



UNIVERSIDADE FEDERAL RURAL DO SEMIÁRIDO
DEPARTAMENTO DE CIÊNCIAS ANIMAIS – DCA_n
BACHARELADO EM BIOTECNOLOGIA

KLÊYDSON RAMOS DE SENA

**METODOLOGIA PARA ANÁLISE METAGENÔMICA DE BACTÉRIAS DO
SEMIÁRIDO RELACIONADAS À FORMAÇÃO DE NUVENS**

MOSSORÓ-RN

2014

KLÊYDSON RAMOS DE SENA

**METODOLOGIA PARA ANÁLISE METAGENÔMICA DE BACTÉRIAS DO
SEMIÁRIDO RELACIONADAS À FORMAÇÃO DE NUVENS**

Monografia apresentada à Universidade Federal Rural do Semiárido (UFERSA), como exigência final para obtenção do título de Bacharel em Biotecnologia.

Orientadora: Prof^a. Dra. Fernanda Matias

MOSSORÓ/RN

2014

Dados Internacionais de Catalogação na Publicação (CIP)
Biblioteca Central Orlando Teixeira (BCOT)
Setor de Informação e Referência

S474m Sena, Klêydson Ramos de

Metodologia para análise metagenômica de bactérias do
semiárido relacionadas à formação de nuvens/ Klêydson Ramos
de Sena -- Mossoró, 2014.
60f.: il.

Orientadora: Prof^a. Dra. Fernanda Matias

Monografia (Graduação em Biotecnologia) – Universidade
Federal Rural do Semiárido. Pró-Reitoria de Graduação.

1. Biotecnologia. 2. DMSP. 3. DNA. 4. Análise
Metagenômica. I. Título.

RN/UFERSA/BCOT/821-14

CDD: 660.6

Bibliotecária: Vanessa Christiane Alves de Souza Borba
CRB-15/452

KLÊYDSON RAMOS DE SENA

**METODOLOGIA PARA ANÁLISE METAGENÔMICA DE BACTÉRIAS DO
SEMIÁRIDO RELACIONADAS À FORMAÇÃO DE NUVENS**

Monografia apresentada à Universidade Federal Rural do Semiárido (UFERSA), como exigência final para obtenção do título de Bacharel em Biotecnologia.

APROVADO EM 04/08/2014

BANCA EXAMINADORA

Prof^a Dra. Fernanda Matias – UFERSA
Orientadora – Presidente

Prof^a Dra. Lidianne Leal Rocha - UFERSA
Primeiro Membro

Prof^a Dra. Emanuelle Fontenele Rabelo - UFERSA
Segundo Membro

A minha mãe, **Sara Ramos de Sena**, e meu pai, **Antônio Sena de Lima**, por todo esforço para que eu conseguisse chegar até aqui e por todas as orações e conselhos que me fortaleceram a cada dia.

DEDICO

AGRADECIMENTOS

A Deus, primeiramente, por sempre me dar forças para continuar minha jornada, mesmo eu não sendo merecedor de toda a Sua misericórdia.

A minha mãe, Sara Ramos de Sena, por ter me dado à vida e por orar todas as madrugadas e me aconselhar para que eu me tornasse o homem que sou. Obrigado mãe! Amo muito a senhora!

A meu pai, Antônio Sena de Lima, por servir de exemplo e me ensinar a respeitar sempre e correr atrás dos meus objetivos.

A minha orientadora, Dra. Fernanda Matias, pela oportunidade de trabalhar ao seu lado, me ensinando e me orientando sempre. Tudo que aprendi, agradeço a ela, por ser paciente e me ajudar em todas as horas, não só nos experimentos, mas nas escolhas profissionais. Muito obrigado!

A minha namorada, Anny Jasmine, por me apoiar e me aguentar em todas as horas de estresse. Pela sua companhia no laboratório tanto nos estudos, como nas coletas e nos experimentos. Por me aconselhar a sempre seguir em frente e não desistir. Muito obrigado, meu amor. Amo você.

A meus colegas de turma (a melhor!), que me proporcionou uma nova família dentro dessa universidade, me fazendo sorrir e chorar, me alegrar e sentir raiva. Especialmente aos meus amigos do peito Jessé, Eduardo e Ramon. Muito obrigado pelas boas risadas e “mungangas”.

Aos companheiros do LABIN, por todo o apoio e ajuda nas horas que precisei. Especialmente Vinícius, que me ajudou bastante nesse fim de jornada e Iara pela ajuda nas coletas e nos experimentos.

As minhas irmãs Késia e Kesyane, e seus filhos Juan, Rennan e Yasmin, por me darem total apoio, orações e momentos de felicidade quando eu mais precisei.

A toda a minha família que sempre me apoiaram, oraram e torceram para que eu conseguisse realizar esse sonho.

Aos meus amigos pessoais, Handeildson (mais que irmão), Heloisa, Niáscara, Larissa e Patrícia Macedo, por torcerem sempre por mim para que eu chegasse onde estou.

Ao LEGPRO, por abrir as portas para que eu pudesse utilizar os equipamentos e aos estagiários, Júlio e Adriana pela paciência de me esperar nas horas da PCR.

Aos companheiros de curso, em especial Hebert e Clara pela companhia, pelos conselhos e pela torcida.

As professoras Dra. Lidianne Leal Rocha e Dra. Emanuelle Fontenele Rabelo, por aceitarem o convite para participar de minha banca.

E a todos que, direta ou indiretamente, me ajudaram a chegar até aqui e contribuíram para o término deste trabalho. Obrigado!

“Tudo tem o seu tempo determinado, e há tempo para todo o propósito debaixo do céu”.

(Eclesiastes 3.1)

RESUMO

A metagenômica é uma estratégia que nos permite o acesso ao potencial genômico de amostras ambientais, permitindo a análise de DNA ambiental sem a necessidade de cultivo dos micro-organismos. Essa abordagem pode ser utilizada para avaliar a presença dos genes *dmdA* que codificam a enzima DMSP Lias. Essa enzima catalisa o gás sulfonilpropionato de dimetila (DMSP) que será biotransformado em sulfeto de dimetila (DMS) auxiliando, assim, na produção de nuvens em determinada região. Buscando-se entender se há relação entre a diversidade microbiana e de genes que codificam a enzima DMSP liase, foram realizadas extrações de DNA metagenômico de solo em três regiões de Mossoró (RN - Brasil) sendo a primeira no bairro Aeroporto de Mossoró (local com menos vegetação – duas amostras), a segunda próxima ao Rio Apodi-Mossoró (local de maior vegetação – duas amostras) e a terceira nas proximidades da cidade de Areia Branca (região de maior salinidade – duas amostras), com o objetivo de se estabelecer um protocolo de extração do DNA das amostras e amplificação dos genes *dmdA* e 16S rDNA. Ao término dos experimentos, conseguiu-se desenvolver protocolos para extração de DNA metagenômico tanto do solo quanto da água do mar, que serviu como controle positivo em relação aos genes, e padronização da amplificação do 16S rDNA.

Palavras-chaves: Metagenômica. DMSP. Extração de DNA.

SUMÁRIO

1. INTRODUÇÃO	11
2. FUNDAMENTAÇÃO TEÓRICA	12
2.1 MUDANÇAS CLIMÁTICAS.....	12
2.2 SEMIÁRIDO NORDESTINO	13
2.3 FORMAÇÃO DE NUVENS.....	15
2.3.1 Teoria Gaia	15
2.3.2 Hipótese CLAW.....	16
2.3.3 Genes <i>Ddd</i>	18
2.4 METAGENÔMICA.....	19
2.4.1 Material Genético – Acesso Indireto X Acesso Direto	21
3. OBJETIVOS.....	23
3.1 OBJETIVO GERAL	23
3.2 OBJETIVOS ESPECÍFICOS	23
4. METODOLOGIA	24
4.1 COLETA DAS AMOSTRAS DO SOLO E DA ÁGUA DO MAR.....	24
4.2 EXTRAÇÃO DE DNA METAGENÔMICO DO SOLO	27
4.3 EXTRAÇÃO DE DNA METAGENÔMICO DA ÁGUA DO MAR	27
4.4 GEL DE AGAROSE PARA ELETROFORESE	28
4.5 AMPLIFICAÇÃO DO GENE 16S rDNA E <i>dmdA</i>	28
4.6 BIOINFORMÁTICA	29
5. RESULTADOS E DISCUSSÃO	30
5.1 COLETA DAS AMOSTRAS	30
5.2 EXTRAÇÃO DE DNA METAGENÔMICO DO SOLO	30
5.3 EXTRAÇÃO DE DNA METAGENÔMICO DA ÁGUA DO MAR	32
5.4 AMPLIFICAÇÃO DO 16S rDNA.....	33
5.5 AMPLIFICAÇÃO DO GENE <i>dmdA</i>	34
5.6 BIOINFORMÁTICA	35
5.6.1 Desenho dos iniciadores	36
6. CONSIDERAÇÕES FINAIS	37
7. PERSPECTIVAS FUTURAS.....	38
8. REFERÊNCIAS	39
9. ANEXOS	47

1. INTRODUÇÃO

O Brasil é um país de dimensões continentais em que cada microrregião possui características climáticas específicas fazendo com que o país seja detentor de uma grande diversidade biológica e genômica, sendo considerado um dos países com megabiodiversidade. Enquanto o sul do país apresenta temperaturas variadas e taxas pluviais médias, o nordeste possui temperaturas constantes além de taxas pluviais mais baixas, constituindo uma variedade vegetal diferenciada. Cada microrregião nordestina possui características climáticas e vegetativas únicas, o que pode indicar uma biodiversidade microbiana diferenciada, tema que tem sido pouco explorado nas pesquisas científicas e/ou tecnológicas. O Rio Grande do Norte por si só apresenta três microrregiões de acordo com o IBGE: Zona da Mata, Agreste e Sertão, sendo que os dois últimos compõem a caatinga. Esse último corresponde a um bioma exclusivamente brasileiro, sendo composto de vegetação típica da região semiárida. Além disso, o estado do Rio Grande do Norte, incluindo a região de Mossoró, é um dos maiores produtores de sal do mundo, o que pode afetar a constituição microbiana local. Mossoró fica na região do Sertão, clima semiárido, a qual é marcada por duas estações, uma seca, de aproximadamente nove meses, e uma úmida, de aproximadamente três meses. Essas características podem afetar a produção de DMSP (sulfonilpropionato de dimetila) que, de acordo com a hipótese CLAW, poderia, de forma indireta, influenciar na quantidade de chuvas, não apenas nessa região, mas no mundo todo. O DMSP é um composto de enxofre orgânico, produzido pelo fitoplâncton marinho como antioxidante. Esse composto é biotransformado a DMS (sulfeto de dimetila) pela enzima DMSP liase presente em bactérias e tem importante papel na formação de nuvens.

