAGAT

>p2c11
GGGCTCGCGTACCATTAGCTTGTTGGTAGGGTAACGGCTTACCAAGGCTACGATGGTTAGCTGGTTTGAGAGGATGAT
CAGCCACACTGGAACGTGAAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGAAACGCCGCTGAGTGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGTGGGAAATGATGGTACCACT
GGAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTTCAAGCGTTGTTCGGAATTACTGGGCG
TAAAGAGCGTG TAGGCGGTTTGGTAAGTCAGATGTGAAAGCCCTGGGCTCAACCCAGGAAGTGCATTTGATACTGCCT
CACTAGAGTATGGGAGAGGAGAGTGAATTCCCAGTGTAGAGGTGAAATTGTAGAT

>p3a2
GGGCTCGCGTACCATTAGCTTGTTAGGTAGGGTAACGGCTTACCAAGCTACGATGGTTAGCTGGTTTGAGAGGATGAT
CAGCCACACTGGAACGTGAAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGCAACGCCGCTGAGTGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGTGGGAAATGATGGTACCACT
GGAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTGTTCGGAATTATTGGGCG
TAAAGAGCGTG TAGGCGGTTTGGTAAGTCAGATGTGAAAGCCCTGGGCTCAACCCAGGAAGTGCATTTGATACTGCCT
CACTAGAGTATGGGAGAGGAGAGTGAATTCCCAGTGTAGAGGTGAAATTGTAGAT

>P3D05
GGGCTCGCGTGCCATTAGCTTGTTGGTGGGGTAATGGCCTACCAAGGCAACGATGGCTAGCTGGTCTGAGAGGATGGC
CAGCCACACTGGAACGTGAAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
TTGACGCAACAACGCCGCTGAGTGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGTGGGAAATGATGGTACCACT
GGAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTGTTCGGAATTACTGGGCG
TAAAGGGCGTG TAGGCGGTTTGGTAAGTCAGATGTGAAAGCCCTGGGCTCAACCCGGAAGTGCATTTGAAACTGCCA
TTCTTGAGTATGGG~GAGGAGAGTGAATTCCCAGTGTAGAGGTGAAATTGTAGAT

>P2H04
GGGCCCCGCGTCCCATTAGCTTGTTGGTGGGGTAATGGCCTACCAAGGCGACGATGGGTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGCACTGAAACACGGGCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGAAACC
CTGACGCAGCAACGCCGCTGAGTGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGTGGGAAATGATGGTACCACT
GGAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTGTTCGGAATTACTGGGCG
TAAAGAGCGTG TAGGCGGTTTGGTCAGTCGAATTGTGGAACCCCTTGGCTTCCAAACGGGGGTGCATTTGGAAGTGCCA
TTCTTGAGTATGGGAGAGGAGAGTGAATTCCCAGTTTAGAGGTGAAATTGTAGAT

>P2G08
GGGCCCCGCGTACCATTAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGACGATGGTTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAACGTGGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTGC GCAATAGGGGAAACC
CTGACGCAGCAACGCCGCTGAGTGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGTGGGACATGATGGTACCGCT

AGAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTCTACGAGGGTGCAAGCGTTGTTCCGGAATTACTGGGCG
TCCAGAGCGTGTAGGCGGTTTGGTAAGTCAGATGTGAAAGCCCTTGGCTCAAACCCGGGAGTGCATTTGAAACTGCCA
TGCTAGAGTATGGGAGAGGAGAGTGGTATTTCCAGTGTAGAGGTGAAATTGTAGAT

>P1G10

GGGCTCGCGTACCATTAGCTTGTTGGTAGGGTAATGGCCTACCAAGGCTACGATGGTTAGCTGGTCTGAGAAGATGAT
CAGCCCACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTGCGCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGAGTGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGTGGGAAATGATGGTACCACT
AGAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTACTACGAGGGTGCGAGCGTTGTTCCGGAATTATTGGGCG
TAAAGAGCGTGTAGGCAGCTTGGTAAGTCAGATGTGAAAGCCCAGGGCTTAACCTGGAAGTGCATTTGAAACTGCCT
AGCTTGAGTATGGGAGAGGAGAGTGAATTCCCAGTGTAGAGGTGAAATTTCGTAGAT

>P1F06

GGGCTCGCGTCCCATTAGCTTGTTGGTGGGTAATGGCCTACCAAGGCTACGATGGGTAGCTGGTCTGAGAGGATGAT
CAGCCCACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGGATTTGCGCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGAGTGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGCGGGAAATGATGGTACCGCT
AGAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAAACGAGGGTGCGAGCGTTGTTCCGGAATTACTGGGCG
TAAAGCGCGTGAGGCGGCTTGGTAAGTCAGATGTGAAAGCCCAGGGCTCAACCTGGAAGTGCATTTGATACTGCCA
TGCTTGAGTATGGGAGAGGGGAGTGAATTCCCAGTGTAGAGGTGAAATTTCGTAGAT

>p3b5

GGGCTCGCGTACCATTAGCTTGTTGGTAGGGTAATGGCCTACCAAGGCGACGATGGTTAGCTGGTCTGAGAGGATGAT
CAGCCCACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTGCGCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGAGTGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGTGGGGAATGATGGTACCACT
GGAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAAACGAGGGTGCAAGCGTTGTTCCGGAATTACTGGGCG
TAAAGAGCGTGTAGGCGGTTTGGCAAGTCAGATGTGAGAGCCCCGGGCTCAACCCGGAAGTGCATTTGAAACTGCCA
TACTTGAGTATGGGAGAGGAGAGTGAATTCCCAGTGTAGAGGTGAAATTTCGTAGAT

>p1e5

GGGCTCGCGTACCATTAGCTTGTTGGTAGGGTAATGGCCTACCAAGGCTACGATGGTTAGCTGGTCTGAGAAGATGGC
CAGCCCACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGCGCAATGGGGGAAACC
CTGACGCAGCGACGCCGTGTGAGTGAAGAAGGACTTCGGGTCGTAAAGCTCTGTCAAGTGGGAAATGAAGGTACCACT
AAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAAACGAGGGTGCAAGCGTTATTTCGGATTTATTGGGCG
TAAAGGGCGTGTAGGCGGTTGGGATAGTCAGATGTGAAAGCCTTCTGCTCAACAGAAGAAGTACGTCTGAAACTGCCC
AACTTGAGTACGAGAGAGGAAAGTGAATTCCCAGTGTAGAGGTGAAATTTCGTAGAT

>p3f1

GGGCTCGCGTACCATTAGCTTGTTGGTAGGGTAATGGCCTACCAAGGCTACGATGGTTAGCTGGTCTGAGAAGATGGC
CAGCCCACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGCGCAATGGGGGAAACC
CTGACGCAGCGACGCCGTGTGAGTGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGTGGGAAATGAAGGTACTACT
AAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAAACGAGGGTGCAAGCGTTATTTCGGATTTATTGGGCG
TAAAGAGCGTGTAGGTGGTTGGGATAGTCAGATGTGAAAGCCTTCTGCTCAACAGAAGAAGTACGTCTGAAACTGCCC
AACTTGAGTACGAGAGAGGAAAGTGAATTCCCAGTGTAGAGGTGAAATTTCGTAGAT

>p1h10

GGGCTCGCGTACCATTAGCTTGTTGGTAGGGTAATGGCCTACCAAGGCTGCGATGGTTAGCTGGTCTGAGAAGATGGC
CAGCCCACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTAAGGAATCTTGCGCAATGGGGGAAACC
CTGACGCAGCGACGCCGTGTGAGTGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGTGGGAAATGAAGGTACCACT
AAAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAAACGAGGGTGCAAGCGTTATTTCGGATTTATTGGGCG
TAAAGAGCGTGTAGGCGGTTGGGATAGTCAGATGTGAAAGCCTTCTGCTCAACAGAAGAAGTACGTCTGAAACTGCCC
AACTTGAGTACGAGAGAGGAAAGTGAATTCCCAGTGTAGAGGTGAAATTTCGTAGAT

>p2d10

GGGCTCGCGTACCATTAGCTTGTTGGTAGGGTAATGGCCTACCAAGGCTACGATGGTTAGCTGGTCTGAGAAGATGGC
CAGCCCACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGCGCAATGGGGGAAACC
CTGACGCAGCGACGCCGTGTGAGTGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGTGGGAAATGAAGGTACCACT
AGAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAAACGAGGGTGCAAGCGTTATTTCGGATTTATTGGGCG
TAAAGAGCGTGTAGGCGGTTAGGTTAGCCAGATGTGAAAGCCTTCTGCTCAACGGAAGAAGTACATCTGAAACTGCCT
AACTTGAGTACGAGAGAGGAAAGTGAATTCCCAGTGTAGAGGTGAAATTTCGTAGAT

>p1c9

GGGCTCGCGTACCATTAGCTTGTTGGTAGGGTAACGGCCTACCAAGGCAACGATGGTTAGCTGGTCTGAGAAGACAAC
CAGCCCACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGCGCAATGGGGGAAACC
CTGACGCAGCGACGCCGCTGGGTGAAGAAGGCCTTCGGGTCGTAAAGCCCTGTCAAGTGGGAAATGAAGGTACCACT
AGAGGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAAACGAGGGTGCTAGCGTTATTTCGGATTTATTGGGCG
TAAAGAGCGTGTAGGCGGTTGGGTTAGTCAGATGTGAAATCCTTCTGCTCAACGGAAGAAGTACGTCTGAAACTGCCC
AACTTGAGTACGAGAGAGGAAAGTGAATTCCCAGTGTAGAGGTGAAATTTCGTAGAT

>p1e2
GGGCTCGCGTACCATTAGCTTGTGGTAGGGTAACGGCCTACCAAGGCAACGATGGTTAGCTGGTCTGAGAAGACAAC
CAGCCACACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGC GCAATGGGGGAAACC
CTGACGCAGCGACGCCGCGTGGGTGAAGAAGGCCTTCGGGTCGTAAAGCCCTGTCAAGTGGGAATTGAAGGTACCACT
AGAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCTAGCGTTATTTCGGATTTATTGGGCG
TAAAGAGCGTGTAGGCGGTTGGGTTAGTCAGATGTGAAATCCTTCTGCTCAACGGAAGAAGTACGTCTGAAACTGCC
AACTTGAGTACGAGAGAGGAAAGTGAATTCCCAGTGTAGAGGTGAAATTCGTAGAT

>p1c12
GGGCTCGCGTACCATTAGCTTGTGGTAGGGTAACGGCCTACCAAGGCGACGATGGTTAGCTGGTCTGAGAAGACAAC
CAGCCACACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGC GCAATGGGGGAAACC
CTGACGCAGTGACGCCGCGTGGGTGAAGAAGGCCTTCGGGTCGTAAAGCCCTGTCAAGTGGGAATTGAAGGTACCACT
AGAGGAAGCACC GGCTACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCCAGCGTTATTTCGGATTTATTGGGCGT
AAAGAGCGTGTAGGCGGTTGGGTTAGTCAGATGTGAAATCCTTCTGCTCAACGGAAGAAGTACATCTGAAACTGCCCA
ACTTGAGTACGAGAGAGGAAAGTGAATTCCCAGTGTAGAGGTGAAATTCGTAGAT

>p1a8
GGGCTCGCGTGCCATTAGCTTGTGGTAGGGTAATGGCCTACCAAGGCTACGATGGTTAGCTGGTCTGAGAAGACAAC
CAGCCACACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGC GCAATGGGGGAAACC
CTGACGCAGCGACGCCGCGTGAGTGAAGAAGGCCTTCGAGTCGTAAAGCTCTGTCAAGTGGGAATTGAAGGTACCACT
GGAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGCAGCGTATTTCGGATTTATTGGGCGTAA
AGAGCGTGTAGGCGGTTGGGTTAGTCAGATGTGAAATCCTTCTGCTCAACAGAAGAAGTACGTCTGAAACTGCCCAGC
TTGAGTACGAGAGAGGAAAGTGAATTCCCAGTGTAGAGGTGAAATTCGTAGAT

>p2a12
GGGCCCCGCGTACCATTAGCTTGTGGTAGGGTAATGGCCTACCGAGGCTACGATGGTTAGCTGGTCTGAGAAGATGAC
CAGCCACACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGC GCAATGGGGGAAACC
CTGACGCAGCGACGCCGCGTGAGTGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGTGGGAATTGAAGGTACCACT
AGAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGCAATACGGAGGGTGCGAGCGTTATTTCGGATTTATTGGGCG
TAAAGAGCGTGTAGGCGGTTGGGTAAGTCAGGTGTGAAATCCTTTTCTGCTCAACAAAAGAAGTGCATCTGAAACTGCTC
AACTTGAGTACGGGAGAGGAGAGTGAATTCCCAGTGTAGAGGTGAAATTCGTAGAT

>p2d8
GGGCCCCGCGTACCATTAGCTAGTTGGTAGGGTAAGGGCCTACCGAGGCTACGATGGTTAGCTGGTCTGAGAAGATGAT
CAGCCACACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGC GCAATGGGGGAAACC
CTGACGCAGCGACGCCGCGTGAGTGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGTGGGAATTGAAGGTACCACT
AGAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCGAGCGTTATTTCGGATTTATTGGGCG
TAAAGAGCGTGTAGGCGGTTGGGTAAGTCAGATGTGAAATCCTTTCTGCTCAACGAGAGAAGTGCATCTGAAACTGCTC
AACTTGAGTACGGGAGAGGAGAGTGAATTCCCAGTGTAGAGGTGAAATTCGTAGAT

>p1a4
GGGCTCGCGTACCATTAGCTTGTGGTAGGGTAAGGCCTACCAAGGCTACGATGGTTAGCTGATCTGAGAAGATGAT
CAGCCACACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGC GCAATGGGGGAAACC
CTGACGCAGCGACGCCGCGTGAGTGAAGAAGGCCTTTGGGTCGTAAAGCTCTGTCAAGTGGGAATTGAAGGTACCACT
AAAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTATTTCGGATTTATTGGGCG
TAAAGAGCGTGTAGGCGGTTGGGTTAGTCAGATGTGAAATCCTTCTGCTCAACGGAAGAAGTGCATCTGAAACTGCCT
GACTTGAGTACGAGAGAGGAAAGTGAATTCCCAGTGTAGAGGTGAAATTCGTAGAT

>P2F03
GGGCTCGCGTACCATTAGCCCGTTGGTAGGGTAAGGCCTACCAAGGCTAGATGGTTAGCTGATCTGAGAAGATGATC
AGCCACACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAGGAATCTTGC GCAATGGGGGAAACC
TGACGCAGCGACGCCGCGTGAGTGAAGAAGGCCTTTGGGTCGTAAAGCTCTGTCAAGTGGGAATTGAAGGTACCACTA
AAGGAAGCACC GGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTATTTCGGATTTATTGGGCGT
AAAGAGCGTGTAGGCGGATGGGTTAGTCAGATGTGAAATCCTTCTGCTCAACGGAAGAAGTGCATCTGAAACTGCCTG
ACTTGAGTACGAGAGAGGAAAGTGAATTCCCAGTGTAGAGGTGAAATTCGTAGAT

>p1b11
GGGCTCGCGTACCATTAGCTTGTGGTAGGGTAATGGCCTACCAAGGCTGCGATGGTTAGCTGGTCTGAGAAGATGGC
CAGCCACACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTAAGGAATCTTGC GCAATGGGGGAAACC
CTGACGCAGCGACGCAGTGTAGTGAAAAAGGACTTCGGGTCGTAAAGCTCTGTCAAGTGGGAATTGAAGGTACCCCA
AGAGAAAGTCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACATAAGGAGCAAGCGTTATTTCGGATTTATTGGGCG
TAAAGAGCGTGTAGGCGGTTGGGATAGTCAGGTGTGAAATTAAGTGGCTCAACCTGTTAAGGTCATCTGAAACTTTAA
GACTTGAGGATAGAAGAGGAAAGTGAATTCCCAGTGTAGCGGTGGAATGCGTAGAT

>p2a5
GAGTTTGC GTACCATTAGCTTGTGGTGGGGTAAGGGCCTACCAAGGCTGTGATGGTTAGCTGGTCTGAGAGGATGAT
CAGTCACACTGGAAGTGGACACGGTCCAGACTCCTACGAGAGGCAGCAGTGAGGAATTTTGC GCAATGGGGGCAACC
CTGACGCAGCAACGCCGCGTGAGTGAAGAAGGCCTTTGGGTCGTAAAGCTCTGTCAACGGGAAGTGACGGTACCGGT

GGAGGGAGCGCCGGCTAACTCCGTGCCAGCAGCCGCGGTAACACGGGGGGCGCAAGCGTTATTTCGGAATTATTGGGCG
TAAAGGGCGCGTAGGCGGTTCTGTGCGTCAGATGTGAAAGCCCGGGGCTCAACCTGGAAGTGCATTTGAAACAGCAG
AACTTGAGTACGGGAGAGGAAAGCGGAATTCTTGGTGTGAGGTGAAATTCGTAGAT

>P1A10

CGGGCTGCGCTCCATTAGCTTGTGGTAGGGTAACGGCCTACCAAGGCGACGATGGATAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAAGTGAACACGGTCCAGACTCCACGGGAGGCAGCAGTTGGGAATATTGGACAATGGGGGAAACC
CTGATCCAGCAATGCCGCGTGAGTGATGAAGGCCTTCGGGTCGTAAAGCTCTTTCAGCGGGGACGTGACGGTACCCGC
AGAAGAAGCACCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGCAAGCGTTGTTTCGGAATTATTGGGCG
TAAAGCGCGTGTAGGCGGTTGCTTAAGTCTGATGTGAAATCCCGGGGCTTAACCCGGAAGTGCATTGGATACTGGGC
AACTAGAGGGCGAGAGAGGAAGTGAATTCTTGGTGTAGGGGTGAAATCCGTAGAT

>P3F07

GGGCCCCGCTCCCATCAGCCTGTTGGTGAGGTAACGGCTCACCAAGGCTAAGACGGGTAGCTGGTCTGAGAGGATGTC
CAGCCACACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTTGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCGACGCCGCGTGAGTGATGAAGGCCCTCGGGTCGTAAAGCTCTGTGCGACCGGAAACTGACGGTACCCG
TTAGTAAAGCACCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTGTTTCGGAATCATTGGTCG
TAAAGAGCGTGTAGGCGGGATGGCATGTCGATGTGAAAGCCCTCGGCTTAACCGAGGAATTGCATCCGAAACTGTGCG
TTCTTGAGTACGGGAGAGGAAGCGGAATTTCCGGTGTAGAGGTGAAATTCGTAGAT

>p2a7

GAGTCTGCGACCTATCAGCTAGTTGGCGGGGTAACGGCCACCAAGGCGAAGACGGGTAGTTGGTCTGAGAGGATGAT
CAGCCACACTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGCACAATGGAGGAAACT
CTGATGCGACGACGCCGCGTGAGTGATGAAGGCCTTTAGGTTCGTAAAGCTCTGTAAGGAGGGAAGCTGACGGTACCTC
CAAGTAAGGACCGGCTAAATACGTGCCAGCAGCCGCGGTAATACGTATGGTCCAGCGTTGTCCGGAATAACTGGGCG
TAAAGGGTGTGTAGGTGATATATATGTCAGGTGTGAAATCCCTCGGCTTAACCGGGGCATTGCGCTTGAAACTGTATA
TCTAGAGCGTGTCAGGGGTAAGCGGAATTTCCGGAGTAGCGGTGAAATGCGTAGAT

>P1E04

GGGCTCGCGGCCGATTAGCTTGTGGTGAGGTAACGGCCTACCAAGGCGATGATCGGTAGCTGGTTTGAGAGAATGGT
CAGCCACACTGGGATTGAGATACGGCCAGACTCCTACGGGAGGCAGCAGTCTAGAAGCTTCGGCAATGGGGGCAACC
CTGACCGAGCGACGCCGCGTGCGGGATGAAGGCCTTCGGGTTGTAAACCGCTTTCGACAGGGAAGTTGACGGTACCTG
GAAAGAAGCCTCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGAGGCAAGCGTTAATCGGATTCTTGGGCG
TAAAGGGCGTGTAGGCGGTTCCCTTAAGTCGGGTGTTAAAGACCACGGCTCACCCGTGGGCCGCGCCCGATACTGAGG
GACTGGAGTGCAGGAGAGGCGAGCGGAATTTCCCGGTGTATCGGTGAAATGCGTAGAT

>p1e10

GGGCCCCGCGGGCTATTAGCTAGTTGGTGAGGTAACGGCCTACCAAGGCGATGATGGCTAGCCGGCCTGAGAGGGCGAC
CGGCCACACAGGGACTGAGACAAGGCCCTGTCCCTACGGGAGGCGCAGGCGGGAATCGTTCACAATGGGCGCAAGCC
TGATGACGCGACGCCGCGTGAGGGATGAAGGCCCTTCGGGTTGTAAACCTCTGCCGGGAGAGAAGCTGACGTTATCTCC
CAGGAAGCCTCGGCTAACTACGTGCCAGCAGCCGCGGTAACAGTAGGAGGCAAGCGTTGTCCGGAATTACTGGGCGT
AAAGGGTGTGTAGGCGGGACGGTAAGTCGGGCGTGAAATCCCACGGCTCAACCGTGGAATGCTCCCGATACTGCCGA
CCTTGAGTACGGGATAGGAGAGTGGAATCACGGTGTAGCGGTAAAATGCGTAGAT

>P1B05

GGGCTCGCGGCCCATTAGCTTGTGGTGAGGTAACGGCTCACCAAGGCTGCGATGGGTAGCCGACCTGAGAGGGTGAT
CGGCCACACTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTCGCAATGGGCGAAAGC
CTGACGGAGCGACGCCGCGTGCGTGATGAAGGCCCTCGGGTCGTAAAGCGCTGTGCGAGGGGAAGCTGACGGTACCCT
CCAGGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGGCTAGCGTTGCTCGGAATTATTGGGCG
TAAAGGGCGCGTAGGCGGCCGCTAAGTTGGTTGTGAAATCCCTCGGCTCAACCGAGGAATGCGACCAATACTGGCG
GGCTTGAGTGCAGGAGGGGAGAGCGGAATTTCCAGTGTAGCGGTGAAATGCGTAGAT

>p1c10

GGGCTCGCGGCCCATTAGCTTGTGGTGAGGTAACGGCTCACCAAGGCTGTGATGGGTAGCCGACCTGAGAGGGTGAT
CGGCCACACTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGAGGAATTTTCGCAATGGGCGAAAGC
CTGACGGAGCGACGCCGCGTGCGTGATGAAGGCCCTCGGGTCGTAAAGCGCTGTGCGAGGGGAAGCTGACGGTACCCT
CCAGGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGGCTAGCGTTGCTCGGAATTATTGGGCG
TAAAGGGCGCGTAGGCGGCCGCTAAGTTGGTTGTGAAATCCCTCGGCTCAACCGAGGAATGCGACCAATACTGGCG
GGCTTGAGTGCAGGAGGGGAGAGCGGAATTTCCAGTGTAGCGGTGAAATGCGTAGAT

>p3c8

GGGCTCGCGTACCATTAGCTAGTTGGCGGGGTAATGGCCACCAAGGTTATGATGGTTAGCCGGCCTGAGAGGGTGTC
CGGCCACACTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTGAGGAATATTGCGCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGATGATGAAGGCCTTCGGGTCGTAAATCCTGTGATGGGAAGTTGACGGTACCAT
CAAGGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGCGAGCGTTGTTTCGGAATTACTGGGCG
TAAAGGGCGCGTAGGCGGCCCTATAAGTTAGATGCTAAAGACCTCGGCTTAACCAAGGACCCGCATCTAATACTGTAG
GGCTAGAGTACAGGAGAGGAGAGTGGAATTTCCAGTGTAGCGGTGAAATGCGTAGAT