O uso de tecnologias aplicadas ao melhoramento climático, sem dúvidas, pode ser a porta para o desenvolvimento das regiões afetadas pelas constantes secas, sendo tais tecnologias nos últimos anos o foco de grandes investimentos governamentais e de grandes empresas interessadas na exploração do semiárido nordestino. Desta forma, este trabalho buscou estabelecer uma metodologia para extração de DNA metagenômico de solo e água marinha. Além disso, buscou-se padronizar a amplificação dos genes 16S rDNA e *dmdA*, sendo o último responsável pela codificação da enzima DMSP liase, das amostras de DNA metagenômico.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 Mudanças Climáticas

O tema “mudanças climáticas” é atual e tem sido constantemente enfatizado pela mídia escrita e televisiva ao longo destes últimos anos. As informações transmitidas ao público em geral têm sido nas maiorias das vezes concisas e imprecisas. Neste sentido, não somente a população, mas principalmente os tomadores de decisão, nem sempre conseguem discernir as certezas e incertezas com relação às variações do clima presente, mas, principalmente, do futuro (Marengo, 2006). O aquecimento global recente tem impactos ambientais intensos (como o derretimento das geleiras e calotas polares), assim como em processos biológicos (como período de floração).

No passado, estudos climáticos sobre determinada área eram aplicados em avaliações restritas, ou seja, não levava em conta os critérios para outra região que poderia ser afetada pelas mesmas mudanças no clima. Por exemplo, um modelo que assumisse desmatamento nas regiões tropicais ou aumento nas concentrações de CO₂ na ordem de 100% tem produzidos climas futuros secos e quentes na América do Sul tropical. Porém, pouco é dito sobre os impactos do desmatamento da Amazônia em outras regiões do planeta. Consequentemente, este tipo de cenário, raramente captura uma ampla gama de incertezas nas projeções de clima. A inadequação na seleção de cenários pode comprometer a interpretação de resultados de estudo de impacto (Marengo, 2006; Marengo e Nobre, 2001).

A Terra sempre passou por ciclos naturais de aquecimento e resfriamento. Da mesma forma, períodos de intensa atividade geológica lançaram à superfície quantidades colossais de gases que, de tempos em tempos, formaram uma espécie de bolha gasosa sobre o planeta, criando um efeito estufa natural. Atualmente, a atividade industrial está afetando o clima terrestre na sua variação natural, o que sugere que a atividade humana é um fator determinante no aquecimento (Marengo, 2006).

O aquecimento das águas superficiais no Pacífico Central interfere no regime de ventos sobre toda a sua região equatorial. As nuvens que normalmente produzem chuvas abundantes na parte oeste do oceano Pacífico, nas vizinhanças da Indonésia, deslocam-se para leste, para o Pacífico Central e, posteriormente, para a costa oeste da América do Sul, trazendo chuvas ao deserto peruano. A mudança de posição das chuvas no oceano provoca alterações nas

condições climáticas de várias regiões continentais ao redor do planeta devido à grande quantidade de energia envolvida no processo de formação da chuva. Grandes secas na Índia, no Nordeste do Brasil, na Austrália, Indonésia e África podem ser decorrentes desse fenômeno (Silva Dias e Marengo, 2002; Marengo e Silva Dias, 2006).

O ciclo anual das chuvas e de vazões no Brasil varia entre bacias e, de fato, à variabilidade interanual do clima, associada aos fenômenos de *El Niño*, *La Niña*, ou à variabilidade na temperatura da superfície do mar do Atlântico Tropical e Sul podem gerar anomalias climáticas. Essas anomalias produzem grandes secas, como em 1877, 1983, 1998 no Nordeste, 2004-2006 no Sul do Brasil, 2001 no Centro-Oeste e Sudeste e, em 1926, 1983, 1998 e 2005 na Amazônia (Marengo e Silva Dias, 2006; Marengo, 2007).

O impacto da variabilidade climática sobre os recursos hídricos no Brasil deverá ser mais dramático. Como por exemplo, pode citar-se o Estado de São Paulo, que neste ano apresentou a mais grave seca já registrada nos últimos 45 anos, num total de 887,2 mm de chuva, deixando os reservatórios com menos de 70% de sua capacidade total. Segundo dados do IAG (Instituto de Astronomia e Geofísica) da USP (Universidade de São Paulo), esta foi a temporada com menos chuvas desde 1969 que atingiu a marca de 814,6 mm. No Nordeste brasileiro, este é um problema recorrente, principalmente no semiárido. Como por exemplo, ano de 2012, o Nordeste sofreu com a maior seca já registrada nos últimos 50 anos, apresentando menos de 600 mm de precipitações e levou a morte de mais de 4 milhões de animais. Atualmente, a disponibilidade hídrica *per capita* na região é insuficiente nos Estados do Rio Grande do Norte, Paraíba, Pernambuco, Alagoas e Sergipe, sem contar as variações regionais, que tornam a situação ainda mais insustentável para a população (Marengo, 2006).

2.2 Semiárido Nordestino

A região Nordeste, com quase 1,56 milhão de km² (18,2% do território nacional), contém a maior parte do Semiárido brasileiro, o qual é formado por um conjunto de espaços que se caracterizam pelo balanço hídrico negativo, resultante das precipitações médias anuais inferiores a 800 mm (Moura et al., 2006). Segundo informações contidas no Dossiê Nordeste Seco (Nordeste, 1999), o Semiárido brasileiro é a área semiárida mais povoada do mundo e, em função das adversidades climáticas, associadas a outros fatores históricos, geográficos e

políticos, que remontam centenas de anos, abriga a parcela mais pobre da população brasileira (Moura et al., 2006).

Em consequência do comportamento das chuvas no Semiárido e da reduzida capacidade de retenção de água na maioria dos solos (Jacomine, 1996), grande parte da população é altamente dependente da água de chuva, de sua captação e de seu armazenamento, uma vez que os rios apresentam regime temporário. A única exceção do rio São Francisco, que se destaca em meio à grande área seca. O clima constitui a característica mais importante do Semiárido, principalmente devido à ocorrência das secas estacionais e periódicas (Mendes, 1997), que determinam o sucesso da atividade agrícola e pecuária e a sobrevivência das famílias.

As mudanças climáticas globais decorrentes do acúmulo de gases de efeito estufa (GEE) na atmosfera, dentre os quais se destacam o dióxido de carbono – CO₂, metano – CH₄ e os óxidos nitrosos – NO_x representam um desafio sem precedentes para a humanidade. O quarto relatório do IPCC (IPCC-AR4) concluiu que o aumento de temperatura média da Terra observada durante os últimos cem anos é devido às atividades antropogênicas relacionadas, principalmente, à queima de combustíveis fósseis e de florestas tropicais (Solomon et al., 2007).

Essas mudanças climáticas só agravam a situação do semiárido, pois torna a região com um clima cada vez mais quente e seco. Isso implica diretamente na agricultura diminuindo as chances do agricultor sobreviver nessas condições de plantio, pois o clima é um dos elementos de maior importância para a agricultura. Desta forma, atividades agrícolas de sequeiro sobre o Semiárido, que em condições passadas já representavam uma incidência significativa de perda em virtude da variabilidade interanual do período chuvoso, num estado futuro de aquecimento global deverão tornar-se cada vez menos inviáveis, até a total inviabilidade de culturas que dependam exclusivamente da ocorrência de chuvas. Isto tem impactos diretos na sociedade nordestina em que, num primeiro momento, atingirá o estrato social mais pobre que vive de culturas de subsistências no Semiárido. Entretanto, com o tempo, esse será um problema para toda sociedade, visto que, após atingir o meio rural, consequentemente irá afetar todos aqueles que são dependentes da produção, não só industrial, como também de pequenas fazendas que sustentam, de alguma forma, as indústrias alimentícias (Lima, Cavalcante e Marin, 2011).

A escassez de chuvas provoca a seca de rios ou a diminuição do nível da água, como é o caso do rio Apodi-Mossoró, presente na região semiárida do Rio Grande do Norte. Este rio, que passa pelas zonas que circundam o município de Governador Dix-Sept Rosado-RN, abastece algumas fazendas, sítios e populações que vivem da agricultura. Portanto, é interessante que as mudanças climáticas sejam acompanhadas pela pesquisa que promova a avaliação da medida dos fatores que possam determinar ou ainda influenciar essas mudanças climáticas como é o caso, respectivamente, dos gases do efeito estufa e o DMS.

2.3 Formação de Nuvens

2.3.1 Teoria Gaia

No início dos anos 1960 o cientista inglês James Lovelock recebeu do laboratório de Propulsão da NASA, um convite para trabalhar numa pesquisa que tinha como finalidade investigar a existência de vida em outros planetas, mais especificamente em Marte. Os estudos tiveram como característica marcante a análise das atmosferas de três planetas, Marte, Vênus e a Terra. Se a intenção do estudo era descobrir vida semelhante a existente na Terra, nada melhor do que analisar as atmosferas dos planetas (Carvalho, 2008; Neto e El-Hani, 2006).

Ao final da pesquisa pôde-se concluir que as composições químicas das atmosferas de Marte e Vênus eram muito parecidas entre si e bem diferentes quando comparadas à atmosfera da Terra. As atmosferas de Marte e de Vênus têm como gás mais abundante na sua superfície o dióxido de carbono (CO_2), com aproximadamente 98% de concentração, o gás nitrogênio (N_2) com cerca de 2%, o oxigênio (O_2) com muito menos de 1%, algum vapor de água (Margulis, 1987) e ausência de metano (CH_4) (Lovelock, 1987). Ao passo que na atmosfera terrestre a porcentagem desses gases é totalmente diferente, com oxigênio mais abundante do que nas atmosferas dos outros dois planetas e na quantidade de 21%, com 0,03% de dióxido de carbono, 79% de nitrogênio (Margulis, 1987) e metano a 1,7ppm (partes por milhão) (Lovelock, 1987).

A partir dessas características, Lovelock percebeu que as atmosferas de Marte e Vênus se encontravam em equilíbrio químico, exatamente ao contrário da atmosfera terrestre, e então concluiu que seria impossível existir vida em Marte e Vênus, como existe na Terra (Lovelock, 1987). Esse estudo despertou Lovelock para um fato bem interessante. Ele pôde perceber que

a atmosfera da Terra estava em constante desequilíbrio e então passou a se perguntar como seria possível uma atmosfera com gases redutores e oxidantes em total coexistência abrigar vida da forma como existe (Lima-Tavares; El-Hani, 2001).

Lovelock (1987) notou, também, que, além deste estado de desequilíbrio químico na atmosfera, a Terra manteve um estado constante do clima desde o surgimento da biosfera, aproximadamente 3,6 bilhões de anos atrás, em que a temperatura do nosso planeta não sofreu alteração significativa. Lovelock propôs, então, que a única maneira dessa atmosfera abrigar a vida, e isto ter permanecido nos últimos três bilhões de anos, seria a existência de um mecanismo de autocontrole da biosfera terrestre, controlando não somente o clima, mas também a composição química da Terra. Esta proposta aponta a existência de um sistema cibernético, autorregulador, constituído na totalidade dos organismos, rochas de superfície, oceano e atmosfera estreitamente inter-relacionados e que a vida existente no planeta seria capaz de compor alças de retroalimentação, positivas e negativas, que seriam responsáveis pela manutenção de condições favoráveis para a biosfera (Lenton; Oijen, 2002). Desta forma, o autor considerou que a Terra pudesse ser metaforicamente considerada um “superorganismo” com o objetivo da manutenção de um equilíbrio e de condições sempre favoráveis para existência de vida (Carvalho, 2008). Esse sistema cibernético foi proposto à comunidade científica em 1974, em artigos escritos juntamente com a colaboradora e amiga pessoal Lynn Margulis, recebendo o nome de teoria Gaia (Lovelock, 1987).