>p1h9
TGGCTCGCGGCCTATCAGCTTGTTGGTGAGGTAACGGCTCACCAAGGCGACGACGGGTAGCCGACCTGAGAGGGTGAC
CGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAAGCGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCGACGCCGCTGAGGGATGAAGGCCTTCGGGTCTGTAACCTCTGTTGACGGGGACGCTGACGGTACCCG
TCAGGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCAAGCGTTGTCCGGATTTATTGGGCG
TGAAGCGCGCTAGGCGGTTTATTAAAGTCGGCTGTGAAAGTGCTGGGCTCAACCCAGCAATGTCAGTCGATACTGATG
GACTTGAGTCCGGGAGAGGGAAGTGGACTTGCTGGTGTAGCGGTGGAATGCGTAGAG

>p2f10
GGGCTCGCGGCCTATCAGCTAGTTGGTGAGGTAACGGCCACCAAGGCGACGACGGGTAGCCGACCTGAGAGGGTGAC
CGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAAGCGGGAATATTGCGCAATGGGCGAAAGC
CTGATGACGACGACGCCGCTGAGGGATGAAGGCCTTCGGGTCTGTAACCTCTGTTGATGGGGACGCTGACGGTACCCA
TCAGGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCAAGCGTTGTCCGGATTTACTGGGCG
TAAAGAGCGCTAGGTGGTCCATTAAAGTCGGCTGTGAAATCGCTGGGCTTAACCCAGCAATGTCAGTCGATACTGGTG
GACTCGAGTCCGGGAGAGGGAAGTGAATTGCTGGTGTAGCGGTGGAATGCGTAGAT

>p2b8
GGGCTCGCGGCCCATCAGCTTGTTGGTAGGGTAACGGCCTACCAAGGCAACGACGGGTAGCCGACCTGAGAGGGTGAC
CGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAAGTGGGAATATTGTCCAATGGGCGAAAGC
CTGAGACAGCGACGCCGCTGAGGGATGAAGGCCTTCGGGTCTGTAACCTCTGTTGACGGGGACGCTGACGGTACCCG
TCAGGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCAAGCGTTGTCCGGATTTACTGGGCG
TAAAGAGCAGCTAGGCGGTCTGCTAAGTCGGCTGTGAAAGCGCTGGGCTCAACCCAGCAATGTCAGTCGATACTGGTA
GACTTGAGTCCGGGAGAGGGAAGTGAATAGCTGGTGTAGCGGTGGAATGCGTAGAT

>P3H06
CGGCCCGCGGTCCATCAGCTAGTTGGTGGGTAATGGCCACCAAGGCGACGACGGATAGCCGACCTGAGAGGGTGAC
CGGCCACAGGGGACTGAGACACGGGGCCCACTCCTACGGGAGGCAGCAGTTAGGAATCTTCCACAATGGGCGAAAGC
CTGATGGAGCGACACTGCGTGAGGGATGAAGGCCTTCGGGTCTGTAACCTCTAAAACGGGGACGCTGACGGTACCCGTT
AGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTATCCGGAATCACTGGGCGTAA
AGGGCGCTAGACGGTCTGTTAAGTCGGGTGTGAAATCTCCTGACTCAATCAGGAGAGGTCGTTGATACTGAGGAGC
TTGAGGGCAGCAGAGGAAGTGGAATTCCCGGTGTAGCGGTGATATGCGTAGAT

>P2F05
GGGCCCGCGGCCCATCAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTAAC
CGGCCACACGGGGACTGAGACATGGTCCCGCTCCTACGGGAGCTAGCATCGGGGAATCTTGCGCAATGGGCGAAAGC
CCGACGCAGTGACGCCGCTGGGGGACGAAGGCCCTCGGGTCTGTAACCCCTGACAGGAGGGAAGCTGACGGTACCTC
CAAGGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCGAGCGTTGTCCGGAATTACTGGGCG
TAAAGGGCGCTAGGCGGTCCGTTAAGCGGAGGGTGAAAGCCACGGCTCAACCTTGGAATTGCTCTCCGAACCTGACG
GACTAGAGGCCGGGAGAGGGCAGTGGAATTGCCGGTGTAGCGGTGACATGCGTAAAT

>p1g3
GGGCCCGCGGCCCATCAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTAAC
CGGCCACACGGGGACTGAGACATGGTCCCGCTCCTACGGGAGCTAGCATCGGGGAATCTTGCGCAATGGGCGAAAGC
CCGACGCAGTGACGCCGCTGGGGGACGAAGGCCCTCGGGTCTGTAACCCCTGACAGGAGGGAAGCTGACGGTACCTC
CAAGGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCGAGCGTTGTCCGGAATTACTGGGCG
TAAAGGGCGCTAGGCGGTCCGTTAAGCGGAGGGTGAAAGCCACGGCTCAACCTTGGAATTGCTCTCCGAACCTGACG
GACTAGAGGCCGGGAGAGGGCAGTGGAATTGCCGGTGTAGCGGTGACATGCGTAAAT

>p3e12
GGGCCCGCGGCCCATCAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTAAC
CGGCCACACGGGGACTGAGACATGGTCCCGCTCCTACGGGAGCTAGCATCGGGGAATCTTGCGCAATGGGCGAAAGC
CCGACGCAGTGACGCCGCTGGGGGACGAAGGCCCTCGGGTCTGTAACCCCTGACAGGAGGGAAGCTGACGGTACCTC
CAAGGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCGAGCGTTGTCCGGAATTACTGGGCG
TAAAGGGCGCTAGGCGGTCCGTTAAGCGGAGGGTGAAAGCCACGGCTCAACCTTGGAATTGCTCTCCGAACCTGACG
GACTAGAGGCCGGGAGAGGGCAGTGGAATTGCCGGTGTAGCGGTGACATGCGTAAAT

>p1f3
GGGCCCGCGGCCCATCAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTAAC
CGGCCACACGGGGACTGAGACATGGTCCCGCTCCTACGGGAGCTAGCATCGGGGAATCTTGCGCAATGGGCGAAAGC
CCGACGCAGTGACGCCGCTGGGGGACGAAGGCCCTCGGGTCTGTAACCCCTGACAGGAGGGAAGCTGACGGTACCTC
CAAGGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCGAGCGTTGTCCGGAATTACTGGGCG
TAAAGGGCGCTAGGCGGTCCGTTAAGCGGAGGGTGAAAGCCACGGCTCAACCTTGGAATTGCTCTCCGAACCTGACG
GACTAGAGGCCGGGAGAGGGCAGTGGAATTGCCGGTGTAGCGGTGACATGCGTAAAT

>p1d2
GGGCCCGCGGCCCATCAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTAAC
CGGCCACACGGTGACTGAGACATGGTCCCGCTCCTACGGGAGCTAGCATCGGGGAATCTTGCGCAATGGGCGAAAGC
CCGACGCAGTGACGCCGCTGGGGGACGAAGGCCCTCGGGTCTGTAACCCCTGACAGGAGGGAAGCTGACGGTACCTC

CAAGGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCGAGCGTTGTCCGGAATTACTGGGCG
TAAAGGGCGCGTAGGCGGTCCGTTAAGCGGAGGGGTGAAAGCCACGGCTCAACCTTGGAATTGCTCTCCGAACGTACG
GACTAGAGGCCGGGAGAGGGCAGTGAATTGCCGGTGTAGCGGTGACATGCGTAAAT

>p2h1

GGACCCGCGGCCGATTAGCTTGGTTGGTAGGGTAATGGCCTACCAAGGCGACGATCGGTAGCCGGTCTGAGAGGATGAT
CGGCCACACTGGGACTGCGACATGGTCCAGACTCCTACGGGAGTTAGCATCAGGAATCTTGCGCAATGGGCGAAAGCC
TGACGCAGCGACGCCGCGTGGGTGATGAAGGCCTGCGGGTCGTAAAGCCCTGTCGGGAGGGAAGCTGACTGTACCTCC
GAGGAAGCCCCGGCTAATTACGTGCCAGCAGCCGCGGTAACACGTAAGGGGCCAGCGTTGTCCGGAATTACTGGGCGT
AAAGTGCACGTAGGTGGTGGGTAAGCGATAGGTGAAAGCCTCCGGCTCAACCGGGGAACGTCTATCGAACTGCTCA
GCTAGAGCACGGGAGAGGGCACTGGAATTCCTGGTGTAGCGGTGAAATGCGTATAT

>P2H07

GGTCCCGCTCTGATTAGCTTGGTTGGTAGGTAACGGCTCACCAAGGCGACGATCAGTAGCCGGCCTGAGAGGGTGTC
CGGCCACACTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTCGAGAAGCTTCGGCAATGGGCGAAAGC
CTGACCGAGCGACGCCGCGTGTGCGATGAAGGCCCTAGGGTCGTAAAGCACTGTCAGGAGGGAATCTGAAGGTACCTC
CAAGGAAGCACCGGCCCTAACTACGTGCAGCAGCCGCGGTAATACGTAGGGTGCAAACGTTGTACGGATTCACTGGGCTA
TCAAGGGTTCGTAGGCGGCTGGTAAGTCAGAAGGTGAAAGCCACTGGCTCAACCGTTGAATTGCCTCTGATACTGTCA
GTCTTGAGTCCGTCCGGGGATGGTGGGATTTCCTGGTGTAGCGGTGAAATGCGTAAAT

>P3B11

AAGCCTGCGTCCTATTAGCTAGTTGGTAAGGTAATGGCTTACCAAGGCTATGATGGGTAGCCGGCCTGAGAGGGTGAA
CGGCCACACTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCGACACCGTGTGGAGGATGAAGGTCTCGGATCGTAAACTCCTGTCAATTGGGGACCTGACGGTACCTA
ATAGGAAGTCTCGGCTAACTCCGTGCCAGCAGCCGCGGTAAGACGGAGGAGGCAAGCGTTGTCCGGAATCATTGGGCG
TAAAGAGCGCGTAGGCGGCTTTTCAAGTCATGTGTGAAAACCTTCAGGCTCAACTTGAGGATTGCATATGAAACTGAAA
AACTAGAGTACGGCAGAGGGAAGTGGGATTTCCTGGTGTAGCGGTGAAATGCGCAGAT

>p3g2

AAGCCTGCGTCCTATTAGCTAGTTGGTAAGGTAATGGCTTACCAAGGCTATGATGGGTAGCCGGCCTGAGAGGGTGAA
CGGCCACACTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCGACACCGTGTGGAGGATGAAGGTCTCGGATCGTAAACTCCTGTCAATTGGGGACCTGACGGTACCTA
ATAGGAAGTCTCGGCTAACTCCGTGCCAGCAGCCGCGGTAAGACGGAGGGGGCAAGCGTTGTCCGGAATCATTGGGCG
TAAAGAGCGCGTAGGCGGCTTTTCAAGTCATGTGTGAAAACCTTCAGGCTCAACTTGAAGATTGCATATGAAACTGGAA
AACTAGAGTACGGCAGAGGAAAGTGGGAATTCCTGGTGTAGCGGTGAAATGCGCAGAT

>P2H05

GAGCCTGCGTCCTATTAGTTTGGTTGGTAAGGTAATGGCTTACCAAGGCAATGATGGGTAGCCGGCCTGAGAGGGTGAC
CGGCCACACTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCGACACCGTGTGAAGGATGAAGGTCTTCGGATCGTAAACTTCTGTCAATTGGAGACCTGACGGTATTCA
ATAGGAAGTCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAAGACGGAGGAGGCAAGCGTTGTCCGGAATATTATGGGCG
TAAAGAGCGCGTAGGTGGTTTTCAAGTCATATGTGAAAACCTTCAGGCTCAACCTGAAGATTGTATATGAAACTGGAT
AACTTTAGTACAGCAGAGGAAAGTGGGAATTCCTGGTGTAGCGGTGAAATGTGCAGAT

>P3C10

GGACCCGCGGCCCATTTAGCTTGGTTGGTAGGTAACGGCTCACCAAGGCTATGATGGGTAGCCGGCCTGAGAGGGTGTA
CGGTCACACTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCGACGCCGCGTGGACGATGAAGGCCTTCGGGTTGTAAAGTCTGTCAAGGAGGACTGACGGTACCTCGG
AGGAAGCTCCGGCTAACTACGTGCCAGCAGACGCGGTAAGACGTAGGGGGCAAGCGTTGCCCGGAATCACTGGGCGTA
AAGAGCTCGTAGACGGTCTGTTTGGTTCGGATGTGAAAGCCCTCGGCTCAAACCGAGGATTGCATTCGATACTGGCAGG
CTAGAGTCTAGCAGAGGGAAGTGGGAATTCCTGGTGTAGCGGTGAAATGCGCAGAT

>p3g9

GATCCCGCGTCCCATTAGTTAGTTGGTAGGTAACGGCTCACCAAGGCAGCGATGGGTAGCCGGCCTTAGCGGGTGAC
CGGCCACACTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTCGAGAGGCTTCCACAATGGGGGAAACC
CTGATGGAGCGACGCCGCGTGGGTGATGAAGGTTTTTCGGATCGTAAAGCCCTTTCAGCAGGGAGCAGACGGTACCTGC
AAAGAAGCTCCGGCTAACTCCGTGCCAGCAGCCGCGGTAAGACGGAGGGAGCAAGCGTTGTTTCGGATTTACTGGGCGT
AAAGGGCGTCTAGGCGGGTTGGCTAGTCAGATGTGAAATCCCCCGGCTCAACCGGGGAACGTGCATTTGATACTGCCAG
TCTTGAGTACGGAAGAGGGAAGCGGAATTCCTGGTGTAGCGGTGAAATGCGTAGAT

>p1a1

GAGCCCGCGGCCTATTAGCTAGTTGGTGGGTAATGGCTCACCAAGGCGATGATGGGTAGCCGGCCTGAGAGGGCGAC
CGGCCACATTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATTTTTCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGAAGGATGAAGGCCTTCGGGTCGTAAAGCTCTGTCAAGATGGGAAGCTGACGGTACCAT
CAATGAAGTCCCGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGAGCAAGCGTTGTTTCGGATTTACTGGGCG
TAAAGAGCGCGTAGGCGGTACGGTAAGTCAGATGTGAAAACACCGGCTCAACCGGTTGGCTGCATTTAAGACTGCCT
GACTTGAGTTCAGGAGAGGGGAGTGGGAATTCCTAGTGTAGCGGTGGAATGCGTAGA

>p1d1

GAGCCCCGCGCCTATTAGCTTGTGGTAGGGTAATGGCCTACCAAGGCGATGATGGGTAGCCGGCCTGAGAGGGTGAC
CGGCCACATTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATTTTGC GCAATGGGCGAAAGC
CTGACGCAGCAACGCCCGGTGAGTGATGAAGGCCTTCGGGTCTGTAAGCTCTGTGAGATGGGAAGCTGACGGTACCAT
CAAGGAAGCTCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGAGCCAGCGTTGTTCCGATTCACTGGGCG
TAAAGAGCGCGTAGGCGGTTCCGTAAGTCAGATGTGAAAACCCAGGCTCAACCGGTTGGCTGCATTTGATACTGCCT
TACTTGAGTTCAGGAGAGGGGAGCGGAATTCCCAGTGTAGCGGTGGAATGCGTCGCT

>P2C05

GGACCTGCGCCCCATTAGCTAGTTGGTGAGGTAATAGCTCACCAAGGCGACGATGGGTAGCCGGCCTGAGAGGGTGAC
CGGCCACACTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGCGCAATGGGCGAAAGG
CTGACGCAGCAACGCCCGGTGGATGATGAAGCTCTAAGGAGCGTAAAGTCTGTAGGAGGGAAGTGACGGTACCTCC
AGAGAAAGCCCCGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGGCGAGCGTTGTCCGGAATTATTGGGCG
TAAAGGGCGCGTAGGCGGAATAGTAAGTCAGGGGTAAAACCATAGGCTCAACCTATGGCCTGCCTTTGAAACCGCTAT
TCTTGAGTGCGGAAGAGGGAAGTGGAATTCTTGGTGTAGCGGTGAAATGCGTAGAT

>P1B06

GGGCTCGCGTCCTATTAGGTAGTTGGTGAGGTAACGGCTCACCAAGCTGACGATGGGTAGCCGACCTGAGAGGGTGAC
CGGCCACACTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCCGGTGGAGGATGAAGTCTTCGGGGTGTAACCTCTTTCAGGGGGGAAGTGACGGTACCCCC
AGAAGAAGTCCGGCTAACCCCGTGCCAGCAGCCGCGGTAATACGGGGGGAGCAAGCGTTGTCCGGAATCACTGGGCG
TAAAGGGTGCGTAGGCGGAATGCAAGTCCGTGGTGAAACCTGATGGCTCAACTGTGAGCCTGCCCCGAAACTGCAT
TTCTTGAGTGCTGCAGAGGAGACCGGAATTCTTGGTGCAGCGGTGAAATGCGTAGAT

>p2h12

AGGCCGCGCTCCATTAGCTTGTGGTGAGGTAACGGCTCACCAAGGCTGCGATGGATAGCCGTCCTGAGAGGGTGAT
CGGACAAATTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGCGCAATGGGCGCAAGC
CTGACGCAGCAGCAGCCCGGTGGGTGATGAAGCTCCTTGGAGTGTAACCTCTGTCAGTGGGGAAGTGACGGTACCCGC
AGAGGAAGTCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGAGCTAGCGTTGTTCGGATTTACTGGGCG
TAAAGGGCGTGAGGTGGATTAGCAAGTCACATGTGAAAACCAAGTGGCTCAACCGCTGGCCTGCACTTGATACTGTTG
ATCTTGAGTACGGGAGAGGAAGCGGAATTCTTGGTGTAGCGGTGAAATGCGTAGAT

>P3A12

GGGCCTGCGGTTCCATTAGTTAAATGGTAGGGTAACGAGCTACCATGACTGCGATGGATAGCTGATCTGAGAGGGTGAT
CAGCCACACTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGCGCAATGGACGAAAGT
CTGACGCAGCAACGCCCGGTGGGGGATGAAGCACTATGGTGTGTAAACCTCTGTAGACAGGGAAGTGACGGTACCTGT
AAAGTAAGTCCGGACTTAACTGTGCCAGCAGCCGCGGTAATACGGGGGGAGCGAGCGTTGTCCGGAATTACTGGGCG
TAAAGGGCGGTGTAGGCGGCTTGACAAGTCAGATGTGAAAGCCCGCGGCTTAACCGCGGAAATGCATTTGAAACTGTTG
AGCTTGAGTGCGAGAGAGGGTAGTGGAATTCCCAGTGTAGCGGTGAAATGCGTAGAT

>p1e9

GGGCCTGCGGCGCATTAGTTAGTTGGTGAGGTGAGAGCTCACCAAGACGATGATGCGTAGCCGGACTGAGAGGTTGAA
CGGCCACATTGGGACTGAGAGACGGCCCAAACCTCCTACGGGAGGCAGCAGTAGGGAATTTTCGGCAATGGAGGAACT
CTGACCGAGCAACGCCCGGTGAATGACGAAGTACTTCGGTATGTAAAGTTCTTTATCAAGGAAGTGACGGTACTTGA
TGAATAAGCCCCCGCTAAGTATGTGCCAGCAGCCGCGGTAATACATAGGGGGCAAGCGTTATCCGGAATTATTGGGCG
TAAAGGGTGCGTAGGCGGTCATTTAAGTCTTGGTGTAAGTGCAAGGCTCAACCTGTGATGCTAAGGAACTGGATGA
CTTGAGTGTGATAGAGGTAAGTGGAATTCATGTGTAGCGGTAAAATGCGTAAAT

>p2h2

GGGCCTGCGGCGCATTAGTTAGTTGGCGGGGTGAAGGCCCAACCAAGACGATGATGCGTAGCCGGACTGAGAGGTTGAA
CGGCCACATTGGGACTGAGAGACGGCCCAAACCTCCTACGGGAGGCAGCAGTAGGGAATTTTCGGCAATGGGGCAACC
CTGACCGAGCAACGCCCGGTGAACGACGAAGTACTTCGGTATGTAAAGTTCTTTATCAGGGAAGTGACGGTACCTGA
TGAATAAGCCCCCGCTAAGTATGTGCCAGCAGCCGCGGTAATACATAGGGGGCAAGCGTTATCCGGAATTATTGGGCG
TAAAGGGTGCGTAGGTGGCGTTTAAAGTCTTGGTGTAAGTGCAAGGCTCAACCTGTGATGCTAAGGAACTGGATG
GCTTGACTGTGATAGGTAAGCGGAATTCTTGGTGTAGCGGTGAAATGCGTAGAT

>p1b8

GGGTCTGCGTCTGATTAGCTTGTGGTGAGGTAATGGCTCACCAAGGCTACGATCAGTAGCTGGTCTGAGAGGATGAT
CAGTCACATTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGTAGTGGGGAATTTTGC GCAATGGGCGAAAGC
CTGACGCAGCAACGCCCGGTGACCGATGAAGCATTTAGGTGTGTAAAGGTCTGTGTTAGGGAAGTGACGGTACCTGA
CAAGAAAGTCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGAGCAAGCGTTATCCGGAATTACTGGGCG
TAAAGGGCACGTAGGCGGGCTCTTAAGTCAGATTTTAAATCCAACAGCTTAACTGTTGATATGGATCTGAAACTGGGA
GTCTTGAATGCAAGAGAGGAAAGCGGAATTCTTGGTGTAGCGGTGAAATGCGTAGAT

>P2A10

GAGCCTGCGGCCCATTAGTTAGTTGGTGGGGTAACGGCCTACCAAGGCAGCGATGGGTAGCCGGCCTGAGAGGGTGAT
CGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAAGAATCTTCGCAATGGGCGAAAGC
CTGACGGAGCGACGCCCGGTGGGTGATGAAGCGGAAACGTTGTAAAGCCCTTTTCTGGAAGTGACATGAACCGAGGA
ATAAGCCACGGCTAATTACGTGCCAGCAGCCGCGGTAACACGTAAGTGGAAGCGTTGTTCCGGAATTATTGGGCGTAA

AGGGCGTGTAGGCGGCCCTTGTAAAGCCTGGTGTGAAAGCCCGCAGCTTAACTGCGGAGACGCGCTGGGAACACAAGGC
TGGAGTCTAAGAGGGGAGCTGGAATTCAGGTGTAGGGGTGAAATCTGTAGAT

>p2a9

GAGCCTGCGGCCCATTTAGTTAGTTGGTGGGGTAACGGCCTACCAAGGCAGCGATGGGTAGCCGGCCTGAGAGGGTGAT
CGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCTAAGAATCTTCCGCAATGGGCGAAAGC
CTGACGGAGCGACGCCGCGTGGGTGATGAAGCGGAAACGTTGTAAAGCCCTTTTCTTGGGGAAGTGACATGAACCGA
GGAATAAGCCACGGCTAATTACGTGCCAGCAGCCGCGGTAACACGTAAGTGGCAAGCGTTGTTCGGAATTATTGGGCG
TAAAGGGCGTGTAGGCGGCCCTTGTAAAGCCTGGTGTGAAAGCCCGCAGCTTAACTGCGGAGACGCGCTGGGAACACAA
GGCTGGAGTCTAAGAGGGGAGCTGGAATTCAGGTGTAGGGGTGAAATCTGTAGAT

>p2h10

GAGCCTGCGGCTCATTAGTTAGTTGGTGGGGTAAAGGCCTACCAAGGCGACGATGAGTAGCCGGCCTGAGAGGGTGAT
CGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCTAAGAATCTTCCGCAATGGGGGAAACC
CTGACGGAGCGACGCCGCGTGGACGAAGAAGGCCGGAAGGTTGTAAAGTCCCTTTTCTTGGGGAGTGACGTGAACCAG
GGAATAAGCCCCGGCTAATTACGTGCCAGCAGCCGCGGTAACACGTAAGGGGCGAGCGTTGTTTCGGATTTATTGGGCG
TAAAGGGCGTGTAGGCGGTGAGGTAAAGTCTGGTGTGAAATGTCCCGGCTCAACCGGGAAGCGCGCTGGAAACTGTCT
TACTTGAGTGATGGAGGGGAGCTGGAATTCCTGGTGTAGGGGTGAAATCTGTAGAG