2.3.2 Hipótese CLAW

Utilizando a teoria Gaia como base, Charlson e colaboradores (1987) propuseram a hipótese CLAW, em que um composto denominado sulfonilpropionato de dimetila (DMSP) produzido por algas e fitoplâncton é catabolizado por bactérias ao gás sulfeto de dimetila (DMS) que desempenha papel importante no ciclo do enxofre. Em 1972 (Lovelock, Maggs e Rasmussen, 1972) foi proposto que o DMS seria o gás complementar (aquático) ao sulfeto de hidrogênio (H_2S) (terrestre) necessário ao fluxo do ciclo de enxofre. O DMS é rapidamente oxidado a aerossóis ácidos sulfatados, os quais servem como núcleos condensadores de nuvens (NCN). As nuvens, por serem brancas, refletem parte da radiação solar proveniente do espaço, resfriando a Terra como um todo. O resfriamento da superfície oceânica pelo aumento de nuvens diminui a liberação de DMS, reduzindo a formação de nuvens e vice-versa, controlando o clima, conforme Figura 1 (Curson et al., 2011). Alguns fatores influenciam na

produção de DMSP particulado intracelular das algas, incluindo: variação da luminosidade (Karsten et al., 1990), baixas temperaturas (Karsten et al., 1992), limitação de nutrientes (Sundra et al., 2002; 2007) e o aumento da osmolaridade causado pela salinidade (Lyon et al., 2011).

As nuvens tem um albedo alto e reduzem a temperatura e a luminosidade da superfície refletindo grande fração da radiação solar incidente. Portanto, já uma redução da quantidade de radiação solar que alcança a superfície da água, levando a uma diminuição na produção de DMS e nuvens e, novamente, a um aumento da incidência de raios solares sobre a superfície da água, o que fecha a alça de retroalimentação (Nunes-Neto et al., 2009).

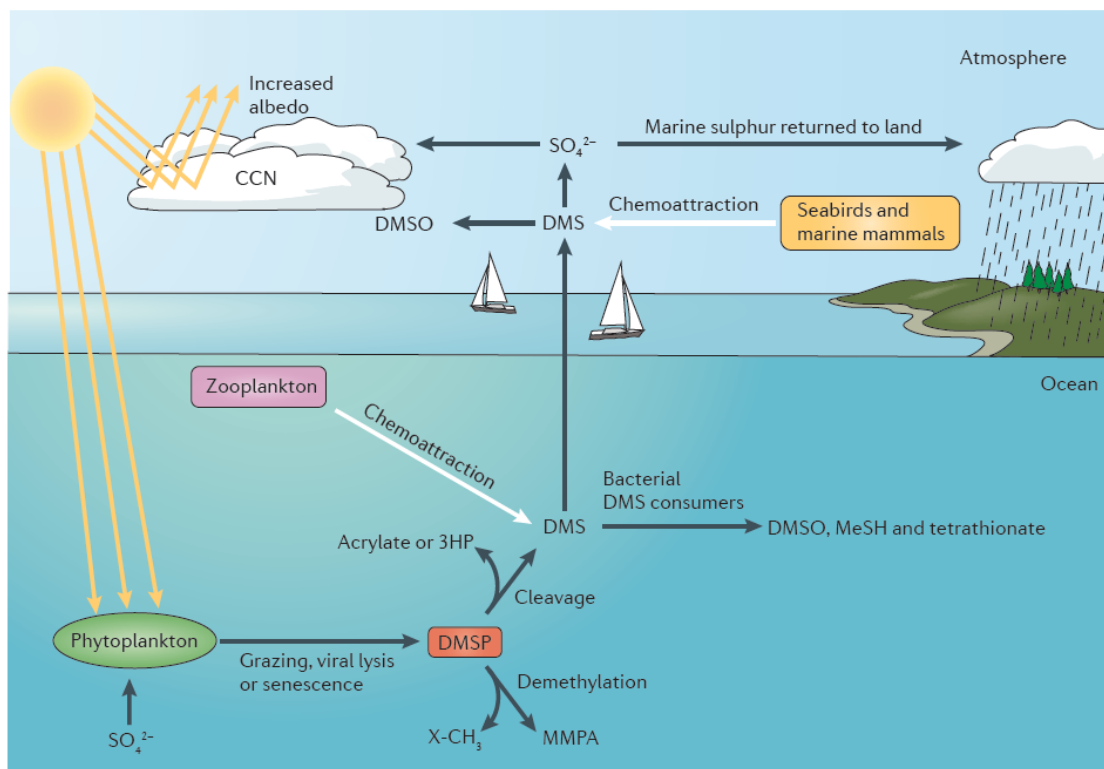


Figura 1: Formação e catabolismo de sulfonilpropionato de dimetila (DMSP). O DMSP gerado pelo fitoplâncton pode ir para dois caminhos: o primeiro é sua desmetilação com formação de mercaptopropionato de metila (MMPA) e o segundo a quebra em sulfeto de dimetila (DMS) e acrilato ou 3-hidroxipropionato (3HP). O DMS será utilizado por bactérias ou liberado na atmosfera onde será transformado em dimetilsulfóxido (DMSO) ou aerossóis sulfato, os quais agirão como núcleo condensadores de nuvens (NCN) que aumentam o reflexo de radiação solar no espaço (albedo) e retornam à superfície terrestre através de chuvas ou neve (Curson et al., 2011).

Diferentes tipos de bactérias marinhas podem catabolizar DMSP, sendo responsáveis pela maior parte do catabolismo do DMSP nos mares (Kiene e Taylor, 1988; Zubkov et al., 2001). Alguns anos atrás, um gene (*dmdA*) que é necessário para a via de desmetilação, foi descrito por Howard et al. (2006) e está envolvidos na via de um emissor de DMS (Todd et al., 2007), em diferentes gêneros bacterianos. Homólogos desse gene, ou seja, numa mesma bactéria existe uma família de homólogos que são capazes de clivar diferentes tipos de proteínas, são encontrados em outras bactérias, como, por exemplo, no mar dos Sargaços no Atlântico Norte, com capacidade DMSP-catabolizante e ocorrem com frequência entre as sequências de genes em bibliotecas metagenômicas de clones feitos diretamente a partir de bactérias marinhas, colhidas de seus ambientes naturais (Howard et al., 2006).

2.4 Metagenômica

Existe cerca de 10^{30} micro-organismos na Terra, sendo os procariotos os mais abundantes (Turnbaugh et al., 2008). Eles colonizaram diversos ambientes durante milhares de anos, incluindo aqueles que, para a maioria dos organismos, são considerados extremos. Além de colonizar o ambiente, os micro-organismos habitam outros organismos e são imprescindíveis para manutenção da vida na Terra, visto que são os principais componentes das cadeias alimentares, participam de ciclos biogeoquímicos, mantêm diversas relações simbióticas com organismos superiores (Peixoto et al., 2011) e representam uma fonte inexplorada de diversidade genética e metabólica com potencial biotecnológico. Entretanto, apenas 1% deles pode ser cultivado em laboratório pelos métodos tradicionais (Amann et al., 1990).

Esta pequena porcentagem de organismos que é isolada através das técnicas de cultivo padrão é conhecida como “organismo erva-daninha” (do inglês *Microbial Weeds*) e, em raros casos, é numericamente dominante nas comunidades das quais foi obtida (Hugenholtz, 2002). O isolamento e cultivo dos demais micro-organismos não podem ser efetuados por diversas razões, dentre as quais encontramos a falta de nutrientes ou superfície adequada, necessidade de simbiose, combinações incorretas de temperatura e/ou pressão e excesso de compostos inibitórios, entre outros (Simu e Hagstrom, 2004). Com base nessas razões, há uma opinião crescente de que o termo não-cultivável seja empregado de forma inapropriada (Rappe e Giovannoni, 2003; Stevenson et al., 2004) e que, na realidade, o que falta é a descoberta de condições adequadas. De fato, a maioria das espécies em muitos ambientes nunca foi descrita

e esse quadro não pode ser alterado até que novas tecnologias de cultivo sejam criadas (Amann et al., 1995).

Nesse contexto, a metagenômica tem revolucionado a Microbiologia abrindo caminho para acesso pleno a comunidades microbiológicas de ecossistemas complexos, tornando-se uma potente ferramenta para exploração da ecologia (Biddle et al., 2008), do perfil metabólico dessas comunidades (DeLong et al., 2006), bem como para identificação de novas biomoléculas a partir da construção de bibliotecas contendo DNA de diversas amostras ambientais tais como solo (Voget et al., 2006; Waschowitz et al., 2009), sedimentos de superfícies contaminadas (Abulencia et al., 2006), águas subterrâneas (Uchiyama et al., 2005), águas superficiais de rios (Wu e Sunda., 2009), geleiras (Simon et al., 2009), solo deserto antártico (Heath et al., 2009), rúmen de búfalo (Duan et al., 2009), dentre outros.

O termo Metagenoma é derivado do conceito estatístico de meta-análise, o processo de combinar estatisticamente análises separadas, e genômica, a análise ampla do material genético de um organismo (Schloss; Handelsman, 2003). Embora atribua-se a Handelsman et al (1998) a elaboração do termo, este pode ser adequadamente aplicado para alguns estudos anteriores como o trabalho de Torsvik et al (1990) que purificaram DNA de uma fração bacteriana isolada do solo; a pesquisa de Schmidt et al (1991) que, com o intuito de minimizar a seleção de sequências particulares, caracterizaram sequências de rRNA 16S de uma população de picoplankton a partir da clonagem direta em *E. coli* como cepa hospedeira; e o trabalho de Healy et al (1995), no qual obtiveram clones expressando celulase de um ambiente termofílico a partir de montagem das então denominadas zoobibliotecas.

Sendo assim, a metagenômica consiste na análise, independente do cultivo, de genomas presentes em dado hábitat (Riesenfeld et al., 2004). Esse campo de pesquisa vem obtendo significado progresso nos últimos anos graças aos avanços, não apenas no desenvolvimento de tecnologias de sequenciamento de última geração, cuja aplicação resultou na criação de grandes de dados derivados de vários ambientes, tais como solo e mar (Simon e Daniel, 2009), mas também na geração de novos algoritmos, *softwares* de bioinformática, permitindo análises mais robustas da enorme quantidade de dados gerados, o que vem abrindo caminho para novas aplicações, tais como metagenômica comparativa, metatranscriptômica e metaproteômica (Chistoserdova, 2010).