>P2F04

GAGTCCGCGTCCCATTTAGCTTGTGGTGGGGTAAGAGCCTACCAAGGCTGCGATGGGTAGCCGGCCTGAGAGGGTGAA
CGGCCACATTTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTTAAGAATCTTCCGCAATGGGCGAAAGC
CTGACGCAGCGACGCCGCGTGGACGATGAAGGCCTTCGGGTTGTAAAGTCCAATAAGCAGGGAAGTGACGGTACCTGC
CTAAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTATGGGGCAAGCGTTGTTTCGGAATTATTGGGCGTA
AAGGGCGTCTAGGCGGGGACTAAGTTAGGTGTGAAATCCACGGGCTCAACCCATGAATTGCACCTAAACTGGTTCT
CTTGAGTTCGGGAGAGGAGAGCGGGATTCCCGGTGTAGCGGTGAAATGCGTAGAT

>p1a2

GGTCCCGCGTCCCATTTAGCTAGTTGGTAGGGTAACTGCCTACCAAGGCGACGATGGGTAGCCGGCCTGAGAGGGTGAC
CGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTCGAGAGGCTTCGGCAATGGGGGAAACC
CTGACCGAGCGACGCCGCGTGGAGGATGAAGTTTTTCGGAACGTAAAGTCCGTGCATACGGGATGTGACAGTACCGTA
AGAGGAAGTCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGAGCGAGCGTTATTTCGGATTCATTGGGCG
TAAAGGGTTCTAGGCGGTGAGATAAGTCAGGGGTGAAATCTAACGGCTCAACCGTTAAATTGCCTTTGAAACTATTT
GACTTGAGTACGAGAGAGGAAAGCGGAATTCCTGGTGTAGCGGTGAAATGCGTGGAT

>p1d9

GGGCTCGCGTCTATCAGCTTGTGGTGGGGTAATGGCCTACCAAGGCGTTGACGGGTAGCTGGTCTGAGAGGATGGT
CAGCCACACTGGGACTGAGACACTGCCCAGACTCCTACGGGAGGCTGCAGTCGAGAATATTTAGCAATGGGCGAAAGC
CTGACTATGCGACGCCGCGTGGAGGATGAAGTTTTTCGGATCGTAAACCTCTTTTATCAGGAAATGACGGTACCTGAG
GAATAAGCCACGGCCAACCTCCGTGCCAGCAGCCGCGGTAATACGGAGGTGGCGAGCGTTACTCGGATTCATTGGGTGT
AAAGGGCAAGTAGGCGGTCCAATGGGTGAGGGGTGAAATCTTCCGACTCAATCGGAAAACCTGCCTTTCAAACATATTGG
ACTTGAGGTTGGAAGAGGAAAATGGAATTCCTAGTGTAAGGGTGAAATCTGTGGAT

>P3E04

CGGGCTATATGGTATCAGCTAGTTGGTGAGGTAATGGCTCACCAAGGCAATGACGCCTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAAGTGAAGACAGGTCCAGACTCCTACGGGAGGCAGCAGTGGGAATATTGCACAATGGAGGAAACT
CTGATGCAGCAACGCCGCGTGGAGGATGACGCATTTCCGGTGTGTAAACTCCTTTTATATAGGAAGTGACGGTACTATA
TGAATAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTACTCGGAATCACTGGGCG
TAAAGCGCGCGCAGGCGGCCCTTTTAAGTTGGATGTGAAAGCCTATGGCTCAACCATAGAAGTGCATCCAAAACATATCA
GGCTAGAGTGTGGGAGAGGAAAGATGGATTAGTTGTGTAGGGGTCAAATCCGTAGAG

>p2h8

CGGGCTATATGGTATCAGCTAGTTGGTGAGGTAATGGCTCACCAAGGCAATGACGCCTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAAGTGAAGACAGGTCCAGACTCCTACGGGAGGCAGCAGTGGGAATATTGCACAATGGAGGAAACT
CTGATGCAGCAACGCCGCGTGGAGGATGACGCATTTCCGGTGTGTAAACTCCTTTTATATAGGAAGTGACGGTACTATA
TGAATAAGCACCGGCTAACTCCGCGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTACTCGGAATCACTGGGCG
TAAAGCGCGCGCAGGCGGCCCTTTTAAGTTGGATGTGAAAGCCTATGGCTCAACCATAGAAGTGCATCCAAAACATATCA
GGCTAGAGTGTGGGAGAGGAAAGATGGATTAGTTGTGTAGGGGTAAAATCCGTAGAG

>p1c2

CGGCCTATGTCGTATCAGCTAGTTGGTGGGGTAAGAGCTCACCAAGGCAATGACGCGTAACTGGTCTGAGAGGATGAT
CAGTCACACTGGAAGTGAAGACAGGTCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGCACAATGGAGGAAACT
CTGATGCAGCAACGCCGCGTGGAGGATGACGCATTTCCGGTGTGTAAACTCCTTTTATATAGAAAGTGACGGTACTATA
TGAATAAGCGCCGCTAATCCGTGCCAGCAGCCGCGGTAATACGGGGGCGCAAGCGTTACTCGGAATCACTGGGCG
TAAAGGGAGCGTAGGCGGCCCTAATAAGTGTCTTGTGAAAGCCTGTGGCTTAACCATAGAAGTGCAGACAAACTGTTA
GGCTAGAGTTCCGAAGAGGTAAGTGGATTGGTGGTGTAGGGGTAAAATCCGTAGAG

>P3F10

TGGCCTGTATAGTATCAGTTAGTTGGTGAGGTAATGGCTCACCAAGACAATGACGCTTAACTGGTTTGAGAGGATGAT
CAGTCACACTGGAACCTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACC
CTGATGCAGCAACGCCGCGTGGAGGATGACACATTTCCGTTGCGTAAACTCCTTTTATAATAGAAGTGACGGTATATTA
TGAATAAGCTCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGGCAAGCGTTACTCGGAATCACTGGGCG
TAAAGAGCATGTAGGCCGATAGATAAGTTGGAAGTGAAATCCTATGGCTTAACCATAGAAGTCTTCCAAAACGTCT
ATCTAGAATATGGGAGAGGTAGATTTTCTGGTGTAGGGGTAAAATCCGTAGAG

>p3g11

CGGGCTATATCGTATCAGCTTGTGGTGAGGTAACGGCTCACCAGGGCTATGACGCGTATCTGGTTTGAGAGGATGAT
CAGACACACTGGAACCTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGCACAATGGGGGAAACC
CTGATGCAGCAACGCCGCGTGGAGGATGACACATTTCCGTTGCGTAAACTCCTTTTAAACAGAGAAGTGACGGTATCTGT
TGAATAAGCAGCCGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGGGGGCGCAAGCGTTATTCCGAATCACTGGGCG
TAAAGGGTGCGTAGGCGTCTAATAAGTCTCCTGTGAAAGCCACGGCTCAACCGTGGAACCTGCAGGGGAAACTGTTA
GGATAGAGTGTGGAAGAGGTAGATGGAATTTCCGGTGTAGGGGTAAAATCCGTAGAG

>p2a11

CGGCCTATATTGTATCAGTTAGTTGGTGGGGTAAGAGCCTACCAAGGCGATGACACATAGCGGGTCTGAGAGGATGAT
CCGCCACACTGGAACCTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGCACAATGGGGGAAACC
CTGATGCAGCAACGCCGCGTGGAGGATGACGCATTTCCGTTGTGTAATAACTCCTTTTATCAGAGAAGTGACGGTATCTGA
TGAATAAGTCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGAGCAAGCGTTACTCGGAATCACTGGGCG
TAAAGGGCGCGTAGGCGGGTGTCAAGTCAGGTGTGAAATCCTATGGCTCAACCATAGAAGTGCACCTTGAAACTGTCA
ACCTAGAGTATGGAAGGGGCAGATGGAATTAGTGGTGTAGGGGTAAAATCCGTAGAT

>p3d4

GAGGCTATGTCTATCAGCTAGTTGGTGAGGTAATGGCTCACCAAGGCTATGACGGGTAGCTGGTTTGAGAGGATGAT
CAGCCACACTGGTACTGAGACACGGTCCAGAATCCTACGGGAGGCAGCAGTGGGGAATATTGGGCAATGGAGGCAACT
CTGACCCAGCAACGCCGCGTGGAGGATGAAGCATTTCCGTTGTGTAATAACTCCTTTTATACAAGAAGTGACGGTATTGTA
TGAATAAGCAGCCGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTACTCGGAATCACTGGGCG
TAAAGGGTGTGTAGGCGGATTAGTAAGTCTTGTGTGAAATCCTACGGCTTAACCGTAGAAGTGCATGAGAACTGTGA
ATCTAGAGTGTGAAAGAGGCAGGTGGAATTAGTGGTGTAGGGGTAAAATCCGTAGAT

>P2B05

GAGCCTATATTCCATCAGGTAGTTGGTAGGGTAACGGCCTACCAAGCCTACGACGGATAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAACCTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGATGATGAAGTCTTCCGGAATGTAAAATCCTTTTGCAGAGGAAGTGACGGTACTCTG
CGAATAAGCCACGGCTAACTCTGTGCCAGCAGCCGCGGTGATACAGGGGTGGCAAGCGTTGTCCGGATTTACTGGGTG
TAAAGGGTGCGCAGGCGGGCTCATAAGTCGGGGGTTAAATCCATGTGCTTAACACATGCATGGCTTCCGATACTGTGA
GTCTAGAGTCTCGAAGAGGAAGATGGAATTTCCGGTGTACGGTGAATGTGTAGAT

>P2E07

GAGCTCGCGGCCTATCAGGTAGTTGGTGAGGTAATGGCTCACCAAGCCAATGACGGGTAGCTGGTCTGAGAGGATGGT
CAGCCATACTGGAACCTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGCCGAGAATCTTCCGCAATGGGCGAAAGC
CTGACGGAGCGACACCGCGTGTAGGACGAAAGCCTTCGGGTTGTAAACTACTGTCACTTGGGACGGGACTGTACCAAG
AGAGGAAGCCACGGCTAACTCTGTGCCAGCAGCCGCGGTAATACAGAGGTGGCGAGCGTTGTTCGGATTGACTGGGTG
TAAAGGGTGCGTAGGCGGCTTGGAAGTCGGATGTGAAATCCTGCAGCTCAACTGCAGAACGGCATTTCGAAACTGCCA
GGCTAGAGTGCGAGAGAGGTGAAGGGAATTCCCGGTGTAAAGGTGAAATCTGTAGAG

>p2e8

GAGCTCGCGGCCTATCAGGTAGTTGGTGAGGTAATGGCTCACCAAGCCAATGACGGGTAGCTGGTCTGAGAGGATGGT
CAGCCATACTGGAACCTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGCCGAGAATCTTCCGCAATGGGCGAAAGC
CTGACGGAGCGACACCGCGTGTAGGACGAAAGCCTTCGGGTTGTAAACTACTGTCACTTGGGACGGGACTGTACCAAG
AGAGGAAGCCACGGCTAACTCTGTGCCAGCAGCCGCGGTAATACAGAGGTGGCGAGCGTTGTTCGGATTGACTGGGTG
TAAAGGGTGCGTAGGCGGCTTGGAAGTCGGATGTGAAATCCTGCAGCTCAACTGCAGAACGGCATTTCGAAACTGCCA
GGCTAGAGTGCGAGAGAGGTGAAGGGAATTCCCGGTGTAAAGGTGAAATCTGTAGAG

>p1g2

GGGCTCATGTCTATCAGCTAGTTGGTGAGGTAACGGCTCACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTGAA
CGGCCACACTGGAACCTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGGCAATGGGCGAGAGC
CTGACCCAGCGACGCCGCGTGGAGGATGACAGCCTTCGGGTTGTAAACCTCTGTTGTATGGGAGATGACGGTACCATA
CGAGGAAGTCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGAGCGAGTGTGTCCGGAATTACTGGGCG
TAAAGGGCAGCGAGGCGGTCATATAAGTCAGCTGTTAAATGTATGGGCTTAACCGATGCATGCGGTTAACTGTATGAC
TGGAGTACTGGAGAGGTAAGCGGAATTCCCGGTGTAGCGGTGAAATGCGTAGAT

>p3b8

GTGCTCATGTCTATCAGCTGGTTGGTGGGGTAACGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTGAA
CGGCCACACTGGAACCTGAGATACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGGCAATGGGGGCAACC
CTGACCCAGCGACGCCGCGTGGGGGAAGACGGCCTTCGGGTTGTAAACCTCTTTTGTACGGGAAGTGACGGTACTGTA
CGAATAAGTCCCGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGACGAGCGTTGTCCGGAATTACTGGGCG

TAAAGGGCGCGCAGGCTGTGCTGCAAGTCAGCTGTTAAATGTCTGGGCTTAACCCGGGCATGCGGTTGACTGCGGTGC
TGGAGTGCTGTAGAGGCAAGTGGAATTCCCAGTGTAGCGGTGAAATGCGTAGAT

>p1c11

GAGCTCGCGTCCGATTAGCCAGTTGGCGGGGTAACGGCCACCAAGCAACGATCGGTAGCCGGCCTGAGAGGGCGAT
CGGCCACACTGGAAGTGAAGACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGAATATTGGGCAATGGGCGAAAGC
CTGACCCAGCAACGCCGCGTGAGGGAAGAAGGTTTTTCGGATTGTAAACCTCTGTCTTGGTGACGTGACGGTAGCCAA
GGAGGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGCAAGCGTTGTCCGGATTTACTGGGCG
TAAAGAGCGCGTAGGCGGCTTGGTAAGTCGGCTGTGAAAGCCAGGGCTCAACCCTGGAAGTGCATTTGAAACTGCCT
AGCTTGAGTATGGGAGAGGAGAGTGGAATTCCCAGTGTAGAGGTGAAATTCGTAGAT

>p2b1

GGTGACTTGTTCTTATAACCATTGCGGGTCGGCCTACCAAGCACGATCGTTACGGCCTGAGAGGGCTATCGCCACAC
TGGAAGTGAAGACGGTCCAGATTTCGTACGGGAGGCAGCAGTGGGAATATTGGGCAATGGGCGAAAGCCTGACCCAT
CAACGCCGCGTGAGGAAGAAGGTCTTCGGATTGTAAACCTCTGTCTTGGTGACGTGACGGTAGCCAAGGAGGAAGCC
CCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGGAGGGGGCAAGCGTTGTCCGGATTTACTGGGCGTAAAGAGCGC
GTAGGTGGCCTGTAAAGTCGGCTGTGAAATCGCTGGGCTCAACCCAGCTGGGTGAGTCGATACTGGTAGGCTCGAGTC
CGGGAGAGGAAGCGGAATTGCTGGTGTAGCGGTGGAATGCGTAGAT

>P1D04

GGACTCGCATCCCATTAGCTAGTTGGTGAGGTAACGGCCACCAAGGCAACGATGGGTAGCCGACCTGAGAGGGTGAT
CGGCCACACTGGAAGTGAAGACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGAATATTGGGCAATGGGCGCAAGC
CTGACCCAGCAACGCCGCGTGAGGGAAGAATGTCTTCGGATTGTAAACCTCTGTCTTGGGGACGTGACGGTACCCAA
GGGAGGAAGCTCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGGGTGCGTAGGCGGCGCTGCAAGTCAGATGTGAAATCCCCGGGCTTAACTCGGGAGTTGCATTTGAAACTGCAG
TGCTTGAGTGAAGAAGAGGCAAGTGGAATTCCTAGTGTAGCGGTGGAATGCGTAGAT

>p1h11

GAGCTCGCGTCTGATTAGCTAGTTGGCGGGGTAAGGCCACCAAGGCGACGATCAGTAGCCGGACTGAGAGGTTGAA
CGGCCACATTGGGACTGAGATACGGCCAGACTCCTACGGGAGGCAGCAGTGGGAATATTGGGCAATGGGCGAAAGC
CTGACCCAGCAACGCCGCGTGAAAGGAAGAAGGCTCTCGGGTTGTAAACTTCTTTGACCAGGGAAGTGACGGTACCTGG
AAAACAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCGAGCGTTGTCCGGATTTACTGGGTG
TAAAGGGCGTGAGGCGGGATTGCAAGTCAGATGTGAAATGCCGGGCTTAAACCCGGAGCTGCATTTGAAACTGCAG
TTCTTGAGTGATGAGAGGCAGGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGAT

>p3e5

GAGTCCGCCCCCATTAGCTAGTTGGCGGGATAATAGCCCACCAAGGCAACGATGGGTAGCCGAGCTGAGAGGTTGAT
CGGCCACACTGGGACTGAGATACGGCCCAGACACCTACGGGTGGCAGCAGTAGGGAATATTGGACAATGGGCGAAAGC
CTGATCCAGCCATGCCGCGTGAGTGATGAAGGCCTTCGGGTGCTAAAACCTTTTCTGAGGGAAGTGACGGTACCTCA
GGAATAAGCAACGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGTTGCGAGCGTTGTTCGGATTAACTGGGCG
TAAAGGGAGCGCAGGTGGTTTGGTTAGTCTAAAGTGAAAGCCCCGGGCTTAACTCGGAAGTGCTTTGTATACTGCCA
AAGTTGAGTTTCGTGATAGGAAAGCGGAATTCAGGTGTAGCGGTGAAATGCGTAGAT

>p3e2

GGGCTTGCGTCCGATTAGCTGGTTGGTAGGGTAACGGCTTACCAAGGCGAGGATCGGTAGCTGGTTTGAGAGGACGGT
CAGCCACACTGGGACTGAGAAACGGCCCAGACTCCTACTGGAGGCAGCAGCAAGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCGACGCCGCGTGGGGGATGAAGGCTTCGGGTTGTAAACCCCTTTTCTCAGGGAAGTGACGGTACCTGA
GGAATAAGTCTCGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAGCGTTATCCGGATTTACTGGGCGT
AAAGCGGGCGTAGGCGGTCTTCAAGTCGGATGTAAAATCTCCCGGCTCAACTGGGAGAGACCATTGAAACTGCTGG
GCTTGAGGATAGCAGAGGGAGGTGGAATTCGGGTGTAGTGGTGATATGCGTAGAT

>p3d11

GGGCTTGCGTCCGATTAGCTGGTTGGTAGGGTAACGGCTTACCAAGGCGAGGATCGGTAGCTGGTTTGAGAGGACGGT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCGACGCCGCGTGGGGGATGAAGGCTTCGGGTTGTAAACCCCTTTTCTCAGGGAAGTGACGGTACCTGA
GGAATAAGTCTCGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAGCGTTATCCGGATTTACTGGGCG
TAAAGCGGGCGTAGGCGGTCTTCAAGTCGGATGTAAAATCTCCCGGCTCAACTGGGAGAGACCATTGAAACTGCTG
GGCTTGAGGATAGCAGAGGGAGGTGGAATTCGGGTGTAGTGGTGATATGCGTAGAT

>P3D06

GGGCCTGCGTCCGATTAGCTGGTTGGTAGGGTAACGGCTTACCAAGGCGAAGATCGGTAGCTGGTTTGAGAGGACGGT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCGACGCCGCGTGGGGGATGAAGGCTTTCGGGTTGTAAACCCCTTTTCTCAGGGAAGTGACGGTACCTGA
GGAATAAGTCTCGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAGCGTTATCCGGATTTACTGGGCG
TAAAGCGGGCGTAGGCGGTCTTCAAGTCGGATGTAAAATCTCCCGGACCAATGGGAGAGAACATTGAAAATGCTG
GGCTTGAGGATAGCAGAAGGAGGTGGAATTCGGGTGTAGTGGTGATATGCGTAGAT

>P1A06
GGGCTTGTGTCCGATTAGCTGGTTGGTAGGGTAAAGGCTTACCAAGGCGATGATCGGTAGCTGGTTTGAGAGGACGAT
CAGCCACACTGGGATTGAGGTACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGAATATTGTGCAATGGGCGAAAGC
CTGACACAGCGACGCCGCGTGGGGGATGAAGGCTTTCGGGTGTGTAACCCCTTTTCTCAGGGAAGTGACGGTACCTGA
GGAATAAGTCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCGAGCGTTATCCGGATTTACTGGGTG
TAAAGAGGGCGCAGGCGGCTTTTCAAGTCGAATATAAAATCTCCCGGCTTAAGAGGGAGGGACTATTGATACTGTTG
GGCTAGAGGGCAGCAGGGGGAGGTGGAATTCCCGGTGTAGTGGTGAAATGCGTAGAT

>p1g8
GGGCTTGTGTCCGATTAGCTGGTTGGTAGGGTAAAGGCTTACCAAGGCGATGATCGGTAGCTGGTTTGAGAGGACGAT
CAGCCACACTGGGATTGAGGTACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGAATATTGTGCAATGGGCGAAAGC
CTGACACAGCGACGCCGCGTGGGGGATGAAGGCTTTCGGGTGTGTAACCCCTTTTCTCAGGGAAGTGACGGTACCTGA
GGAATAAGTCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCGAGCGTTATCCGGATTTACTGGGTG
TAAAGAGGGCGCAGGCGGCTTTTCAAGTCGAATATAAAATCTCCCGGCTTAAGAGGGAGGGACTATTGATACTGTTG
GGCTAGAGGGCAGCAGGGGGAGGTGGAATTCCCGGTGTAGTGGTGAAATGCGTAGAT

>P1A05
GGGCTTGCGTCCGATTAGCTAGTTGGTAGGGTAATGGTCTACCAAGGCGAAGATCGGTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGAATATTGTGCAATGGGCGAAAGC
CTGACACAGTGACGCCGCGTGGGGGATGAAGGCTTTCGGGTGTGTAACCCCTTTTCTCAGGGAAGTGACGGTACCTGA
GGAATAAGTCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGATTTATTGGGCG
TAAAGAGGGCGTAGGCGGCTTTTCAAGTCGATGTAAAATCTCCCGGCTCACTGGGAGCGATCATTGATACTGCTA
GGCTAGAGGACGGCAGAAGGAGGCGGAATTCCCGGTGGAGTGGTGAAATGCGTAGAT

>p1d8
GGGCTTGCGTCCGATTAGCTAGTTGGTAGGGTAATGGTCTACCAAGGCGAAGATCGGTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGAATATTGTGCAATGGGCGAAAGC
CTGACACAGTGACGCCGCGTGGGGGATGAAGGCTTTCGGGTGTGTAACCCCTTTTCTCAGGGAAGTGACGGTACCTGA
GGAATAAGTCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGATTTATTGGGCG
TAAAGAGGGCGTAGGCGGCTTTTCAAGTCGATGTAAAATCTCCCGGCTCACTGGGAGCGATCATTGATACTGCTA
GGCTAGAGGACGGCAGAAGGAGGCGGAATTCCCGGTGGAGTGGTGAAATGCGTAGAT

>p1e1
GGGCTTGCGTCCGATTAGCTAGTTGGTAGGGTAAGGGCTTACCAAGGCGATTATCGGTAGCTGGTCTGAGAGGGCGAC
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGAATTTTGCACAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGTCTTCGGGTGTGTAACCCCTTTTCTCAGGGAAGTGACGGTATCTGA
GGAATAAGTCTCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCAAGCGTTATCCGGATTTACTGGGCG
TAAAGGGGGCGCAGGCGGCTTCTCAAGTCGGGTGTAAAATCTCCCTGCTCAACAGGGAGGGATCACCAATACTGTCA
AGCTAGAGAGCAGCAGAGGGAAGCGGAATTCCCGGTGGAGTGGTGATATGCGTAGAT

>P3A08
GGGATTGCGTCCGATTAGCTAGTTGGTAGGGTAAGGGCTTACCAAGGCGATGATCGGTAGCTGGTCTGAGAGGACGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGAATATTGGGCAATGGGCGAAAGC
CTGACCCAGCGACGCCGCGTGAGGGAATAAGGCCTTCGGGTGTGTAACCTCTTTTCTCAGGGAATGACGGTACCTGA
GGAATAAGTCTCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCAAGCGTTATCCGGATTTACTGGGCG
TAAAGCGGGCGCAGGCGGCTTTTCAAGTCGATTGTAAAATCTCCCGCTCAACGGGGAGAGGTCAATTCGATACTGTTG
AGCTAGAGAGCAGCAGAGGGAAGCGGAATTCCCGGTGTAGTGGTGAAATGCGTAGAT