Além disso, abordagens metagenômicas são ferramentas plausíveis para o acesso ao perfil taxonômico e metabólico dos mais variados ambientes, como também para descoberta

de produtos de interesse biotecnológico provenientes da inexplorada diversidade genética de 99% das espécies de micro-organismos, embora esses não possam ser cultivados pelos métodos padrões da microbiologia (Amann et al., 1990).

2.4.1 Material Genético – Acesso Indireto X Acesso Direto

A etapa inicial da metodologia da Metagenômica consiste na obtenção do material genético ambiental para que este seja usado como matéria-prima em processos posteriores. Por ser o primeiro passo, a disponibilidade de extração de DNA ambiental efetivos (eDNA) torna-se um fator essencial para estudos da microbiologia ambiental (Orsini e Romano-Spica, 2001), uma vez que a eficácia do uso dessa técnica molecular neste tipo de estudo depende exatamente da obtenção de amostras de ácidos nucleicos representativas da comunidade microbiana como um todo (Miller et al., 1999; Schmeisser et al., 2007).

Existem basicamente duas formas diferentes de se extrair eDNA de solo: ou as células são separadas da amostra ambiental antes do procedimento de lise celular, denominado método indireto, ou os micro-organismos são lisados dentro da própria amostra, consistindo no método direto de extração de eDNA (Daniel, 2004; Schmeisser et al., 2007). Dentro dessas duas possibilidades, as variações de técnicas empregadas são extremamente variadas, e isso deve à enorme quantidade de tipos de solo e dos mais variados objetos dos pesquisadores em relação aos estudos posteriores com o material recuperado. A extração do eDNA depende consideravelmente das características físicas, químicas e biológicas do solo, assim como dos diferentes graus de susceptibilidade à lise celular apresentados pelas diferenças de estrutura das paredes celulares microbianas (Kauffmann et al., 2004) e da capacidade dos organismos em aderir às partículas do solo (Zhou et al., 1996). Dessa forma, nenhum método de lise celular ou de purificação de eDNA será apropriado para todos os tipos de solo e aplicações (Daniel, 2004).

Na extração indireta, as amostras de solo são primeiramente expostas a rompimentos físicos e/ou químicos suaves com o objetivo de separar os micro-organismos que estão aderidos às superfícies do solo, para então o procedimento de extração ser iniciado (Holben et al., 1988; Holben e Harris, 1995; Courtois et al., 2001). Nesta técnica, o tamanho médio do material genético isolado ao final do procedimento é maior em comparação à extração direta (Leff et al., 1995), além de ser uma forma vantajosa para amostras de solo que contenham

uma grande quantidade de substâncias que interferem no isolamento do eDNA. Isso porque a recuperação e concentração dos micro-organismos antes da etapa de lise podem evitar algumas das dificuldades impostas por essas impurezas (Daniel, 2004; Schmeisser et al., 2007).

Os procedimentos de extração direta são mais comumente usados por serem apropriados para um número maior de solos, por exigirem menos tempo de trabalho e produzirem uma maior quantidade de eDNA (Holben et al., 1988; Leff et al., 1995) a qual é tida como um produto de maior representatividade da diversidade microbiana, uma vez que os micro-organismos aderidos ao solo também são lisados. A desvantagem da técnica direta é o menor tamanho do material genético recuperado, comparado ao tamanho obtido pela extração indireta (Daniel, 2004) e a co-extração de material contaminante que interfere negativamente dos processos de detecção do eDNA, o que acaba exigindo uma etapa final de purificação (Tsai e Olson, 1992; Tebbe e Vahjen, 1993; Zhou et al., 1996; Schmeisser et al., 2007).

3. OBJETIVOS

3.1 Objetivo Geral

- Estabelecer um protocolo para extração de DNA metagenômico do solo e da água do mar.

3.2 Objetivos Específicos

- Avaliar a pureza do DNA extraído das amostras;
- Padronizar a amplificação do gene 16S rDNA;
- Padronizar a amplificação do gene *dmdA*.

4. METODOLOGIA

4.1 Coleta das Amostras do Solo e da Água do Mar

Foram realizadas três coletas em três períodos diferentes, sendo a primeira realizada no dia 19 de dezembro de 2013, a segunda realizada no dia 31 de março de 2014 e a terceira coleta realizada no dia 20 de junho de 2014. As amostras foram adquiridas de três regiões diferentes, na qual as amostras 1 e 2 foram coletadas nos bairros Aeroporto e Bom Jesus da cidade de Mossoró-RN, que apresenta uma vegetação mais típica do semiárido, com pouca cobertura vegetal e solo mais seco; as amostras 3 e 4 na margem do rio Apodi-Mossoró, próximo a Ufersa, na cidade de Mossoró-RN, local com maior cobertura vegetal contendo árvores com uma maior quantidade de folhas e um solo mais úmido; e as amostras 5 e 6 foram coletadas na salina Augusto Severo na cidade de Areia Branca-RN, região com um solo hipersalino devido a presença da salina. A figura 3 mostra os locais das coletas e as figuras 4, 5 e 6 mostram as diferenças de vegetação e período de coleta.

Todas as amostras foram obtidas com uma pequena pá, colocadas em sacos plásticos de congelamento e guardadas na geladeira para posterior extração. As amostras do bairro Aeroporto ($5^{\circ}11'34''\text{S}$ $37^{\circ}22'26''\text{W}$) foram coletadas a uma profundidade de 7 cm; as do bairro Bom Jesus ($5^{\circ}11'36''\text{S}$ $37^{\circ}22'49''\text{W}$) com 12 cm; próximo ao rio Apodi-Mossoró ($5^{\circ}11'49''\text{S}$ $37^{\circ}20'8''\text{W}$), 12 cm; nas margens do rio ($5^{\circ}11'49''\text{S}$ $37^{\circ}20'9''\text{W}$), profundidade de 20 cm; na entrada da Salina Augusto Severo ($4^{\circ}57'39''\text{S}$ $37^{\circ}8'18''\text{W}$), com 8 cm; e no lavador de sal da mesma salina ($4^{\circ}57'49''\text{S}$ $37^{\circ}8'16''\text{W}$), 10 cm.

Como controle positivo, foi utilizada a água do mar. Para isso, foi realizada a coleta de quatro amostras de água foram retiradas da praia da Redonda, Icapuí-CE, numa profundidade de aproximadamente um metro, em tubos *falcon* num volume de 15 ml. Após retirados do mar, os tubos foram resfriados em caixa de isopor contendo gelo até a chegada do laboratório onde foram imediatamente utilizados.



Figura 3: Bairro Aeroporto (A), Bairro Bom Jesus (B), Margem do Rio Apodi-Mossoró (C), Entrada da Salina Augusto Severo (D), Lavador de Sal da Salina Augusto Severo (E).



Figura 4: Imagens da primeira coleta. Bairro Aeroporto (A), bairro Bom Jesus (B), margem do rio Apodi-Mossoró (C), salina Augusto Severo (D).



Figura 5: Imagens da segunda coleta. Bairro Aeroporto (A), bairro Bom Jesus (B), margem do rio Apodi-Mossoró (C), salina Augusto Severo (D).



Figura 6: Imagens da terceira coleta. Bairro Aeroporto (A), bairro Bom Jesus (B), margem do rio Apodi-Mossoró (C), salina Augusto Severo (D).

4.2 Extração de DNA Metagenômico do Solo

A extração de DNA metagenômico foi realizada conforme descrito por Roger e Bendich (1985) com modificações. Em resumo, pesou-se 1 g de cada amostra em uma balança semi-analítica e adicionou em um tubo cônico de 15 ml, em seguida aqueceu-se o CTAB (brometo de cetil-trimetilamônio) (Tabela 1) no micro-ondas por 30 seg e adicionou 750 µl do mesmo mais 0,2% (15 µl) de β-mercapto-etanol em cada tubo. As amostras foram postas em banho Maria a 65 °C por 10 min e, em seguida, adicionou-se 500 µl de clorofórmio:álcool isoamílico gelado numa proporção de 24:1. Centrifugou-se a 4000 rpm por 10 min a 10 °C. Retirou-se o sobrenadante e adicionou em um novo tubo cônico de 15 ml. Adicionou-se 400 µl de clorofórmio gelado em cada amostra e centrifugou mais uma vez a 4000 rpm por 10 min a 10 °C. Retirou-se o sobrenadante e adicionou em um microtubo de 1,5 ml e adicionou 330 µl de isopropanol gelado. As amostras foram colocadas no freezer a -20 °C por, no mínimo, uma noite. No dia seguinte, as amostras foram centrifugadas mais uma vez a 4000 rpm por 10 min a 10 °C. Deixou-se os microtubos secarem em papel toalha por, no mínimo 2 h. Em seguida, adicionou-se 100 µl de etanol 70% gelado em cada microtubo e centrifugou mais uma vez a 4000 rpm por 10 min a 10 °C. Deixou-se secar mais uma vez em papel toalha, até que todo o álcool evaporasse e foi adicionado 50 µl de água pura para ressuspender o DNA. Todas as amostras foram armazenadas no freezer a -20 °C.

Tabela 1: Fórmula do tampão CTAB

CTAB BUFFER (1L)	
Tris HCl 1M, pH 8.0	10 ml
NaCl 5M	208 ml
EDTA 0,5M	40 ml
CTAB	20 g
Água Destilada q.s.p.	1000 ml

4.3 Extração de DNA Metagenômico da Água do Mar

A extração do DNA metagenômico da água do mar segue, basicamente, os mesmos passos da extração do DNA do solo com modificações. Ao invés da amostra ser pesada em balança, a água do mar foi filtrada em membrana de 0,22 µm, utilizando uma seringa. Depois

de filtrar 15 ml de água, retirou-se a membrana de dentro do invólucro e colocou dentro de um microtubo de 1,5 ml. As etapas seguintes são as mesmas da extração (Item 3.3), desde a adição de CTAB e β -mercatpo-etanol, até a adição final de água pura para criopreservação.

4.4 Gel de Agarose para Eletroforese

Para análise da integridade e da pureza do DNA, as amostras foram adicionadas em gel de agarose 1% (p/v) em tampão TBE 1X (Tabela 2), na qual foi também utilizado como tampão de corrida. As amostras de DNA foram aplicadas no gel com 1/8 do volume de corante de corrida *Blue Gree* (LGC Biotecnologia) tendo como padrão de peso molecular *DNA Ladder* 100pb (Green Bio Resource). As amostras foram analisadas no gel de corrida com os seguintes parâmetros: 100 V, 80mA e 30min.