>p2b12
GGGCTTGCGTCCGATTAGCTAGTTGGTGGGGTAGTGGCTTACCAAGGCAGTGATCGGTAGCTGGCTTGAGAGGGTGGC
CAGCCACACTGGGACTGAGACACGGNCCAGACTCCTACGGGAGGCAGCAGCAGGAATTTTGGGCCATGGGCGAAAGC
CTGACCCAGCGACGCCACGTGAGGAAGAAGGCTCTCGGTTGTGTAACCCCTTTTCTCAGGGAAGTGACGGTACCTGAG
GAATAAGTCTCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTGGGTGGCAAGCGTTATCCGGATTTACTGGGCGT
AAAGAGGGCGCAGGCGGCTTTTCAAGTCGGATGTAAAATCTCCCGCTTAACTGGGAGGGATCATTGATACTGTAA
GCTAGAGGTCAACAGGGGGAGGTGGAATTCCCGGTGTAGTGGTGATATGCGTAGAT

>P2E10
GGGCTTGCGTCCGATTAGCTAGTTGGTGGGGTAATGGCTTACCAAGGTTATGATCGGTAGCTGGTCTGAGAGGACGGT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAGGGAATATTGGGCAATGGGCGAAAGC
CTGACCCAGCGACGCCGCGTGAGGGATGAAGGGCTTCGGTTCGTAAACCTCTTTTCTCAGGGATGTGACAGTACCTGG
GGAATAAGTCACGGCTAACTACGTGCCAGCAGCCGCGGTAACGTAGGTGGCGAGCGTTATCCGGATTTACTGGGCG
TAAAGCGGGCGCAGGCGGCTTCTCAAGTCGGATGTAAAATCTCCCGCTTAACTGGGAGGGACCATCCGATACTGTTG
AGCTAGAGGACAGCAGAGGAAGGTAGAATTCCCGGTGTAGTGGTGGAATGCGTAGAT

>p2b11
GGGCTTGCGTCCGATTAGCTAGTTGGTGGGGTAATGGCTTACCAAGGCAAGGATCGGTAGCTGGTCTGAGAGGACGTC
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGAATAGTAGGCAATGGGCGAAAGC
CTGACCTCGCGACGCCGCGTGGAGGATGAAGGTCTTCGGGTGTGTAACCTCTTTTCTCAGGGATGTGACAGTACCTGA
GGAATAAGTCACGGCTAACTACGTGCCAGCAGCCGCGGTAACGTAGGTGGCGAACGTTATCCGGATTTACTGGGCG

TAAAGCGTGCAGGCGGCTTTTCAAGTCGGATGTAAAATCTCTTGGCTCAACTGGGAGCGACCGTCCGATACTGCTA
AGCTGGAGTGCAACAGAGGAGAGTGGAATTCCCAGTGTAATGGTGGAATGTGTAGAT

>p2d2

GGGCTTGCGTCCGATTACTTGTGGTGGGGTAACGGCTACCAAGGCGATGATCGGTAGCTGGTCTGAGAGGATGGTCA
GCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGAATTTTGCACAATGGGCGAAAGCCT
GACGCAGCAACGCCGCGTGGGGGATGAAGGCCTTCGGGTTGTAAACCCCTTTTCTCAGGGAAGTGACGGTACCTGAGG
AATAAGCCTCGGCTAAATACGTGCCAGCAGCCGCGGTAATACGTATGAGGCGAGCGTTATTCGGATTTACTGGGCGTA
AAGAGGGCGTAGGCGGCCCTTCAAGTCGGATGTGAAATCTCCCGGCTCAACCGGGAGGGGCCATCCGATACTGCTGGG
CTAGAGGGCAGCAAGGGAAGGTAGAATTCCCAGTGTAATGGTGGAATGCGTAGAT

>P1C04

GGGCTTGTGGCCGATTAGCTAGTTGGTAGGGTAATGGCTTACCAAGGTTACGATCGGTAGCTGGTCTGAGAGGATGGT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGAATTTTGCACAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGATGACGGCCTTCGGGTTGTAAACCCCTTTTCTCAGGGAAGTGACGGTACCTGA
GGAATAAGCCTTGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGAAGGCGAGCGTTATTCGGATTTACTGGGCG
TAAAGAGGCGTAGGCGGCCCTTCAAGTCGGATGTGAAATCTCTCGGCTTAACCGGGAGAGGTCATCCGATACTGATG
GGCTGGAGGGCAGCAGGAGAAAGCGGAATTCCCAGTGTAATGGTGATATGCGTAGAT

>p1f1

GGGCTTGCGGCCGATTAGCTAGTTGGTAGGGTAACGGTTTACCAAGGCGATGACCGGTAGCTGGTTTGAAGAGGATGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGAATTTTGCACAATGGGCGAAAGC
CTGACGCAGCAGCAGCCGCGTGGGGGATGAAGGTCTTCGGATTGTAAACCCCTTTTCTCAGGGAAGTGACGGTACCTGA
GGAATAAGCCTCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAGCGTTATCCGGAATTATTGGGCG
TAAAGAGCGGTGTAGGCGGTCTGCAAGTTGGATGTGAAATCTCCCGGCTTAACCGGGAGAGGTCATCCAATACTGCTA
GGCTAGAGTATAACAGAGGGAGATGGAATTCCCAGTGAGTGGCGAAATGCTTAGAT

>P2E12

GGGCTTGCGTTCAATTAGCTAGTTGGTGGGGTAATGGCCTACCAAGGCGATGATTGGTAGCTGGTCTGAGAGGATGTC
CAGCCAGATGGGACTGAGACACGGCCCCTAACTCCTACGGGAGGCAGCAGCAAGGAATATTGGGCAATGGGCGAGAGC
CTGACCCAGCGACACCGCGTGAGGGAAGAAGGCCTTCGGGTTGTAAACCTCTTTTCTCGGGGAAGTGACAGTACCCGA
GGAATAAGTCTCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAACGTTATCCGGATTTACTGGGCG
TAAAGCGGGTGTAGGCGGTTTCTAAGTCAGGTGTGAAATCTTCTGACTCAATCAGAAGGGGTCATTTGATACTGGAG
AACTTGAGGGCGGCAGAAGGAAATGGAATTCCCAGTGATAGCGGTGAAATGCGTAGAT

>p1c8

GGGCTTGCGTCTGATTAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGATGATCAGTAGCTGGTCTGAGAGGATGTC
CAGCCAGATGGGACTGAGATACGGCCCCAACTCCTACGGGAGGCAGCAGCAAGGAATATTGGGCAATGGGCGGAAGC
CTGACCCAGCGACACCGCGTGAGGGAAGAAGGCCTTCGGGTTGTAAACCTCTTTTATTGGGGAAGTGACGGTACCCGA
TGAATAAGTCTCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAACGTTATCCGGATTTACTGGGCG
TAAAGGGGGTGTAGGCGGCTTCTTAAGTCAGGTGTGAAATCTTCTGACTCAATCAGAAGGGGTCATCTGATACTGGGA
AACTTGAGGGCAGCAGAAGGAAATGGAATCCCAGTGATAGCGGTGAAATGCGTAGAT

>p1b9

GGGCTTGCGTCTGATTAGTTAGTTGGTGGGGTAATAGCCTACCAAGGCGATGATCGGTAGCTGGTCTGAGAGGATGGT
CAGCCAGATGGGACTGAGATACGGCCCCAACTCCTACGGGAGGCAGCAGCAAGGAATTTTGGGCAATGGGCGAAAGC
CTGACCCAGCGACACCGCGTGAGGGAAGAAGGCCTTCGGGTTGTAAACCTCTTTTATTGGGGAAGTGACGGTACCCAA
GGAATAAGTCTCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTGGGAGGCGAACGTTATCCGGATTTACTGGGCG
TAAAGGGGGTGTAGGCGGCCCTTTAAGTCAGGTGTGAAATTTTTCGACTTAATCGGAAAGGGTCATCTGATACTGGAA
GGCTTGAGGGCGGCAGAAGGAAATGGAATTCCCAGTGATAGCGGTGAAATGCGTAGAT

>P1D06

TGGCTTGCGTCTGATTAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGATGATCAGTAGCTGGTCTGAGAGGATGGT
CAGCCAGATGGGACTGGGACACGGCCCCAACTCCTACGGGAGGCAGCAGCAAGGAATTTTGGGCAATGAGCGAAAGC
TTGACCCAGCGACACCGCGTGAGGGAATAAGGCCTTCGGGTTGTAAACCTCTTTTCTTGGGGAAGTGACGGTACCCAA
GGAATAAGCCTCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCAAACGTTATTCGGATTTACTGGGCG
TAAAGGGAGCGTAGGCGGTTCCCTAAGTCGGATGTGAAATCTTCTGGCTCAACCAGAAGGCGCCATTTGATACTGAGG
GACTTGAGAGCAGCAGAGGGAAATGGAATCTCCAGTGATAGCGGTGAAATGCGTAGAT

>P2G04

TGGCTTGCGTCTGATTAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGATGATCAGTAGCTGGTCTGAGAGGATGGT
CAGCCAGATGGGACTGAGACACGGCCCCAACTCCTACGGGAGGCAGCAGCAAGGAATTTTGGGCAATGAGCGAAAGC
TTGACCCAGCGACACCGCGTGAGGGAATAAGGCCTTCGGGTTGTAAACCTCTTTTCTTGGGGAAGTGACGGTACCCAA
GGAATAAGCCTCGGACAACACGCGCCAGCAGCCGCGGTAATACGTAGGAGGCAAACGTTATTCGGATTTACTGGGCG
TAAAGGGAGCGTACGCGGTTCCCTAAGTCGGATGTGAAATCTTCTGGCTCAACCAGAAGGCGCCATTTGATACTGAGG
GACTTGAGAGCAGCAGAGGGAAATGGAATCTCCAGTGATAGCGGTGAAATGCGTAGAT

>P1G05
TGGCTTGCGTCTGATTAGCTAGTTGGTAGGGTAACGGCCTACCAAGGCGATGATCAGTAGCTGGTCTGAGAGGATGGC
CAGCCAGATGGGGACTGAGACACGGCCCCAACTCCTACGGGAGGCAGCAGCAAGGAATTTTGGGCAATGAGCGAAAGC
TTGACCCAGCGACACCGCGTGAGGGAATAAGGCCTTCGGGTTGTAAACCTCTTTTCTTGGGGAAGTGACGGTACCCAA
GGAATAAGCCTCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAACGTTATTTCGGATTTACTGGGCG
TAAAGGGAGCGTAGGCGGTTCTTTAAGTCAGATGTGAAATCTTTCGGCTCAATCGGAAGGCGCCATTTGATACTGGGG
GACTTGAGAGCAGCAGAGGGAAATGGAATCTCCAGTGTAGCGGTGAAATGCGTAGAT

>P1G06
GGACTTGCGTCTGATTAGCTAGTTGGCAGGGTAATGGCCTACCAAGGCGATGATCAGTAGCTGGTCTGAGAGGATGGC
CAGCCAGATGGGGACTGGGACACGGCCCCAACTCCTACGGGAGGCAGCAGCAAGGAATTTTGGGCAATGAGCGAAAGC
TTGACCCAGCGACACCGCGTGAGGGAAGAAGGCCTTCGGGTTGTAAACCTCTTTTCTTGGGGAAGTGACGGTACCCAA
GGAATAAGCCTCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAACGTTATTTCGGATTTACTGGGCG
TAAAGGGAGCGTAGGCGGTTCTTTAAGTCAGATGTGAAATCTTCTGGCTCAATCGGAAGGTGTCATTTGATACTAAGG
GACTTGAGAGCAGCAGAGGGAAATGGAATCCCCGGTGTAGCGGTGAAATGCGTAGAT

>p1c5
GGACTTGCGTCTGATTAGCTAGTTGGCAGGGTAATGGCCTACCAAGGCGATGATCAGTAGCTGGTCTGAGAGGATGGC
CAGCCAGATGGGGACTGGGACACGGCCCCAACTCCTACGGGAGGCAGCAGCAAGGAATTTTGGGCAATGAGCGAAAGC
TTGACCCAGCGACACCGCGTGAGGGAAGAAGGCCTTCGGGTTGTAAACCTCTTTTCTTGGGGAAGTGACGGTACCCAA
GGAATAAGCCTCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAACGTTTTTCGGATTTACTGGGCGT
AAAGGGAGCGTAGGCGGTTCTTTAAGTCAGATGTGAAATCTTCTGGCTCAATCGGAAGGTGTCATTTGATACTAAGG
ACTTGAGAGCAGCAGAGGGAAATGGAATCCCCGGTGTAGCGGTGAAATGCGTAGAT

>p1f9
GGACTTGCGTCTGATTAGCTAGTTGGCAGGGTAATGGCCTACCAAGGCGATGATCAGTAGCTGGTCTGAGAGGATGGC
CAGCCAGATGGGGACTGGGACACGGCCCCAACTCCTACGGGAGGCAGCAGCAAGGAATTTTGGGCAATGAGCGAAAGC
TTGACCCAGCGACACCGCGTGAGGGAAGAAGGCCTTCGGGTTGTAAACCTCTTTTCTTGGGGAAGTGACGGTACCCAA
GGAATAAGCCTCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAACGTTATTTCGGATTTACTGGGCG
TAAAGGGAGCGTAGGCGGTTCTTTAAGTCAGATGTGAAATCTTCTGGCTCAATCGGAAGGTGTCATTTGATACTAAGG
GACTTGAGAGCAGCAGAGGGAAATGGAATCCCCGGTGTAGCGGTGAAATGCGTAGAT

>p1e8
GGACTTGCGTCTGATTAGCTAGTTGGCAGGGTAATGGCCTACCAAGGCGATGATCAGTAGCTGGTCTGAGAGGATGGC
CAGCCAGATGGGGACTGGGACACGGCCCCAACTCCTACGGGAGGCAGCAGCAAGGAATTTTGGGCAATGAGCGAAAGC
TTGACCCAGCGACACCGCGTGAGGGAAGAAGGCCTTCGGGTTGTAAACCTCTTTTCTTGGGGAAGTGACGGTACCCAA
GGAATAAGCCTCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAACGTTATTTCGGATTTACTGGGCG
TAAAGGGAGCGTAGGCGGTTCTTTAAGTCAGATGTGAAATCTTCTGGCTCAATCGGAAGGTGTCATTTGATACTAAGG
GACTTGAGAGCAGCAGAGGGAAATGGAATCCCCGGTGTAGCGGTGAAATGCGTAGAT

>P3B04GGGCTCGCGTCTGATTAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCAATGATCGGTAGCTGGTCTGAGAG
GATGGTCAGCCAGATGGGGACTGGGATACGGACCCAACTCCTACGGGAGGCAGCAGCAAGGAATCTTGGGCAATGAGC
GAAAGCTTGACCCAGCGACACCGCGTGAGGGAAGAAGGCCTTCGGGTTGTAAACCTCTTTTCTTGGGGAAGTGACGGT
ACCCAAAGCCTCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAACGTTATTTCGGATTTACTGGGCG
TGGGCGTAAAGTGAGCGTAGGCGGTTCTTTAAGTCAGATGTAAAATCTTCTGACTTAATCAGAAGGCGTCATTTGATA
CTGAGGGACTTGAGAACAGCAGAGGGAAATGGAATCCCCAGTGTAGCGGTGAAATGCGTAGAT

>P3B10
GGGCTTGCGTCTGATTAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCAATGATCGGTAGCTGGTCTGAGAGGATGGT
CAGCCAGATGGGGACTGGGATACGGCCCCAACTCCTACGGGAGGCAGCAGCAAGGAATCTTGGGCAATGAGCGAAAGC
TTGACCCAGCGACACCGCGTGAGGGAAGAAGGCCTTCGGGTTGTAAACCTCTTTTCTTGGGGAAGTGACGGTACCCAA
GGAATAAGCCTCGGCCAACCACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAACGTTATTTCGGATTTACTGGGCG
TAAAGTGAGCGTAGGCGGTTCTTTAAGTCAGATGTAAAATCTTCTGACTTAATCAGAAGGCGTCATTTGATACTGAGG
GACTTGAGAACAGGAGGGGAAAATGGAATCCCCAGTGTAGCGGTGAAATGCGTAGAT

>p1g4
GGGCTTGCACTCTGATTAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCTATGATCGGTAGCTGGTCTGAGAGGATGGT
CAGCCAGATGGGGACTGGGACACGGCCCCAACTCCTACGGGAGGCAGCAGCAAGGAATCTTGGGCAATGGGCGAAAGC
CTGACCCAGCGACACCGCGTGAGGGATGAAGGCCTTCGGGTTGTAAACCTCTTTTCTTGGGGAAGTGACGGTACCCAA
GGAATAAGCCTCGGCCAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAACGTTATTTCGGATTTACTGGGCG
TAAAGGGAGCGTAGGCGGTTCTTTAAGTCAGATGTGAAATCTTCTGATTCAATCAGAAGGGGTCATTTGATACTGAGG
GACTTGAGAACAGCAGAGGGGAAAATGGAATCCCCAGTGTAGCGGTGAAATGCGTAGAT

>P3C11
GGGCTTGCACTCTGATTAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCTATGATCGGTAGCTGGTCTGAGAGGATGGT
CAGCCAGATGGGGACTGGGACACGGGCCCCAACTCCTACGGGAGGCAGCAGCAAGGAATCTTGGGCAATGGGCGAAAGC
CTGACCCAGCGACACCGCGTGAGGGATGAAGGCCTTCGGGTTGTAAACCTCTTTTCTTGGGGAAGTGACGGTACCCAA
GGAATAAGCCTCGGACAAACACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAACGTTATTTCGGATTTACTGGGCG
GGAATAAGCCTCGGACAAACACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAACGTTATTTCGGATTTACTGGGCG

TAAAGGGAGCGTAAGCGGTTCCCTTAAGTCAGATGTGAAATCTTCTGATTCAATCAGAAGGGGTCATTTGATACTGAGG
GACTTGAGAACAGCAGAGGGGAAATGGAATCCCCAGTGTAGCGGTGAAATGCGTAGAT

>P3G03

GGGCTTGCGTCTGATTAGCTAGTTGGTAGGGTAATGGCCTACCAAGGCGATGATCAGTAGCTGGTCTGAGAGGACAAC
CAGCCAGACGGGGACTGAGACACGGCCCTGACTCCTACGGGAGGCAGCAGCAAGGAATTTTGC GCAATGGGCGAAAGC
CTGACGCAGCGACGCCGCGTGGGGGATGAAGGCCTTCGGGTGTAAACCCCTTTTCTCGAGGAAGTGACGGTACTCGA
GGAATAAGTCTTGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGAAGGCAAACGTTATTCGAATTTACTGGGCG
TAAAGCGGGTGTAGACGGCCCTTAAGTCGGGTGTGAAATCTCCTGACTCAATCAGGAGAGGTCGTTTCGATACTGAGG
AGCTTGAGGGCAGCAGAGGAAGGTGGAATTCCCGGTGTAGCGGTGATATGCGTAGAT

>p2a4

GGGCTTGCGTCTGATTAGCTAGTTGATAGGGTAATGGCCTACCAAGGCGATGATCAGTAGCTGGTCTGAGAGGACAAG
CAGCCAGACGGGGACTGAGACACGGCCCTGACTCCTACGGGAGGCAGCAGCAAGGAATTTTGC GCAATGGGCGAAAGC
CTGACGCAGCGACGCCGCGTGGGGGATGAAGGCCTTCGGGTGTAAACCCCTTTTCTCGGGGAAGTGACGGTACTCGA
GGAATAAGTCTTGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGAAGGCAAACGTTATTCGGATTTACTGGGCG
TAAAGCGGGGTAGACGGTCCCTTAAGTCGGGTGTGAAATCTCCTGACTCAATCAGGAGAGGTTATTCGATACTGAGG
AGCTTGAGGGCAGCAGAGGAAGGCGGAATCCCGGTGTAGCGGTGATATGCGTAGAT

>P3D08

GGGCTTGCGTCTGATTAGCTTGTGGTGGGATAATAGCCTACCAAGGCAATGATCAGTAGCTGGTCTGAGAGGACAAC
CAGCCAGACGGGGACTGAGACACGGCCCTGACTCCTACGGGAGGCAGCAGCAAGGAATTTTGC GCAATGGGCGAAAGC
CTGACGCAGCGACGCCGCGTGAAGGATGAAGGCCTTCGGGTGTAAACCTCTTTTCTCGAGGAAGTGACGGTACTCGA
GGAATAAGCTTGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGAAGGCAAACGTTATTCGAATTTACTGGGCG
TAAAGCGGGTGTAGACGGCCCTTAAGTCGGGTGTGAAATCTCCCGACTCAATCGGGAGAGGTCATTTCGATACTAAGG
AGCTTGAGGACAGCAGAGGAGGGCGGAATCCCGGTGTAGCGGTGATATGCGTAGAT

>p1a9

GGGCTTGCGTCTGATTAGCTAGTTGGTGGGGTAATGGCCTACCAAGGCGATGATCAGTAGCTGGTCTGAGAGGACAAC
CAGCCAGACGGGGACTGAGATACGGCCCTGACTCCTACGGGAGGCAGCAGCAAGGAATTTTGC GCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGATGAAGGCCTTCGGGTGTAAACCCCTTTTCTCGGGGAAGTGACGGTACTCGA
GGAATAAGTCTTGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGAAGGCAAACGTTATTCGGATTTACTGGGCG
TAAAGAAGTGCAGACGGTTTTTTAAGTCGGGTGTGAAATCTCCTGACTCAATCAGGAGAAGTCATTTCGATACTAAGG
AGCTTGAGGGCAGCAGAGGAAGGCGGAATCCCGGTGTAGCGGTGAAATGCGTAGAT

>p2c12

GGGCTTGCGTCTGATTAGCTAGTTGGTGGGGTAATGGCCTACCAAGGCGATGATCAGTAGCTGGTCTGAGAGGACAAC
CAGCCAGACGGGGACTGAGATACGGCCCTGACTCCTACGGGAGGCAGCAGCAAGGAATTTTGC GCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGATGAAGGCCTTCGGGTGTAAACCCCTTTTCTCGGGGAAGTGACGGTACTCGA
GGAATAAGTCTTGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGAAGGCAAACGTTATTCGGATTTACTGGGCG
TAAAGAAGTGCAGACGGTTTTTTAAGTCGGGTGTGAAATCTCCTGACTCAATCAGGAGAAGTCATTTCGATACTAAGG
AGCTTGAGGGCAGCAGAGGAAGGCGGAATCCCGGTGTAGCGGTGAAATGCGTAGAT

>p2d1

GGGCTTGCGTCCGATTAGCTAGTTGGTGGGGTAAAGGCTTACCAAGGCGGCGATCGGTAGCTGGTCTGAGAGGATGAT
CAGCCAGACGGGGACTGAGATACGGCCCCGACTCCTACGGGAGGCAGCAGCAAGGAATTTTGC GCAATGGGCGAAAGC
CTGACGCAGCGACGCCGCGTGGGGGAAGAAGGCCTTCGGGTGTAAACCCCTTTTCTCGAGGAAGCGACGGTACTCGG
GGAATAAGTCTCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCAAACGTTATCCGGATTTACTGGGCG
TAAAGGGGGCGCAGGCGGAATCTAAGTCGAGTGTGAAATCTCCTAACTCAATTAGGAGGGGTCATTTCGATACTAGGG
AGCTTGAGGGCAGCAGAGGAGGTGGAATTCCCGGTGTAGCGGTGAAATGCGTAGAT

>p3d2

GGGCTTGCGTCTGATTAGCTAGTTGGTGGGGTAATGGCCACCAAGGCGATGATCAGTAGCTGGTCTGAGAGGACAAC
CAGCCAGACAGGGACTGAGACACGGGCCCTGACTCCTACGGGAGGCAGCAGCAAGGAATTTTGC GCAATGGGCGAAAGC
CTGACGCAGCGACGCCGCGTGGGGGATGAAGGCCTTCGGGTGTAAACCCCTTTTCCCGAGGAAGTGACGGTACTCGG
GGAATACGTCTTGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCAAACGTTATTCGGATTTACTGGGCG
TAAAGCGGGCGTAGGCGGCTTCTTAAGTCTGTGTGAAATCCTACGGCTTAACCGTAGAATGCATGAGAACTGTGTA
ATCTAGAGTGTGAAAGAGGCAGGTGGAATTAGTGGTGTAGGGGTAAATCCGTAGAT