Tabela 2: Tampão de Gel de Agarose e de Corrida

TAMPÃO TBE (10X) pH 8.0	
Tris Base	108 g
Ácido Bórico	55 g
EDTA 0,5M	40 ml
Água Destilada q.s.p.	1000 ml

4.5 Amplificação do gene 16S rDNA e *dmdA*

Para amplificação do DNA ribossomal (rDNA) 16S, foram utilizados os iniciadores universais 27F (5' AGGGTTTGATCMTGGCTGAG 3') e o 1100R (5' GCAACGAGCCGCAACCC 3') conforme Lane (1991). E para avaliação dos genes *dmdA* foram utilizados os iniciadores universais DMDAUR (5' TCIATICKITCIATIAIRTTDGG 3') e o DMDAUF (5' GTICARITITGGGAYGT 3') conforme Varaljay et al (2010). A reação de amplificação foi realizada com 12,5 μ l da mistura – 2X PCR Mix (1,5 mM $MgCl_2$) (LGC Biotecnologia), 9 μ l de água pura, 8% de DMSO, 10 pmol de cada iniciador e entre 50 e 100 ng de cada amostra de DNA, apresentando um volume total de 25 μ l. As amplificações, tanto do gene 16S como para o gene *dmdA*, foram realizadas no termociclador (Amplitherm

Thermal Cyclers TX96Plus) programado para 40 ciclos após desnaturação inicial a 94 °C por 5 min. O ciclo de amplificação para o gene 16S rDNA foi composto de 3 etapas: desnaturação a 94 °C por 1 min, anelamento a 55 °C por 1 min e extensão a 72 °C por 2 min, com uma extensão final a 72 °C por 5 min. E para amplificação do gene *dmdA*, segundo Varaljay et al (2010) também foi composto 3 etapas: desnaturação a 94 °C por 1 min, anelamento a 32 °C por 1 min e extensão a 72 °C por 2 min, com uma extensão final a 72 °C por 5 min. Para a análise da amplificação, aplicou-se 4 µl do produto final da reação em cadeia da polimerase (PCR) com o corante Blue Green (LGC Biotecnologia) em gel de agarose a 1%.

4.6 Bioinformática

Para avaliação das sequências da família dos genes *dmdA* e *ddd* depositadas em bancos de dados genômicos, foram utilizados os bancos de dados *National Center for Biotechnology Information* (NCBI) e o *Basic Local Alignment and Search Tool* (BLAST). Todas as sequências foram recuperadas no formato FASTA. Inicialmente, as sequências foram utilizadas para avaliar possíveis famílias de genes existentes nos bancos de dados. Em seguida, as sequências foram alinhadas no programa Bioedit (Hall, 1999) e submetidas ao programa MEGA 6.0 (Tamura et al., 2011) para construção da árvore filogenética baseada em Máxima Verossimilhança. Para o alinhamento foi utilizado penalidade de início 15 e de continuação 5. Para construção da árvore filogenética foi utilizado o teste Tamura-3-parâmetros (Tamura, 1992) com 500x de bootstrap.

Para desenho dos iniciadores degenerados, primeiramente realizou-se uma busca no NCBI por sequências depositadas dos genes e das proteínas DMSP liase. Essas sequências foram guardadas no formato FASTA, cada uma num arquivo para uso posterior. As sequências de proteínas foram analisadas no programa CODEHOP (*CO*nssensus-*DE*generate *Hybrid Oligonucleotide Primers*) para desenhar novos primers degenerados. Após essa etapa, os iniciadores foram testados *in silico* utilizando o programa FastPCR (Kalendar e Lee, 2014), para avaliar os melhores iniciadores para um futuro experimento de amplificação.

5. RESULTADOS E DISCUSSÃO

5.1 Coleta das Amostras

A partir da análise das figuras 4, 5 e 6, percebe-se que não há tanta diferença na vegetação entre a primeira e a segunda coleta, já entre a primeira e a terceira coleta, há uma grande diferença na vegetação, principalmente no bairro Bom Jesus (Fig. 4B, 5B e 6B), devido ao período chuvoso na cidade de Mossoró. Já na região salina (Fig. 5D) no período de chuvas é possível ver que a quantidade de sal depositada no solo é diminuta quando comparada às outras coletas (Fig. 4D e 6D). Até mesmo a região do rio Apodi-Mossoró, sofre uma mudança na vegetação, pois percebe-se na primeira coleta (Fig. 4C) uma vegetação mais característica da Caatinga. Na terceira coleta (Fig. 6C) a vegetação está mais vistosa quando comparada com a segunda coleta (Fig. 5C). Essas figuras mostram a mudança de períodos na região, marcada, nesse caso, pela presença de chuvas constantes. A região que teve menos mudanças com relação à vegetação e aparência do solo foi a do bairro Aeroporto (Fig. 4A, 5A e 6A).

5.2 Extração de DNA Metagenômico do Solo

Seguindo a metodologia de Roger e Bendich (1985), as amostras apresentaram pouco ou não foi possível detectar DNA. Além disso, nas amostras que apresentaram uma quantidade mínima de material genético percebeu-se que esse se degradava rapidamente. Desta forma, foi realizada uma limpeza com fenol:(clorofórmio:álcool isoamílico) numa proporção de 50:50 em todas as amostras. Após essa etapa, todas as amostras apresentaram DNA em quantidade satisfatória e estável. Apesar dos métodos diretos de extração (há a separação prévia dos micro-organismos da amostra) apresentarem melhores resultados, todas as amostras deste trabalho apresentaram DNA após a extração indireta (onde não há separação dos micro-organismos), mostrando que a metodologia utilizada foi eficiente (Fig. 7, 8 e 9). Alguns autores citam que o uso de esferas de vidro para extração do DNA metagenômico apresenta resultados mais satisfatórios, já que substâncias no solo se ligam ao DNA, prejudicando a pureza do mesmo (Robe et al., 2003; Rodrigues, 2011). Nesse caso, foi feita uma purificação posterior que estabilizou a presença do DNA extraído, parando a degradação. A análise de integridade e pureza em gel de agarose mostrou que algumas amostras

apresentaram mais DNA do que outras, indicando que as diferenças de solo e vegetação podem influenciar na quantidade de micro-organismos.



Figura 7: Extração de DNA Metagenômico da primeira coleta. A primeira amostra indica o marcador molecular, e as demais, Aeroporto (1), Bom Jesus (2), rio Apodi-Mossoró (3), margem do rio Apodi-Mossoró (4), entrada da salina Augusto Severo (5) e lavador de sal da salina (6).

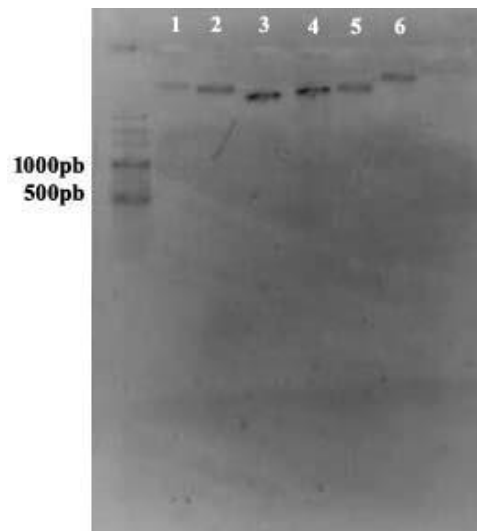


Figura 8: Extração de DNA Metagenômico da segunda coleta. A primeira amostra indica o marcador molecular, e as demais, Aeroporto (1), Bom Jesus (2), rio Apodi-Mossoró (3), margem do rio Apodi-Mossoró (4), entrada da salina Augusto Severo (5) e lavador de sal da salina (6).



Figura 9: Extração de DNA Metagenômico da terceira coleta. A primeira amostra indica o marcador molecular, e as demais, Aeroporto (1), Bom Jesus (2), rio Apodi-Mossoró (3), margem do rio Apodi-Mossoró (4), entrada da salina Augusto Severo (5) e lavador de sal da salina (6).

5.3 Extração de DNA Metagenômico da Água do Mar

A água do mar foi utilizada como controle positivo, pois segundo Varaljay (2010), cerca de 10% dos micro-organismos presentes na água, apresentam o gene *dmdA*. Após testar quatro tipos de metodologias diferentes, por fim, conseguiu-se extrair o DNA da água do mar coletada na praia de Redonda, Icapuí-CE. Primeiramente, seguiu-se o mesmo protocolo utilizado por Roger e Bendich (1985), porém não se obteve resultados. O segundo método foi a utilização de um filtro de papel para reter os micro-organismos no filtro e, após as etapas de purificação e precipitação, não houve resultados. Na terceira metodologia empregada, utilizou uma pequena amostra de alga marinha junto com água do mar, para verificar se havia micro-organismos aderidos a alga, e as demais etapas já citadas, porém, mais uma vez, não se obteve resultados. Só então, seguindo o mesmo princípio de Rodrigues (2011) que utilizou um filtro Sterivex (0,22 µl), que foi possível obter a extração do DNA. Em vez do filtro supracitado, foi utilizada uma membrana para seringa (0,22 µl) da marca Kasvi, a qual foi obtido o DNA como mostra a figura 10.



Figura 10: Extração de DNA Metagenômico da água do mar. A primeira amostra indica o marcador molecular, e as demais, membrana 0,22 µl (1), filtro de papel (2), alga marinha (3), água do mar (4).

5.4 Amplificação do 16S rDNA

Em comparação com os resultados obtidos por Lane (1991), que utilizou os iniciadores universais 27F (o mesmo utilizado nesse trabalho) e 1041R, os resultados desse trabalho se mostraram positivos em relação à amplificação do gene 16S rDNA, utilizando os iniciadores 27F e 1100R, tanto das amostras do solo, quando das amostras de água (Fig. 11). Desta forma, o protocolo utilizado nesse trabalho mostrou-se eficiente em relação a trabalhos já publicados. O gene 16S rDNA é o que mais está sendo utilizado para determinar a presença de membros dos domínios *Bacteria* e *Archea* em uma amostra ambiental (Almeida, 2009) já que é o gene que se mantém mais conservado ao longo da evolução dos procariotos (Holt et al., 1994).



Figura 11: Amplificação do gene 16S rDNA de amostras aleatórias do solo e da amostra da água do mar, utilizando os iniciadores universais 27F e 1100R. Onde, o poço 1 representa a amostra 2 da segunda coleta, o poço 2 a amostra 4 da primeira coleta, o poço 3 a amostra 4 da segunda coleta, o poço 4 a amostra 4 da terceira coleta, o poço 5 a amostra 5 da primeira coleta e o poço 6 a amostra 5 da segunda coleta.

Percebe-se que as amostras 5 e 6 apresentaram bandas bem mais baixas que as demais, mostrando que o gene não foi amplificado. Por serem amostras de um ambiente hipersalino, isso pode ser explicado pela presença de micro-organismos extremófilos ou por serem anaeróbicos.