>P1C06

GGGCTTGCGTCCGATTAGCTAGTTGGTGGGGTAAAGGCCTACCAAGGCGAAGATCGGTAGCTGGTCTGAGAGGATGAA
CAGCCACACGGGGACTGAGATACGGCCCCGACTCCTACGGGAGGCAGCAGCAGGGAATATTGC GCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCCTTCGGGTGTAAACCCCTTTTCTTGGGGAAGGACGGTACCCAA
GGAATAAGCTTCGGCTAACTACGTGCCAGCAGCCGCGGTAAGACGTAGGAGGCAAACGTTATTCGGATTTACTGGGCG
TAAAGGGCGTGTAGGCGGCTTTCTAAGTTTCTGGTGAATCTCCCGCTCAACTGGGAGGGCGCTAGGAAAACCTGGAG
GGCTTGAGGGCAGCAGAGGGGGGTGGAATTCCCGGTGTAGTGGTGTATGCGTAGAT

>P2D12
GGGCTTGCGTCCGATTAGCTAGTTGGTGGGGTAAAGGCCTACCAAGGCGAAGATCGGTAGCTGGTCTGAGAGGATGAA
CAGCCACACGGGGACTGAGATACGGCCCCGACTCCTACGGGAGGCAGCAGCAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCTTTTCGGGTGTAAACCCCTTTTCTTGGGGAAGGGACGGTACCCAA
GGAATAAGCCTCGGCTAACTACGTGCCAGCAGCCGCGGTAAAGACGTAGGAGGCAAGCGTTATCCGGAATTACTGGGCG
TAAAGGGCGTGTAGGCGGCTTTCTAAGTTTCTGGTGAAATCTCCCGGCTCAACTGGGAGGGCGCTAGGAAACTGGAG
GGCTTGAGGGCAGCAGAGGGGGGTGGAATTCCCGGTGTAGTGGTGATATGCGTAGAT

>p1g11
GGGCTTGCGTCCGATTAGCTGGTTGGCGGGGTAATGGCCACCAAGGCGAAGATCGGTAGCTGGTCTGAGAGGATGGA
CAGCCACACGGGGGCTGAGATACGGCCCCGACTCCTACGGGAGGCAGCAGCAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCTTTTCGGGTGTAAACCCCTTTTCTTGGGGAAGGGACGGTACCCAA
GGAATAAGCTTCGGCTAACTACGTGCCAGCAGCCGCGGTAAAGACGTAGGAAGCGAGCGTTATCCGGAATTACTGGGCG
TAAAGGGCGTGTAGGCGGCTTTCTAAGTCTCTGGTGAAATCTCCCGGCTCAACTGGGAGGGCGCTGGGGAACTGGGG
GGCTTGAGGGCAGCAGAGGGGAGTGAATTCCCGGTGTAGTGGTGATATGCGTAGAT

>p1d10
GGGCTTGCGTCCGATTAGCTGGTTGGCGGGGTAATGGCCACCAAGGCGAAGATCGGTAGCTGGTCTGAGAGGATGGA
CAGCCACACGGGGGCTGAGATACGGCCCCGACTCCTACGGGAGGCAGCAGCAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCTTTTCGGGTGTAAACCCCTTTTCTTGGGGAAGGGACGGTACCCAA
GGAATAAGCTTCGGCTAACTACGTGCCAGCAGCCGCGGTAAAGACGTAGGAAGCGAGCGTTATCCGGAATTACTGGGCG
TAAAGGGCGTGTAGGCGGCTTTCTAAGTCTCTGGTGAAATCTCCCGGCTCAACTGGGAGGGCGCTGGGGAACTGGGG
GGCTTGAGGGCAGCAGAGGGGAGTGAATTCCCGGTGTAGTGGTGATATGCGTAGAT

>p2c4
GGGCTTGCGTCCGATTAGCTGGTTGGCGGGGTAATGGCCACCAAGGCGAAGATCGGTAGCTGGTCTGAGAGGATGGA
CAGCCACACGGGGACTGAGATACGGCCCCGACTCCTACGGGAGGCAGCAGCAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCTTTTCGGGTGTAAACCCCTTCTCTTGGGGAAGGGACGGTACCCAA
GGAATAAGCTTCGGCTAACTACGTGCCAGCAGCCGCGGTAAAGACGTAGGAAGCGAGCGTTATCCGGAATTACTGGGCG
TAAAGGGCGTGTAGGCGGCTTTCTAAGTCTCTGGTGAAATCTCCCGGCTCAACTGGGAGGGCGCTAGGGAACTGGGG
GGCTTGAGGGCAGCAGAGGGGAGTGAATTCCCGGTGTAGTGGTGATATGCGTAGAT

>p2d4
GGGCTTGCGTCCGATTAGCTGGTTGGCGGGGTAATGGCCACCAAGGCGAAGATCGGTAGCTGGTCTGAGAGGATGGA
CAGCCACACGGGGACTGAGATACGGCCCCGACTCCTACGGGAGGCAGCAGCAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCTTTTCGGGTGTAAACCCCTTCTCTTGGGGAAGGGACGGTACCCAA
GGAATAAGCTTCGGCTAACTACGTGCCAGCAGCCGCGGTAAAGACGTAGGAAGCGAGCGTTATCCGGAATTACTGGGCG
TAAAGGGCGTGTAGGCGGCTTTCTAAGTCTCTGGTGAAATCTCCCGGCTCAACTGGGAGGGCGCTAGGGAACTGGGG
GGCTTGAGGGCAGCAGAGGGGAGTGAATTCCCGGTGTAGTGGTGATATGCGTAGAT

>P3C05
GGGCTTGCGTCCGATTAGCTGGTTGGCGGGGTAATGGCCACCAAGGCGAAGATCGGTAGCTGGTCTGAGAGGATGGA
CAGCCACACGGGGACTGAGATACGGCCCCGACTCCTACGGGAGGCAGCAGCAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCTTTTCGGGTGTAAACCCCTTTTCTTGGGGAAGGGACGGTACCCAA
GGAATAAGCTTCGGCTAACTACGTGCCAGCAGCCGCGGTAAAGACGTAGGAAGCGAGCGTTATCCGGAATTACTGGGCG
TAAAGGGCGTGTAGGCGGCTTTCTAAGTCTCTGGTGAAATCTCCCGGCTCAACTGGGAGGGCGCTAGGGAACTGGGG
GGCTTGAGGGCAGCAGAGGGGAGTGAATTCCCGGTGTAGTGGTGATATGCGTAGAT

>p3b2
GGGCTTGCGTCCGATTAGCTGGTTGGCGGGGTAATGGCCACCAAGGCGAAGATCGGTAGCTGGTCTGAGAGGATGGA
CAGCCACACGGGGACTGAGATACGGCCCCGACTCCTACGGGAGGCAGCAGCAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCTTTTCGGGTGTAAACCCCTTTTCTTGGGGAAGGGACGGTACCCAA
GGAATAAGCTTCGGCTAACTACGTGCCAGCAGCCGCGGTAAAGACGTAGGAAGCGAGCGTTATCCGGAATTACTGGGCG
TAAAGGGCGTGTAGGCGGCTTTCTAAGTCTCTGGTGAAATCTCCCGGCTCAACTGGGAGGGCGCTAGGGAACTGGGG
GGCTTGAGGGCAGCAGAGGGGAGTGAATTCCCGGTGTAGTGGTGATATGCGTAGAT

>P2C10
GGGCTTGCGTCCGATTAGCTAGTTGGCGGGGTAATGGCTCACCAAGGCGAAGATCGGTAGCTGGTCTAAGAGGATGGA
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCTCTCGGGTGTAAACCCCTTTTCTTGGGGAAGGGACGGTACCCAA
GGAATAAGCCTCGGCTAACTACGTGCCAGCAGCCGCGGTAAAGACGTGGGAGGCGAGCGTTATCCGGAATTATTGGGCG
TAAAGGGCGTGTAGGCGGCCCTTTAAGTTCTCGGTGAAATCTCCCGGCTCAACCGGGAGGGGGCTGAGAATACTGGGG
GGCTTGAGGGCAGCAGAGGGGAGTGAATTCCCGGTGTAGTGGTGATATGCGTAGAT

>P3C04
GGGCTTGCGTCCGATTAGCTAGTTGGCGGGGTAATGGCTCACCAAGGCGAAGATCGGTAGCTGGTCTAAGAGGATGGA
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCTCTCGGGTGTAAACCCCTTTTCTTGGGGAAGGGACGGTACCCAA

GGAATAAGCCTCGGCTAACTACGTGCCAGCAGCCGCGGTAAGACGTGGGAGGCGAGCGTTATCCGGAATTATTGGGCG
TAAAGGGCGTGTAGGCGGCCCTTTAAGTTCTCGGTGAAATCTCCCGGCTCAACCGGGAGGGGCTGAGAATACTGGGG
GGCTTGAGGGCAGCAGAGGGAGATGGAATTCGCGGTGTAGTGGTGATATGCGTAGAT

>P3C07

GGGCTTGCGTCCGATTAGCTCGTTGGTGGGGTAACGGCCTACCAAGGCTGAGATCGGTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCCTTCGGGTGTAAACCCCTTTTCTCGGGGAAGTGACGGTACCCGA
GGAATAAGCCTCGGCTAACTACGTGCCAGCAGCCGCGGTAAGACGTAGGAGGCGAGCGTTATCCGGATTTACTGGGCG
TAAAGGGCGTGTAGGCGGCTCTTTAAGTCTGTGGTGAAATCTCCCGGCTCAACCGGGAGGGTGTCATAGAACTGGGG
AGCTTGAGGGCAGCAGAGGGAGGTGGAATTCGCGGTGTAGTGGTGATATGCGTAGAT

>P3E10

GGGCTTGCGTCCGATTAGCTCGTTGGTGGGGTAACGGCCTACCAAGGCTGAGATCGGTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCCTTCGGGTGTAAACCCCTTTTCTCGGGGAAGTGACGGTACCCGA
GGAATAAGCCTCGGCTAACTACGTGCCAGCAGCCGCGGTAAGACGTAGGAGGCGAGCGTTATCCGGATTTACTGGGCG
TAAAGGGCGTGTAGGCGGCTCTTTAAGTCTGTGGTGAAATCTCCCGGCTCAACCGGGAGGGTGTCATAGAACTGGGG
AGCTTGAGGGCAGCAGAGGGAGGTGGAATTCGCGGTGTAGTGGTGATATGCGTAGAT

>P3C06

GGGCTTGCGTCCGATTAGCTCGTAGGTGGGGTAACGGCCTACCAAGGCTGAGATCGGTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGGACTGAGACACGGACCAGACTCCTACGGGAGGCAGCAGCAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCCTTCGGGTGTAAACCCCTTTTCTCGGGGAAGTGACGGTACCCGA
GGAATAAGCCTCGGCTAACTACGTGCCAGCAGCCGCGGTAAGACGTAGGAGGCGAGCGTTATCCGGATTTACTGGGCG
TAAAGGGCGTGTAGGCGGCTCTTTAAGTCTGTGGTGAAATCTCCCGGCTCAACCGGGAGGGTGTCATAGAACTGGGG
AGCTTGAGGCCAGCAGAGGGAGGTGGAATTCGCGGTGTAGTGGTGATATGCGTAGAT

>p2g1

GGGCTTGCGTCCGATTAGCTCGTTGGTGGGGTCACGGCCTACCAAGGCTGCGATCGGTAGCTGGTCTGAGCGGATGAT
CAGCCACCTTGGGCACTGCACCCGGCCACCTCCTCCGGGCAGCAGCAGCAGGGAATCTTGCGCCATGGGCGAAAGCC
TGACGCAGCAACGCCGCGTGGGGGACGAAGGCCTTCGGGTGTAAACCCCTTTTCTCGGGGCGGTGACGGTCCCCGAG
GAATAAGCCTCGGCTCACTACGTGCCAGCCGCGCGGTCCGACGTAGGAGGCGAGCGTCATCCGGCTTTACTGGGCGT
CAAGGGCGTGTAGGCGGCTCTTCAAGTCTGTGGTGAAATCTCCCGGCTCAACCGGGAGGGTGTCATAGAACTGGTGA
GCTTGAGGCTAGCAGAGGGAGTTGGAATTCGCGTGTAGTGGTGATATGCGTAGAT

>p1f4

GGGCTTGCGTCCGATTAGCTAGTTGGTGTGGTAATGGCGCACCAAGGCTGAGATCGGTAGCTGGTTTGAGAGGACGAT
CAGCCACACTGGGATTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGCAGGGGATATTGCGCAATGGGTGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGTCTTCGGGTGTAAACCCCTTTTCTCGGGGAAGTGACGGTACCCGA
GGAATAAGTCTCGGCTAACTACGTGCCAGCAGCCGCGGTAAGACGTAGGAGGCGAGCGTTATCCGGATTTACTGGGCG
TAAAGGGCGTGTAGGCGGCTTCCAAGTCTTTGGTGAAATCTCCCGACTCAATTGGGAGGGTGCCATAGAACTGGAG
GGCTTGAGGACAGCAGAGGGGGGTGGAATTCGCGGTGTAGTGGTGATATGCGTAGAT

>p1h12

GGGCTTGCGTCCGATTAGTTAGTTGGTGGGGTAATGGCCACCAAGGCGTAGATCGGTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAGGGGATATTGCGCAATGGGTGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGATGACGGCCTTCGGGTGTAAACCCCTTTTCTCAGGGGAAGTGACGGTACCTGA
GGAATAAGTCTCGGCTAACTACGTGCCAGCAGCCGCGGTAAGACGTAGGAGGCAAGCGTTATCCGGATTTACTGGGCG
TAAAGGGCGTGTAGGCGGCTCGTTAAGTCTCTGGTGAAATCTCCCGGCTCAACCGGGAGGGTGCCGAGATACTGGCG
GGCTTGAGGGCAGTAGAGGAGAGCGGAATTCGCGGTGTAGTGATGATATGCGTAGAT

>p1h1

GGGCTTGCGTCCGATTAGCTGGTTGGCGGGGTAATGGCCACCAAGGCGAAGATCGGTAGCTGATCTGAGAGGATGGA
CAGCCACACGGGACTGAGATACGGCCCCGACTCCTACGGGAGGCAGCAGCAGGGAATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGAAGAAGGCCTTCGGGTGTAAACCCCTTTTCTTGGGGAAGTGACTGTACCCAA
GGAGGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAAGACGTAGGAAGCGAGCGTTATCCGGAATTACTGGGCG
TAAAGGGCGTGTAGGCGGCTTTCTAAGTCTCTGGTGAAATCTCCCGGCTCAACTGGGGAAGTGCTTATCGAACTGCTC
AGCTTGAGCACGGGAGAGGGCACTGGAATTCCTGGTGTAGCGGTGAAATGCGTATAT

>P3D10

GGGCGCTAAATCCCATCAGGTAGTTGGTGGGGTAATGGCCTACCAAGCTTAAGACGGGTAGCTGGTCTGAGAGGATGGT
CAGCCAGACGGGACTGAGACACGGCCCCGACTCCTACGGGAGGCAGCAGCAAGGAATCTTGCGCAATGGGCGAAAGC
CTGACGCAGCGACACCGCGTGGGGGAAGACGGCTTCGGGTGTAAACCCCTTTTCTTGGTGAAGGGACGGTAGCCAA
GGAATAAGCCTCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAGCGTTATCCGGATTTACTGGGCG
TAAAGAGCGGTAGGCGGCTTTTAAGTTGGTGGTGAAAGCTCTGGCTCAACTGGGAGAGGTCCATCAAACTGGGA
GGCTTGAGGACAGCAGAGGAGGGTAGAATTCGCGGTGTAGCGGTGAAATGCGTAGAT

>p2c2
GGGTTTCGCGCCTATCAGGTTGTTGGTGGGGTAACGGCCACCAAGCCGATGACGGGTAGCTGGTCTGAGAGGACGAC
CAGTCAGAGGGGACTGCGACACGGCCCCACTCCTACGGGAGGCAGCAGCAGGGAATCTTGGGCAATGGGCGAAAGC
CTGACCCACGACGACCCGCTGGGGGAGGAAGGCCTTCGGGTGTAAACCCCTTTTCCAGGGGAAGGGACGGTACCCTG
GGAATAAGCCTCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCAAGCGTTGTCCGGATTTACTGAGCG
TAAAGGGCGTGTAGGCGGCCAAGTAAGTCTCCCGTGAAATCTCCCGGCTTAACTGGGAGGGGTCAGGGGATACTGCTA
GGCTTGAGGACAGCAGAGGAGGGTGAATTTCCCGATGTAGCGGTGACATGCGTAGAT

>P2F08
GGGCTTGCGTCCTATCAGGTAGTTGGTGAGGTAAAGGCCACCAAGCCGAAGACGGGTAGCTGGTCTGAGAGGATGGT
CAGCCCACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGGATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGTGAAGAAGGCCTTCGGGTGTAAAGCCCTTTTCCGGGGGAAGTGACGGTACCCCG
GGAATAAGCCTCGGCTAACTACGTGCCAGCAGCCGCGGTAAGACGTAGGAGGCGAGCGTTATCCGGATTTACTGGGCG
TCAAGAGCGCTAAGGCGGCAGGGAAGTCAGCCCTGAAATCTCCCGGCTCAACCGGGAGGGGAGGGTTGATACTCCCC
CGCTTGAGGGGGGAGAGGGAGGTGGAATTTCCGGAGTAGTGGTGAAATGCGTAGAT

>p2g10
GGGCTTGCGTCCTATCAGGTAGTTGGTGAGGTAAAGGCCACCAAGCCGAAGACGGGTAGCTGGTCTGAGAGGATGGT
CAGCCCACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAAGGGATATTGCGCAATGGGCGAAAGC
CTGACGCAGCAACGCCGCGTGGGTGAAGAAGGCCTTCGGGTGTAAAGCCCTTTTCCGGGGGAAGTGACGGTACCCCG
GGAGGAAGCCTCGGCTAACTACGTGCCAGCAGCCGCGGTAAGACGTAGGAGGCGAGCGTTATCCGGATTTACTGGGCG
TAAAGAGCGCTAAGGCGGCAGGGAAGTCAGCCCTGAAATCTCCCGGCTCAACCGGGAGGGGAGGGTTGATACTCCCC
CGCTTGAGTATGGGAGAGGGAGGTGGAATTTCCGGAGTAGAGGTGAAATGCGTAGAT

>p2f11
GGGCCCCGCTACCATTTAGCTTGTGGTAGGGTAATGGCCTACCAAGGCGACGATGGTTAGCTGGTCTGAGAGGATGAT
CAGCCCACTGGAAGTGAACACGGTCCAGACTCCTACGGGAGGCAGCAGTGAAGGAATTTTTCGCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGGGTGAAGAAGGCCTTCGGGTGTAAAGCCCTTTTCCGGGGGAAGTGATGGTACCCCG
GGAATAAGCCTCGGCTAACTACGTGCCAGCAGCCGCGGTAAGACGTAGGAGGCGAGCGTTATCCGGATTTACTGGGCG
TAAAGAGCGCTAAGGCGGCAGGGAAGTCAGATGTGAAATCTCCCGGCTCAACCGGGAGGGGAGGGTTGATACTCCCC
CGCTTGAGGGGGGAGAGGGAGGTGGAATTTCCGGAGTAGTGGTGAAATGCGTAGAT

>P1D05
GGGCTCGCGCCGATTAGCTAGTTGGTGTGGTAACGGCGCACCAAGGCGACGATCGGTAGCTGGTCTGAGAGGATGAT
CAGCCCACTGGAAGTGAAGACGGTCCAGACTCCTACGGAAGGCAGCAGTAGGGAATATTGCACAATGGGGGAAACC
CAAATGCGGCAACGCCGCGTGAGCGAAGAAGGCCTTCGGGTGCTAAAGCTCTTTTCCGGGGGAAGGGACGGTACCCCG
GGAACAAGTCTCGGCTAACTACGTGCCAGCAGCCGCGGTAAGGCGTAGGAGGCGAGCGTTATCCGGATTTACTGGGCG
TAAAGGGCAAGTAGGCGGTAGCCAAAGTTGGGCGTGACAGATCTCGGCCTAACTGGGAAAGGTGCTCCAAACTAGGC
TACTAGAGGCGAGGAGAGGGAAGTGAATTTCTAGTGTAGCGGTGGGATGCTCAGAT

>p1h4
GAGCCTGCGGCCGATTAGCTAGTTGGCAGGGTAAAGCCCTACCAAGGCGACGATCGGTAGCTGGTCTGAGAGGATGAT
CAGCCCACTGGAAGTGAACACGGTCCAGACTCCTACGGGGGCGAGCAGTAAGGGATATTGCACAGTGGGGGAAACC
CTGATGCAGCGACGCCGCGTGAGCGAAGAAGGCCTTCGGGTGTAAAGCTCTTTTCTGTGGGACGGGACGGTACCACA
GGAACAAGCCTCGGCTAACTACGTGCAGCAGCCGCGGTAACACGTAGGAGACGAGCGTTATCCGGATTCTACTGGGCGT
AAAGGGCAAGTAGGCGGTTCGAAAGTTGGGCGTGACAATCTCTGGCTTAACTAGGGAAGGACGTTCAAAACTAGGCA
ACTAGAGGCGAGGAGAGGAAAGCGGAATTTCTAGTGTAGCGGTGGAATGCTCAGAT

>P3H03
GGGCCTGCGACCGATTAGCTAGTTGGTGAGGTAAAGGCTACCAAGGCCGAGATCGGTAGCTGTTCTGAGAGGAAGAC
CAGCCACACGGGAAGTGAAGACAGGTCCCGACTCCTACGGGGGCGAGCAGTAGGAATATTGCACAATGGGGGAAACC
CTGATGCAGCAACGCCGCGGAAGGAAGAAGGTCTTTGGATTGTAAACTTCTTTTAAACGGGGGAAGGGACGGTACCCGT
TGAATAAGTCACGGCTAACTACGTGCCAGCAGCCGCGGTAACACGTAGGTGGCGAGCGTTATCCGGAATTATTGGGCG
TAAAGGGCGCGCAGGCGGCGCTCAAGTCTGACGTGAAAGCTCCTGGCTTAACTGGGAGAGGTGCTTGGAAGTGAAG
TACTTGAGGCAATGAGAGGGGTGCGGAATTTCCCGGTGTAGTGGTGGAATGCG

>p2e1
GGGCCTGCGGCGGATTAGGTAGTTGGTGGGGTAACGGCCTACCAAGCCGATGATCCGTAGGGGGCGTGAGAGCGTGAC
CCCCCACTGGCACTGAGACACGGGCCAGACTCCTACGGGAGGCAGCAGTAGGAATATTGCACAATGGGCGAAAGC
CTGATGCAGCAACGCCGCGTGAGGGACGAAGGCCTTCGGGTGCTAAACCTCTTTTCTGGGGGACGGGACGGTACCTCA
GGAATAAGACTCGGCTAACTACGTGCCAGCAGCCGCGGTAACACGTAGGAGGCGAGCGTTATCCGGATTTACTGGGCG
TAAAGCGCGCGCAGGCGGCTTCGTAAGTCGGACGTGAAAGCTCCTGGCTTAACTGGGAGATGTCGTTTCGATACTGCGG
GGCTTGAGGTGGGAGAGGGGTGTGGAATTTCCCGGTGTAGTGGTGGAATGCGTAGAT

>P2B04
GAGCCCCGCTCCTATTAGCTAGTTGGTAAGGTAAAGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAA
CGGCCCACTGGAAGTGAAGACAGGTCCAGACTCCTACGGGAGGCAGCAGTAGGGAATTTTTCGCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGGATGATGAAGGCCTTCGGGTGCTAAAGTCTGTTCTGGGGGAAGTGACGGTACCCCA

GCAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGACGGTTCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGG
AACTTGAGGGTAGAAGAGGAAAGTGAATTCCCAGGTGTAGCGGTGAAATGCGTAGAT