5.5 Amplificação do gene *dmdA*

Mesmo utilizando os iniciadores universais DMDAUR e DMDAUR, como descrito no subitem 2.5, temperatura de anelamento muito baixa (32 °C), para aumentar a inespecificidade do anelamento dos iniciadores e a enzima *Taq polimerase* utilizada por Varaljay et al (2010), as amostras tanto do solo, como da água não apresentaram resultados. Como citado pela mesma autora, os genes testados em seu artigo, não oferecem 100% de certeza na amplificação, visto que vários fatores podem influenciar na determinação da família de genes, como por exemplo, profundidade de onde foram coletados, estação do ano, salinidade, temperatura, pH, entre outros. Como nenhum dos tipos de amostras utilizados neste trabalho (solo e água) apresentou amplificação dos genes que codificam a enzima DMSP liase, algumas hipóteses foram criadas: a) a família de genes (*dmdA*) pode ser diferente das bactérias presentes no meio analisado; b) os iniciadores utilizados podem não ser

adequados para as amostras, visto que nem mesmo a amostra da água do mar apresentou resultado positivo; c) presença em maior quantidade de genes da família *ddd*, que também codificam para DMSP liase, o que iria contra os estudos anteriores em que menos de 10% das bactérias apresentariam essa família de genes (Varaljay, 2012); d) a existência de uma nova família de genes que codificam a mesma enzima.

5.6 Bioinformática

Avaliando-se a árvore filogenética do gene *dmdA* das sequências depositadas nos bancos de dados genômicos (Fig. 12), foi possível observar a grande variabilidade encontrada neste gene. Há uma grande distância entre os genes desses micro-organismos, inclusive entre gêneros semelhantes, como é o caso do gênero *Ruegeria* que forma outra raiz, assim como apresenta maior semelhança com *Dinoroseobacter*. Essas informações serviram de base na escolha dos novos iniciadores a serem testados posteriormente

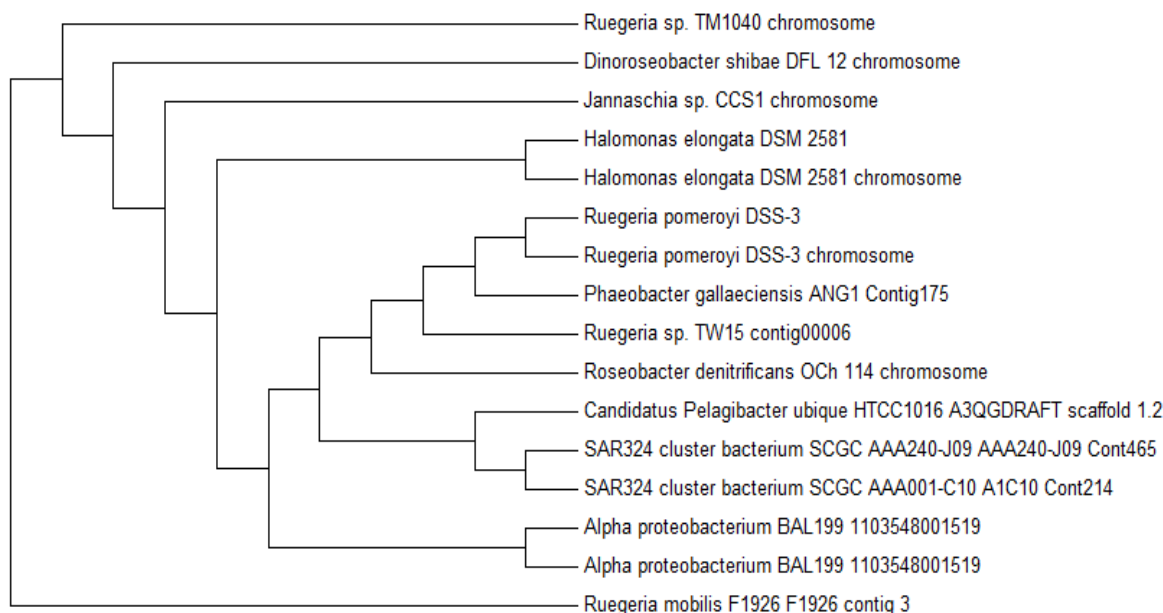


Figura 12: Árvore filogenética das sequências referentes ao gene *dmdA* disponíveis em bancos de dados e analisadas neste trabalho. Análise de Máxima Verossimilhança com teste filogenético Tamura-3-parâmetros (Tamura, 1992).

5.6.1 Desenho dos iniciadores

Todos os iniciadores desenhados (Anexo I) foram testados no *in silico*. Dos iniciadores específicos, dos quais 16 pares foram selecionados (Anexo II), alguns dos genes provenientes de diferentes bactérias geraram a mesma sequência de iniciadores, mas a maioria dos genes analisado apresentou um par de iniciadores específico que seria mais eficiente na sua amplificação. A maior parte desses iniciadores escolhidos reconheciam ou do início ao meio do gene, que tem aproximadamente 1,5 kb, ou do meio ao fim do gene, o que também demonstra grande variabilidade desse gene dentro da família *dmdA*. Dos iniciadores degenerados, dos quais seis iniciadores foram escolhidos (Anexo III), a escolha foi feita baseada na quantidade de genes de diferentes organismos que amplificaria, no tamanho do *amplicon* gerado e na temperatura de anelamento da amplificação.

Cada um dos iniciadores terá de ser testado novamente, tentando-se obter uma maior degeneração destes para aumentar seu espectro de amplificação desses genes. Uma das possibilidades será avaliar o alinhamento das sequências de nucleotídeos e de proteínas buscando regiões conservadas para desenho manual de iniciadores avaliando-os *in silico*.

6. CONSIDERAÇÕES FINAIS

Através deste trabalho, foi possível estabelecer os protocolos de extração e amplificação do gene 16S rDNA. Foram necessárias modificações no protocolo original utilizado para estabilizar o material genético extraído, prevenindo a degradação do mesmo. Não foi possível amplificar os genes *dmdA*, o que pode ter ocorrido por inúmeras razões, tais como: variabilidade muito grande desses genes, mesmo se utilizando iniciadores altamente degenerados e com a presença de muitas inosinas, que conferem maior degeneração ao primer; os micro-organismos apresentarem uma família diferente para os genes que codificam a DMSP liase; esses micro-organismos não apresentarem o gene que codifica tal enzima; entre outros. Essa hipótese foi confirmada com a construção da árvore filogenética dos genes depositados em bancos de dados genômicos, mostrando que o gene não é conservado. A partir desses resultados, foram desenhados novos iniciadores, referentes ao gene *dmdA* que serão utilizados posteriormente.

7. PERSPECTIVAS FUTURAS

- Testar os iniciadores desenhados neste trabalho para amplificar o gene *dmdA*;
- Desenhar iniciadores para outros genes (família *ddd*);
- Tentar amplificar os genes *ddd* com os iniciadores desenhados.

8. REFERÊNCIAS

- ABULENCIA, C. B., et al. Environmental whole genome amplification to access microbial populations in contaminated sediments. **Appl. Environ. Microbiol.**, v.72, p. 3291-3301, 2006.
- ALMEIDA, Bianca Caetano de. **Diversidade de Bactérias em Amostras de Água do Mar no Canal de São Sebastião**. 2009. 198 f. Tese (Doutorado) - Curso de Ciências Biológicas, Universidade de São Paulo, São Paulo, 2009.
- AMANN, R. J., et al. Combination of 16S rRNA targeted oligonucleotide probes with flow-cemetry for analyzing mixed microbial populations. **Appl. Environ. Microbiol.**, v.56, p. 1919-1925, 1990.
- AMANN, R. I.; W. LUDWIG, *et al.* Phylogenetic Identification and in situ Detection of Individual Microbial Cells Without Cultivation. **Microbiol. Rev.**, v.59, n.1, p. 143-69, 1995.
- BIDDLE, J. F., S. FITZ-GIBBON, S. C. SCHUSTER, J. E. BRENCHLEY e C. H. HOUSE. Metagenomic signatures of the Peru Margin sub sea floor biosphere show a genetically distinct environment. **Proc. Natl. Acad. Sci. USA**, v.105, p. 10583-10588, 2008.
- CARVALHO, Anderson Abbehusen Freire de. Uma Relação Entre a Teoria Gaia, o Aquecimento Global e o Ensino de Ciências. **Ver. Edu. em Ciên. Tecno.**, Alexandria, v. 1, n. 2, p.107-120, jul. 2008.
- CHISTOSERDOVA, L. Recent progress and new challenges in meta-genomics for biotechnology. **Biotechnol. Lett.**, v.32, p. 1351-1359, 2010.
- COURTOIS, S., A FROSTEGARD, et al. Quantification of bacterial subgroups in soil: comparison of DNA extracted directly from soil or from cells previously released density gradient centrifugation. **Environ. Microbiol.**, v.3, n.7, p. 431-439, 2001.
- CURSON ,A. R. J.; TODD, J. D.; SULLIVAN, M. J.; JOHNSTON, A. W. B.; Catabolism of Dimethylsulphoniopropionate: Microorganisms, Enzymes and Genes. **Sch. Biol. Scien.**, University of East Anglia. 2011.
- DANIEL, R. The Soil Metagenome – A Rich Resource for the Discovery of Novel Natural Products. **Curr. Opin. Biotechnol.**, v.15, Jun, p.199-204. 2004.

DELONG, E. F., et al. Community genomics among stratified microbial assemblages in the ocean's interior. **Science**, v.311, p. 496-503, 2006.

DUAN, C. J., et al. Isolation and partial characterization of novel genes encoding acidic cellulases from metagenomas of buffalo rumens. **J. Appl. Microbiol.** v.107, p. 245-250, 2009.

EL-HANI, C.N. e EMMECHE, C. On Some Theoretical Grounds for an Organism-centered Biology: Property Emergence, Supervenience, and Downward Causation. **Theo. Biosci.**, v. 119, n° 3-4, p. 234-275, 2000.

HALL, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucl. Acids. Symp.**, v.41, p. 95-98, 1999.

HANDELSMAN, J.; M. R. RONDON, *et al.* Molecular Biological Access to the Chemistry of Unknown Soil Microbes: A New Frontier for Natural Products. **Chem. Biol.**, v.5, n.10, Oct, p.R245-9. 1998.

HANDELSMAN, J. Metagenomics: Application of Genomics to Uncultured Microorganisms. **Microbiol. Mol. Biol. Rev.**, v.68, n.4, Dec, p.660-85. 2004.

HEALY, F. G., et al. Direct isolation of functional genes encoding cellulases from the microbial consortia in a thermophilic, anaerobic digester maintained on lignocellulose. **Appl. Microbiol. Biotechnol.**, v.43, p. 667-674, 1995.

HEATH, C., et al. Identification of a novel alkaliphilic esterase active at low temperatures by screening a meta-genomic library from Antarctic desert soil. **Appl. Environ. Microbiol.** v.75, p. 4657-4659, 2009.

HOLBEN, W. E. & D. HARRIS. DNA-based monitoring of total bacterial community structure in environmental samples. **Mol. Ecol.**, v.4, n.5, p. 627-631, 1995.

HOLBEN, W. E., J. K. JANSSON, et al. DNA probe method for the detection specific microorganisms in the soil bacterial community. **Appl. Environ. Microbiol.**, v.54, n.3, p. 703-711, 1988.