>p1f2

GAGCCCCGCTCCTATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAA
CGACCACACTGGAAGTGAAGACAGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGCACAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGGATGATGAAGGCCTTCGGGTCGTAAAGTCTGTCTGGGGGAAGTGACGGTACCCCA
GCAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCGT
AAAGAGCTCGTAGACGGTTCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGGA
ACTTGAGGGTAGAAGAGGAAAGTGAATTCCCAGGTGTAGCGGTGAAATGCGTAGAT

>P2E05

GAGCCCCGCTCCTATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAA
CGGCCACACTGGAAGTGAAGACAGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGCACAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGGATGATGAAGGCCTTCGGGTCGTAAAGTCTGTCTGGGGGAAGTGACGGTACCCCA
GCAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGACGGTTCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGG
AACTTGAGGGTAGAAGAGGAAAGTGAATTCCCAGGTGTAGCGGTGAAATGCGTAGAT

>p1d12

GAGCCCCGCTCCTATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAA
CGGCCACACTGGAAGTGAAGACAGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGCACAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGGATGATGAAGGCCTTCGGGTCGTAAAGTCTGTCTGGGGGAAGTGACGGTACCCCA
GCAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGACGGTTCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGG
AACTTGAGGGTAGAAGAGGAAAGTGAATTCCCAGGTGTAGCGGTGAAATGCGTAGAT

>p3e3

GAGCCCCGCTCCTATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAA
CGGCCACACTGGAAGTGAAGACAGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGCACAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGGATGATGAAGGCCTTCGGGTCGTAAAGTCTGTCTGGGGGAAGTGACGGTACCCCA
GCAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGACGGTTCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGG
AACTTGAGGGTAGAAGAGGAAAGTGAATTCCCAGGTGTAGCGGTGAAATGCGTAGAT

>p3b9

GAGCCCCGCTCCTATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAA
CGGCCACACTGGAAGTGAAGACAGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGCACAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGGATGATGAAGGCCTTCGGGTCGTAAAGTCTGTCTGGGGGAAGTGACGGTACCCCA
GCAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGACGGTTCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGG
AACTTGAGGGTAGAAGAGGAAAGTGAATTCCCAGGTGTAGCGGTGAAATGCGTAGAT

>p2c1

GAGCCCCGCTCCTATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAA
CGGCCACACTGGAAGTGAAGACAGGTCCAGACTCCTACTGGAGGCAGCAGTGGGGAATATTGCGCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGGACGATGAAGGCCTTCGGGTCGTAAAGTCTGTCTGGGGGAAGTGACGGTACCCCA
GCAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGACGGTTCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGG
AACTTGAGGGTAGAAGAGGAAAGTGAATTCCCAGGTGTAGCGGTGAAATGCGTAGAT

>p1b2

GAGCCCCGCTCCTATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAA
CGGCCACACTGGAAGTGAAGACAGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGCACAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGGACGATGAAGGCCTTCGGGTCGTAAAGTCTGTCTGGGGGAAGTGACGGTACCCCA
GCAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGACGGTTCCGAAAGTCAGCTGTGAAATCACCAGACTCAACCTGGTAGGGCCAGTTGAAACTTCGG
AACTTGAGGGTAGAAGAGGAAAGTGAATTCCCAGGTGTAGCGGTGAAATGCGTAGAT

>p3h5

TAACCCGCGTACCATTAGCTTGTAGGTAAGGTAACGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAAAGGACGAA
CGGACACACTGGAAGTGAACACAGGTCCAGACTCCTACGGGAGGCAGAAAGTGGGGAATTTTGCACAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGGACGATGAAGGCCTTCGGGTCGTAAAGTCTGTACTGAGGGAAGTGACGGTACCCCA
GCAGAAAGCTCCGGCCATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCGT
AAAGAGCTCGTAGACGGTTCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGGA
ACTTGAGGGTAGAAGAGGAAAGTGAATTCCCAGGTGTAGCGGTGATATGCGTAGAT

>p1h8
GAGCCCCGCTCCTATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATGGGTAGCCGGCCTGAGAGGGCGAC
CGGCCACATTGGGACTGAGATACGGCCCCAACTCCTACGGGAGGCAGCAGTGGGGAATATTGCGCAATGGACGCAAGT
CTGACGCAGCAACGCCGCTGGATGATGAAGGCCTTCGGGTCGTAAAGTCCTGTTCTGGGGGAAGTGACGGTACCCCA
GCAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGACGGTCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGG
AACTTGAGGGTAGAAGAGGAAAGTGAATTCCCGTGTAGCGGTGAAATGGTAGAT

>P3F04
GACCCGCGTCTATTAGCTAGTTGGTGAGGAATGGGTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAACGG
CCACACTGGAGCTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGC GCAATGGGGGAAACCTG
ACGCAGCAACGCCGCGTGGACGATGAAGGCCTTCGGGTCGTAAAGTCCTGTTCTGGGGGAAGTGACGGTACCCAGGA
GAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCGTAA
AGAGCTCGTAGACGGTCCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGGAAC
TTGAGGGTAGAAGAGGAAAGTGAATTCCCGTGTAGCGGTGAAATGGTAGAT

>P3D07
GAGCCCCGCTCCTATTAGCTAGTTGGTGAGGTAATGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAA
CGGCCACACTGGAGCTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGC GCAATGGGGGAAAC
CTGACGCAGCAACGCCGCGTGGACGATGAAGGCCTTCGGGTCGTAAAGTCCTGTTCTGGGGGAAGTGACGGTACCCCA
GGAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGACGGTCCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGG
AACTTGAGGGTAGAAGAGGAAAGTGAATTCCCGTGTAGCGGTGAAATGGTAGAT

>p3g1
GAGCCCCGCTCCTATTAGCTAGTTGGTGAGGTAATGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAA
CGGCCACACTGGAGCTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGC GCAATGGGGGAAAC
CTGACGCAGCAACGCCGCGTGGACGATGAAGGCCTTCGGGTCGTAAAGTCCTGTTCTGGGGGAAGTGACGGTACCCCA
GGAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGACGGTCCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGG
AACTTGAGGGTAGAAGAGGAAAGTGAATTCCCGTGTAGCGGTGAAATGGTAGAT

>p2f1
GAGCCCCGCTCCTATTAGCTAGTTGGTGAGGTAATGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAA
CGGCCACACTGGAAGTGAAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGC GCAATGGGGGAAAC
CTGACGCAGCAACGCCGCGTGGACGATGAAGGCCTTCGGGTCGTAAAGTCCTGTTCTGGGGGAAGTGACGGTACCCCA
GGAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGACGGTCCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGG
AACTTGAGGGTAGAAGAGGAAAGTGAATTCCCGTGTAGCGGTGAAATGGTAGAT

>p2g11
GAGCCCCGCTCCTATTAGCTAGTTGGTGAGGTAATGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAA
CGGCCACACTGGAAGTGAAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGC GCAATGGGGGAAAC
CTGACGCAGCAACGCCGCGTGGACGATGAAGGCCTTCGGGTCGTAAAGTCCTGTTCTGGGGGAAGTGACGGTACCCCA
GGAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAACACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGACGGTCCCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGCCAGTTGAAACTTCGG
AACTTGAGGGTAGAAGAGGAAAGTGAATTCCCGTGTAGCGGTGAAATGGTAGAT

>P2D07
GAGCCCCGCTCCTATTAGCTAGTTGGTGAGGTAACAGCTTACCAAGGCAACAATGGGTAGCCGGTCTGAGAGGATGAA
CGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGC GCAATGGGGGAAAC
CTGACGCAGCAACGCCGCGTGAACGATGAAGGCCTTCGGGTCGTAAAGTCTGTTCTGGGGGAAGTGACGGTACCCCA
GCAGAAAGCTCCGGCCAATTATGTGCCAGCAGCCGCGGTAATACATAAGGAGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGGCGGTACCGAAAGTCAGCTGTGAAATCACCAGGCTTAACCTGGTAGGGTCAGCTGAAACTTCGG
GACTTGAGGGTAGAAGAGGAAAGCGGAATTCCCGTGTAGCGGGTGAATGGTAGAT

>P3G05
GAGCCTGCGTCCTATTAGTTAGTTGGTGGGGTAATGGCCTACCAAGACTACAATGGGTAGCCGGTCTGAGAGGATGTA
CGGCCACACTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCGCAATGGGGGAAAC
CTGACGCAGCAGCAGCCGCTGGATGATGAAGGCCTTTGGGTTGTAAAATCCTGTACTGGGGGAAGTGACGGTACCCCA
AGAGAAAGCTCCGGCTAATTATGTGCCAGCAGCCGCGGTAATACATAAGGAGCAAGCGTAATCCGGAATTATTGGGCG
TAAAGAGCTCGTAGGCGTTCTTAAAAGTCAGGTGTGAAATTAACCTGGCTCAACCTGTTAAGGTCATCTGAAACTTTAA
GACTTGAGGATAGAAGAAGGAAGTGAATTCCCTGTAAGCGGTTGAAATGCGTAGAT

>p1f12
GAGCCTGCGTCCTATTAGTTAGTTGGTGGGGTAATGGCCTACCAAGACTACAATGGGTAGCCGGTCTGAGAGGATGTA
CGGCCACACTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCGCAATGGGGGAAAC
CTGACGCAGCAGCAGCCGCTGGATGATGAAGGCCTTTGGGTTGTAAAATCCTGTTCTGGGGGAAGTGACGGTACCCCA

AGAGAAAGCTCCGGCTAATTATGTGCCAGCAGCCGCGGTAATACATAAGGAGCAAGCGTTATCCGGAATTATTGGGCG
TAAAGAGCTCGTAGGCGGTCTTAAAGTCAGGTGTGAAATTAAGTGGCTCAACCTGTTAAGGTCATCTGAACTTTAA
GACTTGAGGATAGAAGAGGAAAGTGGAATTCCTCGTGTAGCGGTGGAATGCGTAGAT

>P3G10

GAGCCTGCGTCCTATTAGTTAGTTGGTGGGGTAATGGCCTACCAAGACTACAATGGGTAGCCGGTCTGAGAGGATGTA
CGGCCACACTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCGCAATGGGGGAAACC
CTGACGCAGCGACGCCGCGTGGATGATGAAGGCCTTTGGGTGTGTAAGTAACTGTTCTGGGGGAAGTGACGGTACCCCA
AGAGAAAGCTCCGGCTAATTATGTGCCAGCAGCCGCGGTAATACATAAGGAGCAAGCGTTATCCGGAATTATTGGGCG
TAAAGAGCTCGTAGGCGGCTTAAAGTCAGGTGTGAAATTAAGTGGCTCAACCTGTTAAGGTCATCTGAACTTTAA
GACTTGAGGATAGAAGAGGAAAGTGGAATTCCTCGTGTAGCGGTGGAATGCGTAGAT

>p3h11

GAGCCCGCTCCTATTAGCTAGTTGGTAAGGTAACGGCTTACCAAGGCGACGATGGGTAGCCGGTCTGAGAGGACGAA
CGGCCACACTGGAAGTGAAGACAGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGCAATGGGGGAAACC
CTGACGCAGCAACGCCGCGTGGATGATGAAGGCCTTCGGGTGCTAAAGTCTGTTCTGGGGGAAGTGACGGTACCCAA
GGAATAAGCTCGGCCAAGTACGTGCCAGCAGCCGCGGTAATACGTAGGAGGCGAAGCTTATTCGGATTTACTGGGCG
TAAAGGGAGCGTAGGCGGTTCTTAAAGTCAGATGTGAAATCTTCTGATTCAATCAGAAGGGTCAATTTGATACTGAGG
GACTTGAGAACAGCAGAGGGGAAATGGAATCCCCAGTGTAGCGGTGAAATGCGTAGAT

>p2c8

CGGCCGATGCCTGATTAGCTAGTTGGTGGGGTAAAGGCCACCAAGGCGACGATCAGTAGCTGGTCTGAGAGGACGAT
CAGCCACACTGGGGCTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGACAATGGGGGCAACC
CTGATCCAGCAATGCCGCGTGTGTGAAGAAGGCCTTCGGGTGTGTAAGCACTTTTGTCCGGAAAGTGACGGTACCCGA
AGAATAAGCACCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCG
TAAAGCGTGCGCAGGCGGTTTGTGAAGACAGGCGTGAAATCCCCGGGCTTAACCTGGGAATTGCGCTTGTGACTGCAA
GGCTAGAGTGCGTCAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAGAG

>P2C09

CGGCCGATGCCTGATTAGCTAGTTGGTGGGGTAAAGGCCACCAAGGCGACGATCAGTAGCTGGTCTGAGAGGACGAT
CAGCCACACTGGGGCTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGACAATGGGGGCAACC
CTGATCCAGCAATGCCGCGTGTGTGAAGAAGGCCTTCGGGTGTGTAAGCACTTTTGTCCGGAAAGTGACGGTACCCGA
AGAATAAGCACCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCG
TAAAGCGTGCGCAGGCGGTTTGTGAAGACAGGCGTGAAATCCCCGGGCTTAACCTGGGAATTGCGCTTGTGACTGCAA
GGCTAGAGTGCGTCAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAGAG

>p2a8

CGGCCGATGTCTGATTAGCTAGTTGGTGGGGTAAAGGCCACCAAGGCGACGATCAGTAGCTGGTCTGAGAGGACGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGACAATGGGGGCAACC
CTGATCCAGCAATGCCGCGTGTGTGAAGAAGGCCTTCGGGTGTGTAAGCACTTTTGTCCGGAAAGTGACGGTACCCGA
AGAATAAGCACCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGCGAGCGTTAATCGGAATTGCTGGGCG
TAAAGCGTGCGCAGGCGGTTTGTGAAGACAGGCGTGAAATCCCCGGGCTTAACCTGGGAATTGCGCTTGTGACTGCAA
GGCTAGAGTGCGTCAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAGAG

>P1H05

CGGCCGATGTCTGATTAGCTAGTTGGTGGGGTAAAGGCCACCAAGGCGACGATCAGTAGCTGGTCTGAGAGGACGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGACAATGGGGGCAACC
CTGATCCAGCAATGCCGCGTGTGTGAAGAAGGCCTTCGGGTGTGTAAGCACTTTTGTCCGGAAAGTGACGGTACCCGA
AGAATAAGCACCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCG
TAAAGCGTGCGCAGGCGGTTTGTGAAGACAGGCGTGAAATCCCCGGGCTTAACCTGGGAATTGCGCTTGTGACTGCAA
GGCTAGAGTGCGTCAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAGAG

>p1b10

CGGCCGATGTCTGATTAGCTAGTTGGTGGGGTAAAGGCCACCAAGGCGACGATCAGTAGCTGGTCTGAGAGGACGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGACAATGGGGGCAACC
CTGATCCAGCAATGCCGCGTGTGTGAAGAAGGCCTTCGGGTGTGTAAGCACTTTTGTCCGGAAAGTGACGGTACCCGA
AGAATAAGCACCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCG
TAAAGCGTGCGCAGGCGGTTTGTGAAGACAGGCGTGAAATCCCCGGGCTTAACCTGGGAATTGCGCTTGTGACTGCAA
GGCTAGAGTGCGTCAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAGAG

>p1b1

CGGCCGATGTCTGATTAGCTAGTTGGTGGGGTAAAGGCCACCAAGGCGACGATCAGTAGCTGGTCTGAGAGGACGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGACAATGGGGGCAACC
CTGATCCAGCAATGCCGCGTGTGTGAAGAAGGCCTTCGGGTGTGTAAGCACTTTTGTCCGGAAAGTGACGGTACCCGA
AGAATAAGCACCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCG
TAAAGCGTGCGCAGGCGGTTTGTGAAGACAGGCGTGAAATCCCCGGGCTTAACCTGGGAATTGCGCTTGTGACTGCAA
GGCTAGAGTGCGTCAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAGAG

>p3d3
CGGCCGACGTCTGATTAGCTAGTTGGTGGGGTAAAGGCCACCAAGGCGACGATCAGTAGCTGGTCTGAGAGGACGAT
CAGCCACACTGGGACTGAGACACGGGCCCTTACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGACAATGGGGGCAACC
CTGATCCAGCAATGCCGCGTGTGTGAAGAAGGCCTTCGGGTGTAAAGCACTTTTGTCCGGAAGTGACGGTACCGGA
AGAATAAGCACCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCG
TAAAGCGTGCGCAGGCGGTTTTGTAAAGACAGGCGTGAAATCCCCGGGCTTAACCTGGGAATTGCGCTTGTGACTGCAA
GGCTAGAGTGCGCTCAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAGAG

>p2a3
CGGCCGATGGCAGATTAGGTAGTTGGTGGGATAAAAGCTTACCAAGCCGACGATCTGTAGCTGGTCTGAGAGGACGAC
CAGCCACACTGGGACTGAGACACGGNCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGACAATGGGGGCAACC
CTGATCCAGCAATGCCGCGTGTGTGAAGAAGGCCTTCGGGTGTAAAGCACTTTTGTCCGGAAGTGACGGTACCGGA
AGAATAAGCACCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAAAGTGCGAGCGTTAATCGGAATTACTGGGCG
TAAAGCGTGCGCAGCCGTTATGTAAAACAGATGTGAAATCCCCGGGCTCAACCTGGGAATGCATTTGTGACTGCAT
GGCTAGAGTACGTTAGAGGGGGATGGAATTCCGCGTGTAGCAGTGAAATGCGTAGAT

>P3B12
AGGCCCAAGTGAGATTGGCTAGTTGGTGAGGTAAGAGCTCACCAAGGCGACGATCTCTAGCTGTTCTGAGAGGAAGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGGAAACC
CTGATGCAGCCATGCCGCGTGTGTGAAGAAGGCCTTCGGGTGTACAGCACTTTCAGTTGTGAGGTGACGTAACAACA
GAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCGAGCGTTAATCGGAATTACTGGGCGT
AAAGCGCACGCAGGCGGTTTTGTAAAGCTAGATGTGAAAGCCCCGGGCTCAACCTGGGAAGGCCATTTAGAAGTGGCAG
ACTAGAGTCTTGAGAGGGGAGTGGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG

>p2h3
ATGCCCAGGTGGGATTAGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGAAGTGAAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGGAAACCC
TGATGCAGCCATGCCGCGTGTGTGAAGAAGGCCTTCGGGTGTAAAGCACTTTCAGCGAGGAGGTGACGTTACTCGCA
GAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGT
AAAGCGCACGCAGGCGGTTTGATAAGTTAGATGTGAAAGCCCCGGGCTCAACCTGGGAATTGCATTTAAAAGTGTCCA
GCTAGAGTCTTGATAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG

>p1h3
GGGCTTGCGTCCGATTAGTTAGTTGGTGGGGTAATGGCCCACCAAGGCGTAGATCGGTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGCAGGGGATATTGCGCAATGGGTGAAAGC
CTGACGCAGCAACGCCGCGTGGGGGATGACGGCCTTCGGGTGTAAACCCCTTTCTCAGGGAAGTGACGGTACCTGAG
GAATAAGTCTCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGT
AAAGCGCACGCAGGCGGCTCGTTAAGTTAGATGTGAAAGCCCCGGGCTCAACCTGGGAATTGCATTTAAAAGTGTCCA
GCTAGAGTCTTGATAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAG

>p2g3
CGGCCCGCGTTAGATTAGTAGTTGGTGGGGTAATGGCCTACCAAGTCTACGATCTATAGCTGGTTTGAGAGGATGAT
CAGCAACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTAGACAATGGGCGCAAGC
CTGATCTAGCCATGCCGCGTGAGTGATGAAGGCCTTAGGGTCGTAAAGCTCTTTCGCCTGTGAAGTGACTGTAGCAGG
TAAAGAAACCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGTTAGCGTTGTTTCGGAATTACTGGGCG
TAAAGCGCGCGTAGGCGGACTTAAAGTCAGGGGTGAAATCCCCGGGCTCAACCCCGGAATGCCTTTGATACTGGGT
GTCTTGAGTTTCAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGAT

>P3E08
CGGCCCGCGTTAGATTAGTAGTTGGTGGGGTAATGGCCTACCAAGTCTACGATCTATAGCTGGTTTTAGAGGATGAT
CAGCAACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTGACAATGGGCGCAAGC
CTGATCCAGCTATGCCGCGTGAGTGATGAAGGCCTTAGGGTCGTAAAGCTCTTTCGCCTGTGATGTGACAGTAGCAGG
TAAAGAAACCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGTTAGCGTTGTTTCGGAATTACTGGGCG
TAAAGCGCGCGTAGGCGGACTAGTCAGTCAGGGGTGAAATCCCCGGGCTCAACCCCGGAATGCCTTTGATACTGCTA
GTCTAGAGTTTCAAGAGAGTGAGTGGAATTCCGAGTGTAGAGGTGACATTCGTAGAT

>p3e9
GAGCCCGCGTAGGATTAGCTTGTTGGTGAGGTAAAGGCTCACCAAGGCGACGATCCTTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGCGAAAGC
CTGATCCAGCAATGCCGCGTGAGTGATGAAGGCCTTAGGGTGTAAAGCTCTTTTACCCGGGATGTGACAGTACCGGG
AGAATAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCG
TAAAGCGCACGTAGGCGGCTTTGTAAAGTTAGAGGTGAAAGCCAGGGCTCAACCTGGAATTGCCTTTAAGACTGCAT
CGCTCGAATCATGGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGAT

>p3g4
GAGCCCGCGTAGGATTAGCTTGTTGGTGAGGTAAAGGCTCACCAAGGCGACGATCCTTAGCTGGTCTGAGAGGATGAT
CAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGCGAAAGC
CTGATCCAGCAATGCCGCGTGAGTGATGAAGGCCTTAGGGTGTAAAGCTCTTTTACCCGGGATGTGACAGTACCGGG

AGAATAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGCGTTGTTTCGGAATTACTGGGCG
TAAAGCGCACGTAGGCGGCTTTGTAAGTTAGAGGTGAAAGCCCAGGGCTCAACCCTGGAATTGCCTTTAAGACTGCAT
CGCTCGAATCATGGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGAT

ANEXO C - SÍNTESE DE SONDAS FLUOROGÊNICAS PARA DETECÇÃO DE MONOXIGENASES E ÁLCOOL DESIDROGENASES

Sonda para detecção de Baeyer-Villiger monoxigenases BV2

As sondas foram sintetizadas segundo a metodologia de Renaud Sicard onde a primeira etapa consiste na alquilação das clorocetonas **8** e **9** com umbeliferona, gerando as sondas **10** e **11**. Em seguida os respectivos controles positivos foram sintetizados pela oxidação na presença de *m*-CPBA levando às lactonas **12** e **13** (Figura 28).

Figura 28. Síntese das sondas **10** e **11** e dos controles positivos **12** e **13**.

Composto **10**: IE/EM: $m/z = 244$ $[M]^+$. RMN de ^1H (300,01 MHz, CDCl_3 , δ_{TMS} 0,00) δ : 7,54 (*d*, $J=9,5$ Hz, 1H); 7,28 (*d*, $J=9,5$ Hz, 1H); 6,83 (*d*, $J=9,5$ Hz, 1H); 6,84 (*s*, 1H); 6,18 (*d*, $J=9,5$ Hz, 1H); 4,6 (*m*, 1H); 2,45 (*m*, 1H); 2,3 (*m*, 2H); 2,1 (*m*, 1H); 1,9 (*m*, 2H); RMN de ^{13}C (75,5 MHz, CDCl_3 , δ_{CDCl_3} 77,0) δ : 216; 160; 161; 158; 142; 129; 115; 116; 117; 102; 79; 35; 29; 17.

Composto **11**: IE/EM: $m/z = 258$ $[M]^+$. RMN de ^1H (300,01 MHz, CDCl_3 , δ_{TMS} 0,00) δ : 7,54 (*d*, $J=9,5$ Hz, 1H); 7,28 (*d*, $J=9,5$ Hz, 1H); 6,83 (*d*, $J=9,5$ Hz, 1H); 6,84 (*s*, 1H); 6,18 (*d*, $J=9,5$ Hz, 1H); 4,6 (*m*, 1H); 2,45 (*m*, 1H); 2,3 (*m*, 2H); 2,1 (*m*, 3H); 1,9 (*m*, 2H); RMN de ^{13}C (75,5 MHz, CDCl_3 , δ_{CDCl_3} 77,0) δ : 206; 161; 160; 155; 143; 128; 113,4; 113,2; 113,1; 102; 80; 40,6; 34; 27; 23.