HOLT, J., KRIEG, N. R., SNEATH, P. H. A. **Bergey's manual of determinative bacteriology**. 9 ed. Baltimore: Williams & Wilkins, 1994.

HOWATD, E. C., HENRIKSEN, J. R., BUCHAN, A., REISCH, C. R., BUERGSMANN, H., et al. Bacterial taxa that limit flux from the ocean. **Science**, v.314, p. 649-652, 2006.

HUGENHOLTZ, P. Exploring prokaryotic diversity in the genomic era. **Genome. Biol.**, v.3, n.2, p. PREVIEWS0003, 2002.

JACOMINE, P. K. T. Solos sob caatingas – Características e uso agrícola. In: ALVAREZ, V. H.; FONTES, L. E. F.; FONTE, M. P. F. (Ed.). **O solo nos grandes domínios morfoclimáticos do Brasil e o desenvolvimento sustentado**. Viçosa: SBSC/UFV. 1996. p.95-111.

Kalendar R., Lee D., Schulman A.H. FastPCR software for PCR, *in silico* PCR, and oligonucleotide assembly and analysis. *DNA Cloning and Assembly Methods*, **Metho. Mol. Biol.**, Svein Valla and Rahmi Lale (ed.), 1116: 271-302, 2014.

KARSTEN, U.; WIENCKE, C.; KIRST, G.O. The Effect of Light Intensity and Daylength on the Beta-dimethylsulphoniopropionate (DMSP) Content of Marine Green Nacroalgae from Antarctica. **Plant. Cell. Environ.**, 13: 989–993. 1990.

KARSTEN, U.; WIENCKE, C.; KIRST, G.O. Dimethylsulphoniopropionate (DMSP) Accumulation in Green Macroalgae from Polar to Temperate Regions: Interactive Effects of Light versus Salinity and Light versus Temperature. **Polar. Biol.** 12, 603–607. 1992.

KAUFFMAN, I. M., SCHMITT, J., et al. DNA isolation from soil samples for cloning in different hosts. **Appl. Microbiol. Biotechnol.**, v.64, n.5, p. 664-670, 2004.

KIENE, R. P.; TAYLOR, B.F. Demethylation of Dimethylsulfonylpropionate and Production of Thiols in Anoxic Sediments. **Appl. Environ. Microbiol.**, v.54, p. 2208–2212, 1988.

LANE, D.J. 16S/23S rRNA Sequencing. In: E. Stackebrandt; M. Goodfellow (Ed.). **Nucleic Acid Techniques in Bacterial Systematics**. New York: John Wiley and Sons, p. 115–175, 1991.

LEFF, L. G., DANA, J. R., et al. Comparison of methods of DNA extraction from stream sediments. **Appl. Microbiol. Biotechnol.**, v.61, n.3, p. 1141-1143, 1995.

LENTON, T. M.; OIJEN, M. V. Gaia as a complex adaptive system. Philosophical Transactions of the Royal Society of London B: **Biol. Scien.**, v.357, n.1421, p. 683-695, 2002.

LIMA, Ricardo da Cunha Correia; CAVALCANTE, Arnóbio de Mendonça Barreto; MARIN, Aldrin Martin Perez. **Desertificação e Mudanças Climáticas no Semiárido Brasileiro**. Campina Grande: Instituto Nacional do Semiárido, p. 211, 2001.

LIMA-TAVARES, M.; EL-HANI, C. N. Um olhar epistemológico sobre a transposição didática da teoria Gaia', *Investigações em Ensino de Ciências* 6, a4. 2001. Disponível em: http://www.if.ufrgs.br/public/ensino/vol6/n3/v6_n3_a4.htm#Top. Último acesso em: 23/04/2014.

LOVELOCK, J.E. Letter to the Editors - Gaia as Seen Through the Atmosphere. **Atm. Env. Pergamon. Press.**, 6: 579-580. 1972.

LOVELOCK, J. E. Gaia – Um modelo para a dinâmica planetária e celular. In: THOMPSON, W. I. *Gaia: uma teoria do conhecimento*. São Paulo: Gaia, p. 77-90, 1987.

LOVELLOCK, J.E.; MAGGS, R.J. & RASMUSSEN, R.A. Atmospheric Dimethyl Sulphide and the Natural Sulphur Cycle. **Nature**, 237: 452-453. 1972.

LYON, B. R.; LEE, P. A.; BENNETT, J. M.; DITULLIO, G. R.; JANECH M. G. Proteomic Analysis of a Sea-Ice Diatom: Salinity Acclimation Provides New Insight into the Dimethylsulfoniopropionate Production Pathway. **Plant. Physiol.**, 157: 1926–1941. 2011.

MARENGO, J., NOBRE, C. A. The Hydroclimatological framework in Amazonia. In **Biogeochemistry of Amazonia**, Richey, J., McClaine, M., Victoria, R., Eds., pp. 17-42. 2001.

MARENGO, J., SILVA DIAS, P. Mudanças climáticas globais e seus impactos nos recursos hídricos. Capítulo 3 em **Águas Doces do Brasil: Capital Ecológico, Uso e Conservação**, pp.63-109. Eds. A.Rebouças, B., Braga e J. Tundisi. Editora Escrituras, SP, 2006.

MARENGO, J. On the Hydrological Cycle of the Amazon Basin: A historical review and current State-of-the-art, no prelo, **Rev. Brasil. Meteorol.**, 2006.

MARENGO, J. A. *Mudanças climáticas globais e seus efeitos sobre a biodiversidade – caracterização do clima atual e definição das alterações climáticas para o território brasileiro ao longo do século XXI*. 2. ed. Brasília: Ministério do Meio Ambiente, v.1, p.214, 2007.

MARENGO, José A.. **Mudanças Climáticas Globais e seus Efeitos sobre a Biodiversidade**: Caracterização do Clima Atual e Definição das Alterações Climáticas para o Território Brasileiro ao Longo do Século XXI. 26. ed. Brasília: Biodiversidade, p. 201, 2006.

MARGULIS, L. Os primórdios da vida. In: THOMPSON, W. I. **Gaia: uma teoria do conhecimento**. São Paulo: Gaia, p. 91-101, 1987.

MENDES, B. V. Biodiversidade e desenvolvimento sustentável do SemiÁrido. Fortaleza: SEMACE, 108 p. il., 1987.

MILLER, D. N., BRYANT, J. E., et al. Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. **Appl. Environ. Microbiol.**, v.65, n.11, p. 4715-4724, 1999.

MOURA, Magna Soelma Beserra de et al. Clima e água de chuva no Semi-Árido. **Semiárido**, Recife, v. 1, n. 2, p.36-59, jul. 2011.

NUNES NETO, Nei Freitas; EL-HANI, Charbel Niño. Gaia, Teleologia e Função. **Episteme**, Porto Alegre, v. 11, n. 23, p.15-48, jun. 2006.

NORDESTE sertanejo: a região semi-árida mais povoada do mundo. Estudos Avançados, São Paulo, v. 13, n. 35 p.60-68, Mai/Ago. 1999.

NUNES-NETO, N. F.; EL-HANI, C. N; CARMO, R. S., Uma Conexão entre Algas e Nuvens: Fundamentos Teóricos da Hipótese CLAW e suas Implicações para as Mudanças Climáticas. **Oecol. Brasil.**, 13(4): p.596-608, Dec. 2009.

ORSINI, M. & V. ROMANO-SPICA. A microwave-based method for nucleic acid isolation from environmental samples. **Lett. Appl. Microbiol.**, v.33, n.1, p. 17-20, 2001.

PEIXOTO, R. S.; VERMELHO, A.; ROSADO, A. S. Petroleum-degrading enzymes: bioremediation and new prospects. **Enzyme. Reseach.**, v.2011, p. 1-7, 2011.

RAPPE, M. S. & J. GIOVANNONI. The uncultured microbial majority. **Annu. Rev. Microbiol.**, v.57, p. 369-394, 2003.

RIENSENFELD, C. S.; P. D. SCHLOSS, *et al.* Metagenomics: Genomic Analysis of Microbial Communities. **Annu. Rev. Genet.**, v.38, p.525-52. 2004.

ROBE, P., R. NALIN, et al. Extraction of DNA from soil. **Eur. J. Soil Biol.**, v.39, p. 183-190, 2003.

RODRIGUES, Thiago Bruce. **DIVERSIDADE METAGENÔMICA MICROBIANA DE BIOMAS TERRESTRES E MARINHOS**. 2011. 269 f. Tese (Doutorado) - Curso de Ciências Biológicas (genética), Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.

ROGERS, Scott O.; BENDICH, Arnold J.. Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. **Plant. Mol. Biol.: Martinus Nijhoff Publishers**, Dordrecht, v.5, n.5, p.69-76, 1985.

SCHLOSS, P. D. & J. HANDELSMAN. Biotechnological Prospects from Metagenomics. **Curr. Opin. Biotech.**, v.14, p.303-10. 2003.

SCHMEISSER, C., STEELE, H., et al. Metagenomics, biotechnology with non-culturable microbes. **Appl. Environ. Microbiol.**, v.75, n.5, p. 955-962, 2007.

SCHMIDT, T. M, DELONG, E. F., PACE, N. R. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. **J. Bacteriol.**, v.173, p. 4371-4378, 1991.

SILVA DIAS, P., MARENGO, J., Águas atmosféricas. Águas Doces no Brasil - capital ecológico, usos múltiplos exploração racional e conservação. Aldo da Cunha Rebouças, Benedito Braga Jr., José Galizia Tundisi, Eds. 2ns Edition, IEA/USP, pp. 65-116, 2002.

SIMON, C. & R. DANIEL. Achievements and new Knowledge Unraveled by Metagenomic Approaches. **Appl. Environ. Microbiol.**, v.85, p.265-67. 2009.

SIMON, C., HERATH, J., ROCKSTROH, S., DANIEL, R. Rapid identification of genes encoding DNA polymerases by function-based screening of metagenomic libraries derived from glacial ice. **Appl. Environ. Microbio.**, v.75, p. 2964-2968, 2009.

SIMU, K. & HAGSTROM, A. Oligotrophic bacterioplankton with a novel single-cell life strategy. **Appl. Environ. Microbiol.**, v.70, n.4, p. 2445-2451, 2004.

SOLOMON, S.; D. QIN; M. MANNING, et al. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. **Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA**, p.996. 2007.

STEVENSON, B. S.; S. A. EICHORST, *et al.* New Strategies for Cultivation and Detection of Previously Uncultured Microbes. **Appl. Environ. Microbiol.**, v.70, n.8, p.4748-55, Aug. 2004.

SUNDA, W. G.; KIEBER, D. J.; KIENE, R. P.; HUNSTMAN, S. An Antioxidant Function for DMSP and DMS in Marine Algae. **Nature**, 418: 317–320. 2002.

SUNDA, W. G.; HARDISON, R.; KIENE, R. P.; BUCCIARELLI, E.; HARADA, H. The Effect of Nitrogen Limitation on Cellular DMSP and DMS Release in Marine Phytoplankton: Climate Feedback Implications. **Aquat. Sci.**, 69: 341–351. 2007.