Composto **12**: RMN de ^1H (300,01 MHz, CDCl_3 , δ_{TMS} 0,00) δ : 7,6 (*d*, $J=9,6$ Hz, 1H); 7,5 (*d*, $J=9,6$ Hz, 1H); 7,0 (*sl*, $J=9,6$ Hz, 1H); 6,9 (*d*, $J=9,6$ Hz, 1H); 6,3 (*d*, $J=9,6$ Hz, 1H); 6,15 (*t*, $J=3,3$ Hz, 1H); 2,7 (*m*, 2H); 2,3 (*m*, 1H); 2,15 (*m*, 2H); 1,95 (*m*, 1H); RMN de ^{13}C (75,5 MHz, CDCl_3 , δ_{CDCl_3} 77,0) δ : 168,8; 160,7; 158,8; 155,4; 143,0; 129,0; 114,4; 114,3; 113,3; 104,5; 98,6; 29,8; 27,3; 15,2.

Composto **13**: RMN de ^1H (300,01 MHz, CDCl_3 , δ_{TMS} 0,00) δ : 7,6 (*d*, $J=9,6$ Hz, 1H); 7,5 (*d*, $J=9,6$ Hz, 1H); 7,0 (*sl*, $J=9,6$ Hz, 1H); 6,9 (*d*, $J=9,6$ Hz, 1H); 6,3 (*d*, $J=9,6$ Hz, 1H); 6,15 (*t*, $J=3,3$ Hz, 1H); 2,7 (*m*, 2H); 2,35 (*m*, 1H); 2,25 (*m*, 2H); 1,9 (*m*, 2H); 1,75 (*m*, 1H); RMN de ^{13}C (75,5 MHz, CDCl_3 , δ_{CDCl_3} 77,0) δ : 172,7; 160,6; 158,5; 155,4; 142,9; 129,1; 114,6; 114,4; 113,1; 104,9; 98,0; 36,6; 33,6; 22,9; 22,7.

Sonda para detecção de epóxido-hidrolases, esterases e lipases

Os ensaios foram realizados utilizando cinco sondas fluorogênicas, as quais foram sintetizadas pela Dra. Luciana de Oliveira e pela Dra. Lu Shi Chen. Para detecção de epóxido-hidrolase foram utilizadas duas sondas, a 7-(1,2-epóxibutóxi)-2H-1-benzopiran-2-ona (20) e 7-(3,4-epóxi-hexilóxi)-2H-1-benzopiran-2-ona (21).

Para a detecção de esterases também foram utilizadas duas sondas, sendo 7-(1,2-di-acetato-butóxi)-2H-1-benzopiran-2-ona (22) para detecção de hidrólise acetatos, e 7-(1,2-di-propionato-butóxi)-2H-1-benzopiran-2-ona (23) para detecção de hidrólise de propionatos. A sonda 7-(1,2-di-octanoato-butóxi)-2H-1-benzopiran-2-ona (24), que contém ésteres de ácido de cadeia longa (C) foi utilizada para detecção de lipases, e além disso, foram sintetizados os controles positivos 7-(1,2-di-hidróxi-butóxi)-2H-1-benzopiran-2-ona (25) e 7-(3,4-di-hidróxi-hexilóxi)-2H-1-benzopiran-2-ona (26) (Figura 29).

Figura 29. Substratos fluorogênicos, e controles positivos para epóxido-hidrolases, esterases e lipases.

ANEXO D - RELAÇÃO DE SEQUÊNCIAS AFILIADAS A ORGANISMOS NÃO-CULTIVADOS

OTUs	Clone repres.	RDP	Blast	Descrição	% similar.	Referência
1	p1a2	Bacteria, unclassified Bacteria.	AB305537	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: p763_b_7.43.	95%	The microbial diversity in hydrothermal sediments from the Yonaguni Knoll IV, the Southern Okinawa Trough.
2	p1d9	Bacteria, unclassified Bacteria.	AF507707	Uncultured soil bacterium clone S042 16S ribosomal RNA gene, partial sequence.	87%	Empirical and theoretical bacterial diversity in four Arizona soils.
3	P1E04	Bacteria, unclassified Bacteria.	EF648019	Uncultured bacterium clone HB5 16S ribosomal RNA gene, partial sequence.	92%	Bacterial community dynamics of aerobic activated sludge during different produced water treatment.
4	P3D12	Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	EF667596	Uncultured bacterium clone LaC15L61 16S ribosomal RNA, partial sequence.	95%	Identification of novel PCE-respiring microorganisms in anoxic river sediment by RNA-based stable isotope probing.
		Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	AY221614	Uncultured soil bacterium clone HS9-74 16S ribosomal RNA gene, partial sequence.	91%	Microbial community analysis of soils contaminated with lead, chromium and petroleum hydrocarbons.
5	p2g5	Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	AY221614	Uncultured soil bacterium clone HS9-74 16S ribosomal RNA gene, partial sequence.	93%	Microbial community analysis of soils contaminated with lead, chromium and petroleum hydrocarbons.
6	P2E03	Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	EF208655	Uncultured bacterium clone CI35cm.2.13a 16S ribosomal RNA gene, partial sequence.	97%	Spatial structure of the microbial community in sandy carbonate sediment.
		Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	AF154096	Uncultured hydrocarbon seep bacterium BPC076 16S ribosomal RNA gene, partial sequence.	97%	16S rRNA analysis of hydrocarbon seep sediment.
7	P3E11	Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	EF125419	Uncultured bacterium clone MSB-1F8 16S ribosomal RNA gene, partial sequence.	96%	16S rDNA diversity analysis of mangrove soil.
8	P1G01	Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	AB305377	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: p763_b_7.43.	94%	The microbial diversity in hydrothermal sediments from the Yonaguni Knoll IV, the Southern Okinawa Trough.
		Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	AJ535229	Uncultured delta proteobacterium partial 16S rRNA gene, clone Hyd89-13.	94%	Activity, distribution, and diversity of sulfate reducers and other bacteria in sediments above gas hydrate (Cascadia Margin, Oregon).

9	P3F07	Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	DQ394958	Uncultured Geobacter sp. clone VHS-B3-70 16S ribosomal RNA gene, partial sequence.	94%	Microbial diversity and community structural shifts of sediments within the Victoria Harbour and its vicinity.
10	p2e11	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfosarcina.	AM229470	Uncultured delta proteobacterium partial 16S rRNA gene, clone B1-32.	99%	Molecular diversity studies of bacterial communities of oil polluted microbial mat from Etang de Berre (France).
11	P3C12	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, unclassified Desulfobacteraceae.	EU047534	Uncultured sulfate-reducing bacterium clone BznS259 16S ribosomal RNA gene, partial sequence.	98%	Anaerobic degradation of benzene by a marine sulfate-reducing enrichment culture, and cell hybridization of the dominant phylotype.
		Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, unclassified Desulfobacteraceae.	AY592396	Uncultured bacterium clone Amsterdam-2B-38 16S ribosomal RNA gene, partial sequence.	98%	Comparison of microbial communities from deep-sea mud volcanoes in the eastern Mediterranean in relation to the conservation of species or conservation of function concept.
12	P3C03	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, unclassified Desulfobacteraceae.	AB177130	Uncultured bacterium gene for 16S rRNA, clone: ODP1230B1.06.	92%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
13	p1g9	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, unclassified Desulfobacteraceae.	AB177131	Uncultured bacterium gene for 16S rRNA, clone: ODP1230B1.06.	94%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
14	P1H06	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, unclassified Desulfobacteraceae.	AM176871	Uncultured bacterium partial 16S rRNA gene, clone SZB46.	95%	Recovery of novel bacterial diversity from mangrove sediment.

15	p1a3	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfobacterium.	AY542205	Uncultured delta proteobacterium clone GoM HDB-20 16S ribosomal RNA gene, partial sequence.	98%	Characterization of Microbial Community Structure in Gulf of Mexico Gas Hydrates: Comparative Analysis of DNA- and RNA-Derived Clone Libraries.
16	p2c11	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfobacterium.	EF125430	Uncultured bacterium clone MSB-2C9 16S ribosomal RNA gene, partial sequence.	98%	16S rDNA diversity analysis of mangrove soil.
17	P1G10	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfobacterium.	AY771951	Uncultured delta proteobacterium clone SI28 16S ribosomal RNA gene, partial sequence.	92%	Diversity and vertical distribution of cultured and uncultured Deltaproteobacteria in an intertidal mud flat of the Wadden Sea.
18	P3D05	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfobacterium.	EF459970	Uncultured bacterium clone 197b2 16S ribosomal RNA gene, partial sequence.	93%	Active bacterial population structure and function along vertical red-ox gradients in Baltic Sea sediment.
19	p3b5	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfobacterium.	DQ394937	Uncultured Desulfobacterium sp. clone VHS-B3-34 16S ribosomal RNA gene, partial sequence.	96%	Microbial diversity and community structural shifts of sediments within the Victoria Harbour and its vicinity.
20	P1F06	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfobacterium.	AB294892	Uncultured delta proteobacterium gene for 16S rRNA, partial sequence, clone: pltb-HW-10.	93%	Culture-dependent & culture-independent analyses of bacterial community at a shallow submarine hydrothermal system off Taketomi Island, Japan.
		Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfobacterium.	DQ811831	Uncultured delta proteobacterium clone MSB-5D12 16S ribosomal RNA gene, partial sequence.	94%	16S rRNA analysis of mangrove soil.

		Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfobacterium.	AJ237601	Desulfobacterium anilini (T); DSM 4660.		Reclassification of Desulfobacterium macestii as Desulfomicrobium macestii comb. nov.
21	P2H04	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfobacterium.	AB294892	Uncultured delta proteobacterium gene for 16S rRNA, partial sequence, clone: pltb-HW-10.	89%	Culture-dependent & culture-independent analyses of bacterial community at a shallow submarine hydrothermal system off Taketomi Island, Japan.
22	P2G08	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfobacterium.	EF393191	Uncultured bacterium clone ORSFAB_a11 16S ribosomal RNA gene, partial sequence.		Phylogenetic Characterization of Polychlorinated Biphenyl Dechlorinating Consortia Under Different Anaerobic Treatments of Ohio River Sediments.
23	p2a5	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfotignum .	EU047541	Uncultured sulfate-reducing bacterium clone BznS334 16S ribosomal RNA gene, partial sequence.	96%	Anaerobic degradation of benzene by a marine sulfate-reducing enrichment culture, and cell hybridization of the dominant phylotype.
		Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfotignum .	EF207159	Desulfotignum toluolica strain H3 clone 7 16S ribosomal RNA gene, partial sequence.	95%	Desulfotignum toluolica sp. nov., a novel toluene degrading sulphate reducing bacteria isolated from an oil reservoir model column.
24	p3f1	Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	AB177158	Uncultured bacterium gene for 16S rRNA, clone: ODP1230B17.18.	98%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
25	p1c9	Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	AB177299	Uncultured bacterium gene for 16S rRNA, clone: ODP1251B1.3	95%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
26	p1a8	Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	AB177299	Uncultured bacterium gene for 16S rRNA, clone: ODP1251B1.3	97%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.

27	p2a12	Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	AB177299	Uncultured bacterium gene for 16S rRNA, clone: ODP1251B1.3	94%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
28	p1b11	Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	AB177146	Uncultured bacterium gene for 16S rRNA, clone: ODP1230B12.01.	90%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
29	P1A10	Bacteria, Proteobacteria, class Deltaproteobacteria, unclassified Deltaproteobacteria.	DQ499320	Uncultured bacterium clone CV90 16S ribosomal RNA gene, partial sequence.	91%	Extremely acidic, pendulous cave wall biofilms from the Frasassi cave system, Italy.
30	p2c8	Bacteria, Proteobacteria, class Betaproteobacteria, order Burkholderiales, family Burkholderiaceae, genus Cupriavidus.	EF205449	Uncultured bacterium clone YCB6 16S ribosomal RNA gene, partial sequence.	99%	Comparison of bacterial diversity among geographically separated geothermal regions in central Tibet.
		Bacteria, Proteobacteria, class Betaproteobacteria, order Burkholderiales, family Burkholderiaceae, genus Cupriavidus.	DQ677012	Ralstonia sp. Iso-A8 16S ribosomal RNA gene, partial sequence.	99%	Phylogenetic and physiological diversity of dissimilatory ferric iron reducers in sediments of the polluted Scheldt estuary, Northwest Europe.
		Bacteria, Proteobacteria, class Betaproteobacteria, order Burkholderiales, family Burkholderiaceae, genus Cupriavidus.	CP000352	Cupriavidus metallidurans CH34.	99%	Complete sequence of the chromosome of Ralstonia metallidurans CH34
31	p2a3	Bacteria, Proteobacteria, class Betaproteobacteria, order Burkholderiales, family Comamonadaceae, genus Delftia.	EF421404	Delftia tsuruhatensis strain A90 16S ribosomal RNA gene, partial sequence.	96%	Molecular Characterization of Class 3 Integrins from Delftia spp
32	P3B12	Bacteria, Proteobacteria, class Gammaproteobacteria, order Alteromonadales, family Alteromonadaceae, genus Alteromonas.	EF587979	Alteromonas sp. UST061013-038 16S ribosomal RNA gene, partial sequence.	97%	Phylogenetic relationship and acyl-homoserine lactones based quorum sensing in bacteria isolated from natural subtidal biofilm.

33	p2h3	Bacteria, Proteobacteria, class Gammaproteobacteria, order Aeromonadales, family Aeromonadaceae , genus Aeromonas.	EU100407	Aeromonas sp. CAX 16S ribosomal RNA gene, partial sequence.	99%	
		Bacteria, Proteobacteria, class Gammaproteobacteria, order Aeromonadales, family Aeromonadaceae , genus Aeromonas.	EU082831	Aeromonas punctata strain D4 16S ribosomal RNA gene, partial sequence.	99%	Degradation of linear alkylbenzene sulfonate by a bacterial consortium isolated from the aquatic environment of Argentina.
		Bacteria, Proteobacteria, class Gammaproteobacteria, order Aeromonadales, family Aeromonadaceae , genus Aeromonas.	AY987737	Aeromonas hydrophila strain AN-3 16S ribosomal RNA gene, complete sequence.	99%	A multigene approach to Aeromonas taxonomy.
34	p1h3	Bacteria, Proteobacteria, class Gammaproteobacteria, order Aeromonadales, family Aeromonadaceae , genus Aeromonas.	EF111197	Uncultured gamma proteobacterium clone RBE2CI-160 16S ribosomal RNA gene, partial sequence.	93%	Bacterial Community Analysis of the Bogota River.
		Bacteria, Proteobacteria, class Gammaproteobacteria, order Aeromonadales, family Aeromonadaceae , genus Aeromonas.	EF631963	Aeromonas veronii bv. sobria strain AE33 16S ribosomal RNA gene, partial sequence.	91%	Characterization of Quinolone Resistance Determinants in Aeromonas spp. isolated from South African Aquaculture Systems.
35	p2g3	Bacteria, Proteobacteria, class Alphaproteobacteria, order Rhodobacterales, family Rhodobacteraceae , genus Roseobacter.	DQ659414	Roseobacter sp. BS90 16S ribosomal RNA gene, partial sequence.	98%	Isolation and characterization of marine photoheterotrophic bacteria belonging to Roseobacter cluster.
36	P3E08	Bacteria, Proteobacteria, class Alphaproteobacteria, order Rhodobacterales, family Rhodobacteraceae , unclassified Rhodobacteraceae.	EF588012	Rhodobacteraceae bacterium UST061013-083 16S ribosomal RNA gene, partial sequence.	97%	Phylogenetic relationship and acyl-homoserine lactones based quorum sensing in bacteria isolated from natural subtidal biofilm.

	Bacteria, Proteobacteria, class Alphaproteobacteria, order Rhodobacterales, family Rhodobacteraceae , unclassified Rhodobacteraceae.	AY345413	Bacterium K2-53B 16S ribosomal RNA gene, partial sequence.	98%	Microbial Communities in the Hawaiian Archipelago: A Microbial Diversity Hotspot.
37 p3e9	Bacteria, Proteobacteria, class Alphaproteobacteria, order Sphingomonadales, family Sphingomonadaceae , unclassified Sphingomonadaceae.	EU133621	Uncultured bacterium clone FFCH7770 16S ribosomal RNA gene, partial sequence.	97%	Phylogenetic diversity and evolutionary relationships between rare and abundant members of the bacterial community in tallgrass prairie soil.
	Bacteria, Proteobacteria, class Alphaproteobacteria, order Sphingomonadales, family Sphingomonadaceae , genus Sphingomonas.	AM411908	Kaistobacter sp. P-55 partial 16S rRNA gene, strain P-55.	97%	Phylogenetic identity, growth-response time and rRNA operon copy number of soil bacteria indicate different stages of community succession.
38 p2a7	Bacteria; Proteobacteria; Deltaproteobacteria; environmental samples.	AM229478	Uncultured delta proteobacterium partial 16S rRNA gene, clone B1-50.	95%	Molecular diversity studies of bacterial communities of oil polluted microbial mat from Etang de Berre (France).
	Bacteria; environmental samples.	EF686977	Uncultured bacterium clone ATB-KS-1503 16S ribosomal RNA gene gene, partial sequence.	90%	Microbial diversity in a biogas-producing co-fermentation of maize silage and bovine manure.
39 p3e5	Bacteria, unclassified Bacteria.	AJ306780	Uncultured bacterium partial 16S rRNA gene, clone SHA-100.	0,91	Mikrobielle Diversitaet und Dynamik einer 1,2-Dichlorpropan dechlorierend en Mischkultur.
	Bacteria, unclassified Bacteria.	AB013271	Unidentified bacterium gene for 16S rRNA, partial sequence, clone: NKB19.	0,9	Microbial diversity in Nankai Trough sediments at a depth of 3,843 m.
40 p1g3	Bacteria, Firmicutes, class Clostridia, order Clostridiales, family Syntrophomonadaceae, unclassified Syntrophomonadaceae.	AB305551	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: p816_b_1.13.	91%	The microbial diversity in hydrothermal sediments from the Yonaguni Knoll IV, the Southern Okinawa Trough.
	uncultured candidate division WS1 bacterium	DQ676417	Uncultured candidate division WS1 bacterium clone MVS-77 16S ribosomal RNA gene, partial sequence.	85%	Archaeal and bacterial community composition of sediment and plankton from a suboxic freshwater pond.
	Bacteria, Firmicutes, class Clostridia, order Clostridiales, family Syntrophomonadaceae, genus Thermaerobacter.	AY094621	Thermaerobacter sp. C4-1 16S ribosomal RNA gene, partial sequence.	83%	Aerothermobacter sp. nov. C4-1 gene for 16S rRNA

41	p2h1	Bacteria, unclassified Bacteria.	AB177253	Uncultured bacterium gene for 16S rRNA, clone: ODP1244B5.24.	97%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
42	p1e10	Bacteria, unclassified Bacteria.	AY711271	Uncultured bacterium clone SIMO-1905 16S ribosomal RNA gene, partial sequence.	99%	Diversity of salt marsh prokaryotes.
		Bacteria, unclassified Bacteria.	AY592662	Uncultured bacterium clone Napoli-2B-40 16S ribosomal RNA gene, partial sequence.	96%	Comparison of microbial communities from deep-sea mud volcanoes in the eastern Mediterranean in relation to the conservation of species or conservation of function concept.
43	p1h9	Bacteria, unclassified Bacteria.	AB177305	Uncultured bacterium gene for 16S rRNA, clone: ODP1251B1.9.	92%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
		Bacteria, unclassified Bacteria.	EF999395	uncultured bacterium clone BotBa31 16S ribosomal RNA gene, partial sequence.	91%	Vertical Distribution and Diversity of Bacteria and Methanotrophs in Pearl River Estuarine Sediments.
44	p2f10	Bacteria, unclassified Bacteria.	AB177247	Uncultured bacterium gene for 16S rRNA, clone: ODP1244B3.4.	95%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
		Bacteria, unclassified Bacteria.	EF999395	uncultured bacterium clone BotBa31 16S ribosomal RNA gene, partial sequence.	91%	Vertical Distribution and Diversity of Bacteria and Methanotrophs in Pearl River Estuarine Sediments.
45	p2b8	Bacteria, unclassified Bacteria.	EF999395	uncultured bacterium clone BotBa31 16S ribosomal RNA gene, partial sequence.	91%	Vertical Distribution and Diversity of Bacteria and Methanotrophs in Pearl River Estuarine Sediments.
46	P3H06	Bacteria, unclassified Bacteria.	AB297400	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: Me0P-20.	88%	Molecular detection of halophilic denitrifying bacteria for saline wastewater treatment system.
		Bacteria, unclassified Bacteria.	DQ463707	Uncultured Chloroflexi bacterium clone TK-SH11 16S ribosomal RNA gene, partial sequence.	89%	Lake Tanganyika characterized by a distinct bacterial community.
47	P1B05	Bacteria, Acidobacteria, class Acidobacteria, order Acidobacteriales, family Acidobacteriaceae, genus Gp22	AY375053	Uncultured bacterium clone B141 16S ribosomal RNA gene, partial sequence.	87%	Bacterial Community in sediment from the western Pacific Warm Pool and its relationship to environment
48	p3c8	Bacteria, unclassified Bacteria.	AY592588	Uncultured bacterium clone Napoli-1B-33 16S ribosomal RNA gene, partial sequence.	88%	Comparison of microbial communities from deep-sea mud volcanoes in the eastern Mediterranean in relation to the conservation of species or conservation of function concept.
49	p1a1	Bacteria, unclassified Bacteria.	AY711125	Uncultured low G+C Gram-positive bacterium clone SIMO-1759 16S ribosomal RNA gene, partial sequence.	93%	Diversity of salt marsh prokaryotes.

50	p1d1	Bacteria; Firmicutes; environmental samples.	AY711125	Uncultured low G+C Gram-positive bacterium clone SIMO-1759 16S ribosomal RNA gene, partial sequence.	92%	Diversity of salt marsh prokaryotes.
		Bacteria, unclassified Bacteria.	AB177243	Uncultured bacterium gene for 16S rRNA, clone: ODP1244B22.13.	89%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
52	p3g9	Bacteria, unclassified Bacteria.	AB177152	Uncultured bacterium gene for 16S rRNA, clone: ODP1230B14.02.	94%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
53	P3B11	Bacteria, unclassified Bacteria.	AF351232	Uncultured actinobacterium clone 4-42 16S ribosomal RNA gene, partial sequence.	91%	Diversity of 16S rDNA and naphthalene dioxygenase genes from coal-tar-waste-contaminated aquifer waters.
54	P2H05	Bacteria, unclassified Bacteria.	AJ853543	Uncultured bacterium partial 16S rRNA gene, clone GZKB49.	96%	Molecular phylogenetic diversity of bacteria associated with the leachate of a closed municipal solid waste landfill.
55	P3C10	Bacteria, Actinobacteria, class Actinobacteria, unclassified Actinobacteria.	AF361216	Uncultured actinomycete clone SUBT-5 16S ribosomal RNA gene, partial sequence.	91%	Phylogenetic diversity analysis of subterranean hot springs in Iceland.
56	P2C05	Bacteria, Deferribacteres, class Deferribacteres, order Deferribacterales, unclassified Deferribacterales.	DQ394944	Uncultured bacterium clone VHS-B3-50 16S ribosomal RNA gene, partial sequence.	92%	Microbial diversity and community structural shifts of sediments within the Victoria Harbour and its vicinity.
57	P1B06	Bacteria, unclassified Bacteria.	AY592659	Uncultured bacterium clone Napoli-2B-37 16S ribosomal RNA gene, partial sequence.	95%	Comparison of microbial communities from deep-sea mud volcanoes in the eastern Mediterranean in relation to the conservation of species or conservation of function concept.
58	p2h12	Bacteria, unclassified Bacteria.	EU048666	Uncultured bacterium clone MD2896-B6 16S ribosomal RNA gene, partial sequence.	87%	Bacterial and archaeal diversity in surface sediment of the South Slope, South China Sea.
59	P3A12	Bacteria, unclassified Bacteria.	AB175571	Uncultured bacterium gene for 16S rRNA, partial sequence, clone:IBC2-7.	89%	Variability in microbial community and venting chemistry in a sediment-hosted backarc hydrothermal system: Impacts of subseafloor phase-separation.
		Bacteria, Deferribacteres, class Deferribacteres, order Deferribacterales, unclassified Deferribacterales.	DQ811936	Uncultured Deferribacteres bacterium clone MSB-5A6 16S ribosomal RNA gene, partial sequence.	88%	16S rRNA analysis of mangrove soil.
60	p1b8	Bacteria, unclassified Bacteria.	AF275926	Uncultured eubacterium AB16 16S ribosomal RNA gene, partial sequence.	96%	Monitoring of activity dynamics of an anaerobic digester bacterial community using 16S rRNA polymerase chain reaction--single-strand conformation polymorphism analysis.

61	p1g2	Bacteria, Firmicutes, class Clostridia, order Clostridiales, family Syntrophomonadaceae, genus Aminobacterium.	AB175355	Uncultured bacterium gene for 16S rRNA, partial sequence, clone:BSA1B-01.	95%	Microbial community analysis of mesophilic anaerobic protein degradation process using bovine serum albumin (BSA)-fed continuous cultivation.
		Bacteria, Firmicutes, class Clostridia, order Clostridiales, family Syntrophomonadaceae, genus Aminobacterium.	AF073521	Aminobacterium mobile (T); ILE-3.	91%	Aminobacterium mobilis sp. nov., a new anaerobic amino-acid-degrading bacterium.
		Bacteria, Firmicutes, class Clostridia, order Clostridiales, family Syntrophomonadaceae, genus Aminobacterium.	AF069287	Aminobacterium colombiense (T); ALA-1.	89%	Isolation of a novel amino acid-degrading obligate anaerobe isolated from anaerobic sludge.
62	p3b8	Bacteria, Firmicutes, class Clostridia, order Clostridiales, family Syntrophomonadaceae, genus Aminobacterium.	DQ478749	Uncultured bacterium clone CLONG96 16S ribosomal RNA gene, partial sequence.	96%	Phenotypic properties and microbial diversity of methanogenic granules from a full-scale upflow anaerobic sludge bed reactor treating brewery wastewater.
63	p1c11	Bacteria, Firmicutes, class Clostridia, unclassified Clostridia.	DQ806167	Uncultured bacterium clone RL186_aah62b01 16S ribosomal RNA gene, partial sequence.	89%	Microbial ecology: human gut microbes associated with obesity
64	p2b1	Bacteria, unclassified Bacteria.	AB177247	Uncultured bacterium gene for 16S rRNA, clone: ODP1244B3.4.	89%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
65	P1D04	Bacteria, Firmicutes, class Clostridia, unclassified Clostridia.	AB069798	Uncultured bacterium 16S rRNA, partial sequence, clone:CS5.21.	94%	Molecular phylogenetic analyses of reverse-transcribed bacterial rRNA obtained from deep-sea cold seep sediments.
		Bacteria, Firmicutes, class Clostridia, order Clostridiales, unclassified Clostridiales.	AM176888	Uncultured bacterium partial 16S rRNA gene, clone SZB67.	93%	Recovery of novel bacterial diversity from mangrove sediment.
66	p1h11	Bacteria, Firmicutes, class Clostridia, order Clostridiales, family Acidaminococcaceae, genus Papillibacter.	AF167711	Papillibacter cinnaminovorans 16S ribosomal RNA gene, partial sequence.	96%	Papillibacter cinnaminovorans gen. nov., sp. nov., a cinnamate-transforming bacterium from a shea cake digester.
67	p3e2	Bacteria, unclassified Bacteria.	AB177311	Uncultured bacterium gene for 16S rRNA, clone: ODP1251B11.22.	98%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
68	P1A05	Bacteria, unclassified Bacteria.	AB094905	Uncultured bacterium gene for 16S rRNA, partial sequence, clone:OHKB7.59.	94%	Microbial communities associated with geological horizons in coastal subseafloor sediments from the sea of okhotsk.

69	P1A06	Bacteria, unclassified Bacteria.	AB177189	Uncultured bacterium gene for 16S rRNA, clone: ODP1230B3.02.	96%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
70	P3A08	Bacteria, unclassified Bacteria.	AY592746	Uncultured bacterium clone Napoli-4B-06 16S ribosomal RNA gene, partial sequence.	95%	Comparison of microbial communities from deep-sea mud volcanoes in the eastern Mediterranean in relation to the conservation of species or conservation of function concept.
71	p2b12	Bacteria, unclassified Bacteria.	AY356377	Uncultured eubacterium N3-40 16S ribosomal RNA gene, partial sequence.	92%	Microbial Reductive Dechlorination of Organochlorine Pollutants in the Marine Environment.
72	p1e1	Bacteria, Chloroflexi, class Anaerolineae, subclass Caldilineae, order Caldilineales, family Caldilineaceae, unclassified Caldilineaceae.	AB300122	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: KM22B-27.	96%	Wide distribution of ancient microbial DNAs in the Quaternary mud sediments.
73	P2E10	Bacteria, unclassified Bacteria.	AY584740	Uncultured bacterium clone ER-A4-2 16S ribosomal RNA gene, partial sequence.	93%	A PCR-based specific assay reveals a population of bacteria within the Chloroflexi associated with the reductive dehalogenation of polychlorinated biphenyls.
74	p2b11	Bacteria, unclassified Bacteria.	AY356373	Uncultured eubacterium N5-8 16S ribosomal RNA gene, partial sequence.	91%	Microbial Reductive Dechlorination of Organochlorine Pollutants in the Marine Environment.
75	p2d2	Bacteria, Chloroflexi, class Anaerolineae, subclass Caldilineae, order Caldilineales, family Caldilineaceae, unclassified Caldilineaceae.	AB300118	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: KM22B-101.	94%	Wide distribution of ancient microbial DNAs in the Quaternary mud sediments.
76	P1C04	Bacteria, Chloroflexi, class Anaerolineae, subclass Caldilineae, order Caldilineales, family Caldilineaceae, unclassified Caldilineaceae.	AB300119	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: KM35B-30.	94%	Wide distribution of ancient microbial DNAs in the Quaternary mud sediments.
77	p1f1	Bacteria, Chloroflexi, class Anaerolineae, subclass Caldilineae, order Caldilineales, family Caldilineaceae, unclassified Caldilineaceae.	AY869675	Uncultured bacterium clone FS117-3B-02 16S ribosomal RNA gene, partial sequence.	91%	Microbial life in ridge flank crustal fluids.

78	P2E12	Bacteria, unclassified Bacteria.	AB177059	Uncultured bacterium gene for 16S rRNA, clone: ODP1227B19.11.	91%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
79	p1c8	Bacteria, unclassified Bacteria.	AY592653	Uncultured bacterium clone Napoli-2B-29 16S ribosomal RNA gene, partial sequence.	92%	Comparison of microbial communities from deep-sea mud volcanoes in the eastern Mediterranean in relation to the conservation of species or conservation of function concept.
80	p1b9	Bacteria, unclassified Bacteria.	EF667658	Uncultured bacterium clone LaP15L27 16S ribosomal RNA, partial sequence.	91%	Identification of novel PCE-respiring microorganisms in anoxic river sediment by RNA-based stable isotope probing.
		Bacteria, unclassified Bacteria.	AY093480	Uncultured bacterium clone MB-C2-127 16S ribosomal RNA, partial sequence.	90%	Microbial Communities from Methane Hydrate-Bearing Deep Marine Sediments in a Forearc Basin.
81	P1D06	Bacteria, unclassified Bacteria.	AY592633	Uncultured bacterium clone Napoli-2B-06 16S ribosomal RNA gene, partial sequence.	97%	Comparison of microbial communities from deep-sea mud volcanoes in the eastern Mediterranean in relation to the conservation of species or conservation of function concept.
82	P1G05	Bacteria, unclassified Bacteria.	EF208708	Uncultured bacterium clone CI75cm.2.15 16S ribosomal RNA gene, partial sequence.	97%	Spatial structure of the microbial community in sandy carbonate sediment.
83	P1G06	Bacteria, unclassified Bacteria.	AB300104	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: KM35B-155.	95%	Wide distribution of ancient microbial DNAs in the Quaternary mud sediments.
84	P3B04	Bacteria, unclassified Bacteria.	AY592653	Uncultured bacterium clone Napoli-2B-29 16S ribosomal RNA gene, partial sequence.	97%	Comparison of microbial communities from deep-sea mud volcanoes in the eastern Mediterranean in relation to the conservation of species or conservation of function concept.
85	p1g4	Bacteria, unclassified Bacteria.	AY592654	Uncultured bacterium clone Napoli-2B-29 16S ribosomal RNA gene, partial sequence.	97%	Comparison of microbial communities from deep-sea mud volcanoes in the eastern Mediterranean in relation to the conservation of species or conservation of function concept.
86	P3G03	Bacteria, unclassified Bacteria.	DQ463707	Uncultured Chloroflexi bacterium clone TK-SH11 16S ribosomal RNA gene, partial sequence.		Lake Tanganyika characterized by a distinct bacterial community.
87	P3D08	Bacteria, unclassified Bacteria.	AB177058	Uncultured bacterium gene for 16S rRNA, clone: ODP1227B19.08.		Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
88	p1a9	Bacteria, unclassified Bacteria.	AF407200	Uncultured bacterium clone GIF9 16S ribosomal RNA gene, partial sequence.	97%	Microbial diversity in an in situ reactor system treating monochlorobenzene contaminated groundwater as revealed by 16S ribosomal DNA analysis.

	Bacteria, unclassified Bacteria.	DQ811891	Uncultured Chloroflexi bacterium clone MSB-4E3 16S ribosomal RNA gene, partial sequence.	93%	16S rRNA analysis of mangrove soil.
89	p2d1	AF524012	Uncultured bacterium clone FW64 16S ribosomal RNA gene, partial sequence.	89%	Recovery of novel bacterial diversity from a forested wetland impacted by reject coal.
	Bacteria, unclassified Bacteria.	DQ463707	Uncultured Chloroflexi bacterium clone TK-SH11 16S ribosomal RNA gene, partial sequence.	88%	Lake Tanganyika characterized by a distinct bacterial community.
90	p3d2	DQ463707	Uncultured Chloroflexi bacterium clone TK-SH11 16S ribosomal RNA gene, partial sequence.	88%	Lake Tanganyika characterized by a distinct bacterial community.
91	P1C06	EU181503	Uncultured bacterium clone 8BAVF8 16S ribosomal RNA gene, complete sequence.	94%	Prokaryotic diversity associated with enhanced organic matter remineralization in continental margin sediments.
	Bacteria, unclassified Bacteria.	AY704394	Uncultured bacterium clone FS118-62B-02 16S ribosomal RNA gene, partial sequence.	91%	Microbial life in ridge flank crustal fluids.
92	P2C10	EU181503	Uncultured bacterium clone 8BAVF8 16S ribosomal RNA gene, complete sequence.		Prokaryotic diversity associated with enhanced organic matter remineralization in continental margin sediments.
	Bacteria, unclassified Bacteria.	AY704394	Uncultured bacterium clone FS118-62B-02 16S ribosomal RNA gene, partial sequence.		Microbial life in ridge flank crustal fluids.
93	P3C07	AY704394	Uncultured bacterium clone FS118-62B-02 16S ribosomal RNA gene, partial sequence.	96%	Microbial life in ridge flank crustal fluids.
94	p2g1	AY704394	Uncultured bacterium clone FS118-62B-02 16S ribosomal RNA gene, partial sequence.	89%	Microbial life in ridge flank crustal fluids.
95	p1f4	AB177238	Uncultured bacterium gene for 16S rRNA, clone: ODP1244B11.11	95%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
96	p1h12	AB177238	Uncultured bacterium gene for 16S rRNA, clone: ODP1244B11.11		Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
97	p1h1	AB177253	Uncultured bacterium gene for 16S rRNA, clone: ODP1244B5.24.	90%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
98	P3D10	DQ811888	Uncultured Chloroflexi bacterium clone MSB-5B10 16S ribosomal RNA gene, partial sequence.	89%	
99	p2c2	AY869681	Uncultured bacterium clone FS118-32B-02 16S ribosomal RNA gene, partial sequence.	94%	Microbial life in ridge flank crustal fluids.
100	P2F08	AY704394	Uncultured bacterium clone FS118-62B-02 16S ribosomal RNA gene, partial sequence.	89%	Microbial life in ridge flank crustal fluids.

101	p2f11	Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfobacterium.	DQ394937	Uncultured Desulfobacterium sp. clone VHS-B3-34 16S ribosomal RNA gene, partial sequence.	86%	Microbial diversity and community structural shifts of sediments within the Victoria Harbour and its vicinity.
		Bacteria, Proteobacteria, class Deltaproteobacteria, order Desulfobacterales, family Desulfobacteraceae, genus Desulfobacterium.	DQ811831	Uncultured delta proteobacterium clone MSB-5D12 16S ribosomal RNA gene, partial sequence.	85%	16S rRNA analysis of mangrove soil.
102	P1D05	Bacteria, Chloroflexi, class Anaerolineae, subclass Caldilineae, order Caldilineales, family Caldilineaceae, unclassified Caldilineaceae.	AY627576	Uncultured bacterium clone Urania-2B-17 16S ribosomal RNA gene, partial sequence.	90%	Comparison of deep-sea microbial communities in the eastern Mediterranean.
103	p1h4	Bacteria, unclassified Bacteria.	AB177241	Uncultured bacterium gene for 16S rRNA, clone: ODP1244B18.9	87%	Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin.
104	P3H03	Bacteria, Chloroflexi, class Anaerolineae, subclass Caldilineae, order Caldilineales, family Caldilineaceae, unclassified Caldilineaceae.	DQ811875	Uncultured Chloroflexi bacterium clone MSB-4D11 16S ribosomal RNA gene, partial sequence.		16S rRNA analysis of mangrove soil.
105	p2e1	Bacteria, Chloroflexi, class Anaerolineae, subclass Caldilineae, order Caldilineales, family Caldilineaceae, genus Caldilinea.	EF464632	Uncultured Chloroflexi bacterium clone ELC_30_7_bac 16S ribosomal RNA gene, partial sequence.	96%	Identification of sulfate reducing bacteria in methylmercury contaminated mine tailings by analysis of SSU ribosomal RNA genes.
106	P2B04	Bacteria; Firmicutes; class Clostridia, unclassified Clostridia.	AF323773	Uncultured bacterium clone BA087 16S ribosomal RNA gene, complete sequence.	95%	Molecular diversity of bacteria members associated with benzoate degradation in the methanogenic consortium.
107	p1h8	Bacteria; Firmicutes; unclassified Firmicutes.	AF323760	Uncultured bacterium clone BA021 16S ribosomal RNA gene, complete sequence.	94%	Molecular diversity of bacteria members associated with benzoate degradation in the methanogenic consortium.
108	P2D07	Bacteria; Firmicutes; class Clostridia, unclassified Clostridia.	EF559145	Uncultured bacterium clone G35_D8_L_B_G08 16S ribosomal RNA gene, partial sequence.		Microbial functional groups in a mesophilic anaerobic solid waste digester.

109	p1f12	Bacteria; Firmicutes; class Clostridia, unclassified Clostridia.	AY592365	Uncultured bacterium clone Amsterdam-2B-05 16S ribosomal RNA gene, partial sequence.	98%	Comparison of microbial communities from deep-sea mud volcanoes in the eastern Mediterranean in relation to the conservation of species or conservation of function concept.
110	p3h11	Bacteria; Firmicutes; unclassified Firmicutes.	AF323760	Uncultured bacterium clone BA021 16S ribosomal RNA gene, complete sequence.	95%	Molecular diversity of bacteria members associated with benzoate degradation in the methanogenic consortium.
111	p1e9	Bacteria; Firmicutes; class Mollicutes, order Acholeplasmatales, family Acholeplasmataceae, genus Acholeplasma.	DQ337046	Uncultured bacterium clone EV818EB5CPSAJJ46 16S ribosomal RNA gene, partial sequence.	95%	The distribution of microbial taxa in the subsurface water of the Kalahari Shield, South Africa.
112	p2h2	Bacteria; Firmicutes; class Mollicutes, order Acholeplasmatales, family Acholeplasmataceae, genus Acholeplasma.	AY953180	Uncultured anaerobic bacterium clone B-1M 16S ribosomal RNA gene, partial sequence.	89%	Unique Microbial Diversity of Anaerobic Swine Lagoons
		Bacteria; Firmicutes; class Mollicutes, order Acholeplasmatales, family Acholeplasmataceae, genus Acholeplasma.	S000484607	Acholeplasma pleciae (T); ATCC 49582; PS-1.		Reclassification of Mesoplasma pleciae as Acholeplasma pleciae comb. nov. on the basis of 16S rRNA and gyrB gene sequence data.
		Bacteria; Firmicutes; class Mollicutes, order Acholeplasmatales, family Acholeplasmataceae, genus Acholeplasma.	S000414577	Acholeplasma palmae (T); ATCC49389; J233.		Phylogeny of mycoplasma-like organisms (phytoplasmas): a basis for their classification.
		Bacteria; Firmicutes; class Mollicutes, order Acholeplasmatales, family Acholeplasmataceae, genus Acholeplasma.	S000427876	Acholeplasma vituli (T); FC-097.		Acholeplasma vituli sp. nov., from bovine serum and cell cultures.
113	P2A10	Bacteria; Spirochaetes; class Spirochaetes, order Spirochaetales, family Spirochaetaceae, genus Spirochaeta.	EF031095	Uncultured bacterium clone SalB63 16S ribosomal RNA gene, partial sequence.	92%	Prokaryotic community of Tirez, a endorheic hypersaline lake in La Mancha, Spain.
114	p2h10	Bacteria; Spirochaetes; class Spirochaetes, order Spirochaetales, family Spirochaetaceae, genus Spirochaeta.	AY605137	Uncultured spirochete clone IE046 16S ribosomal RNA gene, partial sequence.	96%	Phylogenetic diversity and temporal variation in the Spirochetal community in microbial mats of Ebro and Camargue Deltas.
115	P2E07	Bacteria, unclassified Bacteria.	EU134907	Uncultured bacterium clone FFCH17214 16S ribosomal RNA gene, partial sequence.	90%	Phylogenetic diversity and evolutionary relationships between rare and abundant members of the bacterial community in tallgrass prairie soil.

		Bacteria, unclassified Bacteria.	AB294345	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: SWB34.	86%	Microbial community structure in deep coal seam groundwater
116	P3E04	Bacteria, Proteobacteria, class Epsilonproteobacteria, unclassified Epsilonproteobacteria.	DQ112498	Uncultured epsilon proteobacterium clone KorMud-V6C55 16S ribosomal RNA gene, partial sequence.	97%	Bacterial community structure in the intertidal mudflatsediment vegetated by <i>Sueada japonica</i> in Ganghwa Island, Korea.
		Bacteria, Proteobacteria, class Epsilonproteobacteria, order Campylobacterales, family Helicobacteraceae, genus Sulforovun.	EF061975	Uncultured epsilon proteobacterium clone XME6 16S ribosomal RNA gene, partial sequence.	96%	The bacterial diversity in mangrove sediment of Xiamen, China
117	p1c2	Bacteria, Proteobacteria, class Epsilonproteobacteria, order Campylobacterales, family Helicobacteraceae, genus Sulforovun.	EF061976	Uncultured epsilon proteobacterium clone XME6 16S ribosomal RNA gene, partial sequence.	90%	The bacterial diversity in mangrove sediment of Xiamen, China
118	P3F10	Bacteria, Proteobacteria, class Epsilonproteobacteria, order Campylobacterales, family Campylobacteraceae, genus Arcobacter.	AB250570	Uncultured bacterium gene for 16S ribosomal RNA, partial sequence, clone: #4-3.	94%	Molecular and microbial ecology of the Grotta Azzurra of Palinuro Cape (Salerno, Italy).
119	p3g11	Bacteria, Proteobacteria, class Epsilonproteobacteria, order Campylobacterales, family Helicobacteraceae, genus Sulforovun.	EF061976	Uncultured epsilon proteobacterium clone XME6 16S ribosomal RNA gene, partial sequence.	88%	The bacterial diversity in mangrove sediment of Xiamen, China
120	p2a11	Bacteria, Proteobacteria, class Epsilonproteobacteria, order Campylobacterales, family Helicobacteraceae, genus Sulfurimonas.	AY225612	Uncultured epsilon proteobacterium clone AT-co23 16S ribosomal RNA gene, partial sequence.	94%	Bacterial diversity in hydrothermal sediment and epsilonproteobacterial dominance in experimental microcolonizers at the Mid-Atlantic Ridge.
121	p3d4	Bacteria, Proteobacteria, class Epsilonproteobacteria, order Campylobacterales, family Helicobacteraceae, genus Sulforovun.	EF061976	Uncultured epsilon proteobacterium clone XME34 16S ribosomal RNA gene, partial sequence.	89%	The bacterial diversity in mangrove sediment of Xiamen, China

122	P2B05	Bacteria, Chlorobi, class Chlorobia, order Chlorobiales, family Chlorobiaceae, genus Chlorobium.	AM690805	Chlorobium limicola partial 16S rRNA gene, isolate CNg.	99%	Distribution of green sulfur bacteria in various habitats in India
-----	-------	--	----------	---	-----	--

ANEXO E - PROTOCOLO DAPI

Método: Protocolo DAPI

Reagentes / Passos:

1. Aliquotar 5 a 10 μL de sobre nadante (SN) de enriquecimento previamente precipitado e pingar sobre lâmina de vidro escavada;
2. Deixar as amostras secarem na lâmina por 30 minutos;
3. Aplicar 50 μL de metanol absoluto em cada escavação e deixar secar por 3 a 5 minutos;
4. Quando o material estiver bem seco, adicionar 50 μL de solução DAPI de uso, em cada amostra e deixar em temperatura ambiente (TA) por 5 min. sob abrigo da luz;
5. Remover excesso de DAPI por aspiração;
6. Lavar 2x com PBS 150 mM, pH 7.2 ;
7. Lavar 2x com Água- reagente;
8. Deixar 1h em local seco, TA, sob abrigo da luz;
9. Aplicar 10 μL de meio de montagem (glicerol 60%, PBS 150 mM, 40% contendo DABCO (2%);
10. Colocar a lamínula e observar em imersão (aumento de 60 X).

ANEXO F - ARTIGO - HYDROLASE ACTIVITY IN METAGENOMIC LIBRARY FROM PETROLEUM AND HEAVY METALS IMPACTED SEDIMENTS

Simões, T.H.N.^a, Mantovani, S.M.^b, Dimitrov, M. R.^c, Roseto, C.^b, Marsaioli, A.^b, Garboggini, F.F.^a, Oliveira. V.M.^a, Menezes, C.B.A.^a, Simioni K.C.^a, Santos, E.C. Manfio, G.P.^a

^a *Microbial Resource Division – CPQBA – State University of Campinas UNICAMP
P.O. Box 6171 – CEP 13081-970 SP - Brazil.*

^b *Chemistry Institute, P.O. Box 6154, State University of Campinas – UNICAMP,
13084-862, Campinas, SP, Brazil.*

^c *Department of Technology, UNESP- Jaboticabal CEP 14884-900 Jaboticabal – SP-
Brazil.*

*Author for correspondence (e-mail: thnsimoes@yahoo.com.br, phone: +55 19 3884-7500.

Abstract

Recent advances in molecular biology using direct cloning libraries have brought new perspectives to explore the functional diversity of non-cultivated microorganisms in environmental samples, enabling the discovery of novel genes and enzymes for biotechnological applications. This study aimed at exploring the diversity of enzymes in the metagenoma of oil degrading bacterial communities in Guanabara Bay (Rio de Janeiro) sediments. Approximately 900 clones with 35 kb insert fragments were generated from petroleum and heavy metals impacted sediment enrichments. Proteobacteria and sulphate reducing bacteria (BRS) represented the main groups in the sample studied based upon 16S rDNA phylogenetic analyses. High throughput screening (HTS) were performed to detect lipases, esterases and epoxide-hydrolases.

Key Words: metagenome, sediments, HTS, hydrolases.