TAMURA, K. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. **Mol. Biol. Evol.**, 9, p. 678-687, 1992.

TAMURA, K.; PETERSON, D.; PETERSON, N.; STECHER, G.; NEI, M.; KUMAR, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. **Mol. Biol. Evol.**, 28, 2731-2739, 2011.

TEBBE, C. C. & VAHJEN, W. Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant DNA from bacteria and a yeast. **Appl. Environ. Microbiol.**, v.58, n.8, p. 2657-2665, 1993.

TODD J.; ROGERS R.; LI Y. G.; WEXLER M.; BOND P. L.; SUN L.; CURSON A. R. J.; MALIN G.; STEINKE M.; JOHNSTON A. W. B. Structural and Regulatory Genes Required to Make the Gas Dimethyl Sulfide in Bacteria. **Science** v.315, p. 666–669, 2007.

TORSVIK, V.; J. GOKSOYR, *et al.* High Diversity in DNA of Soil Bacteria. **Appl. Environ. Microbiol.**, v.56, n.3, p. 782-7, 1990.

TSAI, Y. L. & OLSON, B. H. Rapid method for direct extraction of DNA from soil and sediments. **Appl. Environ. Microbiol.**, v57., n4., p. 1070-1074, 1991.

TURNBAUGH, P. J. & GORDON, J. I. An invitation to the marriage of metagenomics and metabolomics. **Cell.**, v.134, p. 708-713, 2008.

UCHIYAMA, T., T. ABE, T. IKEMURA, K. WATANABE. Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes. **Nat. Biotechnol.**, v.23, p. 88-93, 2005.

VARALJAY, V. A.; HOWARD, E. C.; SUN, S.; MORAN M. A. Deep Sequencing of a Dimethylsulfoniopropionate-degrading Gene (*dmdA*) by Using PCR Primer Pairs Designed on the Basis of Marine Metagenomic Data. **Appl. Environ. Microbiol.**, v.76, p. 609–617, 2010.

VARALJAY, Vanessa A. et al. Bacterial Dimethylsulfoniopropionate Degradation Genes in the Oligotrophic North Pacific Subtropical Gyre. **Journals Aem: Appl. Environ. Microbiol.**, Georgia, v.12, p.2775-2782, 2012.

VOGET, S., H. L. STEELE, W., STREIT, R. Characterization of a metagenoma-derived halotolerant cellulose. **J. Biotechnol.**, v.126, p. 26-36, 2006.

WASCHKWITZ, T., ROCKTROH, S. DANIEL, R. Isolation and characterization of metalloproteases with a novel domain structure by construction and screening of metagenomic libraries. **Appl. Environ. Microbiol.**, v.74, p. 2506-2516, 2009.

WU, C. & B. SUN. Identification of novel esterase from metagenomic library of Yangtze River. **J. Microbiol. Biotechnol.**, v.19, p. 187-193, 2009.

ZHOU, J. & BRUNS, M. A. DNA recovery from soils of diverse composition. **Appl. Environ. Microbiol.**, v.62, n.2, p. 316-322, 1996.

ZUBKOV, M. V.; FUCHS, B. M.; ARCHER, S. D.; KIENE, R. P.; AMANN, R.; BURKILL, P. H. Linking the Composition of Bacterioplankton to Rapid Turnover of Dissolved Dimethylsulphoniopropionate in an Algal Bloom in the North Sea. **Environ. Microbiol.**, 304–311. 2001.

9. ANEXOS

ANEXO I – LISTA DOS INICIADORES ESPECÍFICOS DESENHADOS NO PROGRAMA FastPCR

1F1_1_68-87	gggtcaagggtatac cgctc
1F2_1_91-110	aatcacatgttgctgcc cac
1F3_1_12-8-149	aggccgattgcgcca tctgaa
1F4_1_23-2-251	atcagcccgcgcgaca tgga
1F5_1_25-5-274	gatggccgatgaccag tggt
1F6_1_27-8-297	acgtgccacggctga tcat
1F7_1_34-1-360	ccgatcattactggctg tcg
1F8_1_36-9-388	tgggacactgctgcaat tcg
1F9_1_42-4-446	gagatcgctgaacccg atgtctc
1F10_1_4-64-483	gaccaggggcccgcg atctg
1F11_1_4-81-503	ctgatggcgcggtctt tggcga
1F12_1_5-03-524	aggcggtgcgcgatat ccgctt
1F13_1_5-40-559	gctggcctttcaggga gtcg
1F14_1_5-74-594	cgctcaggctggtcga aacag
1F15_1_6-15-634	cgagggttcggaactg ggca
1F16_1_6-34-656	atgccgctgtggaacg cgctgtt
1F17_1_6-61-680	gccggtgcggacctga acgt
1F18_1_6-94-713	cccaacaatatcgagc gcgt
1F19_1_7-22-742	ggttgttgagctatggc aacg
1F20_1_7-52-771	gcgagaacacgccgt atgaa
1F21_1_7-84-803	aagtcttgcaattcgcc cga
1F22_1_8-01-822	cgaggactatatcggc aaggca
1F23_1_8-16-837	caaggcagcactggc cgaacag
1F24_1_8-59-880	cagatccgggactgg tgatcg
1F25_1_8-79-898	cggtggcgagattccg ccct
1F26_1_9-51-972	gtcagcgatccattccc ctgaa
1F27_1_9	atggtggatcgagcc

97-1017	attgg
F28_1_10-25-1047	gcaccgggatggaagt ggaaacg
1F29_1_1-068-1088	tacgggtgcgcgaggg gttctg
1R1_1_10-37-1058	atgccgtcgggcgtttc cactt
1R2_1_99-9-1020	cgcccaatggctgcga tccacc
1R3_1_97-9-998	atccgatgcgcacatt cac
1R4_1_96-7-986	acattcacgccgaattc agg
1R5_1_95-0-972	ttcaggggaatggatc gctgacc
1R6_1_89-9-918	cagcggccaggcatc ctgac
1R7_1_88-2-902	tgacagggcggaatct cgcca
1R8_1_83-1-852	cggtccgttcttgccct gttcg
1R9_1_80-5-824	gctgccttgccgatata gtc
1R10_1_7-89-808	agtcctcggcggaatt gcag
1R11_1_7-59-780	caggccgcattcatac ggcgtg
1R12_1_7-30-751	gggtcatgtcgttgcca tagct
1R13_1_6-91-710	cgctcgatattgttggg gca
1R14_1_6-42-661	cggcaaacagcgcgtt ccac
1R15_1_6-11-630	cagttccgaaccctcg acat
1R16_1_5-76-595	cctgtttcgaccagcct gag
1R17_1_5-56-575	cgcgccaccacaagct cgac
1R18_1_5-26-547	aggccagccgcttgta gcggaa
1R19_1_5-07-526	aaaagcggatatcgcg cacc
1R20_1_4-76-497	aagaccgcgccatca gatcgt
1R21_1_4-35-454	ccagcggcgagacat cgggt
1R22_1_3-78-398	atccccagcccgaatt gcagc
1R23_1_3-44-363	cagcgacagccagtaa tgat
1R24_1_3-28-347	tgatcggcggccagttt cac

1R25_1_2-85-304	cgccacgatgatcgac cgtg
1R26_1_2-61-283	gcacgtagtaacactg gtcatcg
1R27_1_1-90-209	tcggggccctggatgc tgac
1R28_1_1-52-171	gtcccagacctgcaca tggt
1R29_1_1-35-154	gttccttcagatgggcg caa
1R30_1_1-08-127	gcaggctgtcgaacac cgtg
1R31_1_9-2-111	cgtgggcagcaacatg tgat
1R32_1_7-4-96	gtgattgtagacggtat agccct
1R33_1_6-25	ggcgggacgggaaga tgctg
2F1_1_1-20	atggcgcttatctctcc gtc
2F2_1_96-116	catgttgctgccaccg tctt
2F3_1_12-0-141	aagtgtggaggctgac taccgt
2F4_1_26-1-280	gccggggcggtgctttt acg
2F5_1_28-6-308	atcgtggatgagacgg gtggcat
2F6_1_31-3-334	aacgaccccggtggcg gtcaagc
2F7_1_35-0-371	ggtggattccattgcg gacag
2F8_1_36-4-383	gcggacagtgatctgc tggt
2F9_1_39-5-417	ggatttccaatggctat cgactg
2F10_1_4-41-463	tgtcagcccgtggcc attcagg
2F11_1_4-64-483	ggccaaaggcgagac aactg
2F12_1_4-87-506	gcccgatatcttggcga cag
2F13_1_5-05-524	agcatccgcgacatca agtt
2F14_1_5-45-566	aattcagggggcgtaga gatggt
2F15_1_5-69-590	tggcgcgggtcgggga ttcaa
2F16_1_5-87-606	ccaacaaggcggtttt gag
2F17_1_6-54-675	gttcaaggcgggcgga ggacctg
2F18_1_6	cgcagggtgccccaat

84-703	ggca
2F19_1_6 97-716	aatggcattgagcgga tcga
2F20_1_7 23-742	tctgctgtcttatggcaa cg
2F21_1_7 53-774	cgacaacacgccgcat gaatgc
2F22_1_7 90-812	tgtgatacgagaccg ctattgg
2F23_1_8 11-833	ggctgcatcggtcggg atgcctt
2F24_1_8 61-881	gcaacagatcagggc gattgc
2F25_1_8 93-912	accggctgcccgtctgt gac
2F26_1_9 23-942	ggctttacgccgggga tacg
2F27_1_9 66-985	tgcctggagcccggat cacc
2F28_1_9 88-1007	accaatgtcgccatcg gcat
2F29_1_1 024-1043	tgggataccggcacia ggct
2F30_1_1 042-1064	ctgcgcgttgaacgc cgatgg
2F31_1_1 081-1101	gtgcaggccggtttctg gcac
2R1_1_10 47-1066	cgccatccggcggttca acg
2R2_1_10 35-1054	tttcaacgcgcagcctt gtg
2R3_1_10 23-1042	gccttgtgccgtatcc caa
2R4_1_98 5-1006	tgccgatggcgacatt gggtgg
2R5_1_95 4-974	ctccaggcagcggaa tgacc
2R6_1_87 8-900	cagccgggtcgccctcg atcgcaa
2R7_1_85 4-874	ccctgatctgttcacc ggtc
2R8_1_80 1-822	accgatgcagccaata gcggtc
2R9_1_78 9-808	tagcgggtctgcgtatca caa
2R10_1_7 60-780	aagcccgcatcatgc ggcgt
2R11_1_7 27-748	tcatgtcgttgccataa gacag
2R12_1_7 01-720	gccctgatccgctca atgc
2R13_1_6 52-673	ggctctgccccgcttg aacag
2R14_1_6 40-659	ttgaacagcgcatccc agag
2R15_1_5 77-596	cctgttttgaatacccc ga
2R16_1_5 42-561	ctcacgcccctcgaatt caa

2R17_1_4 93-513	gcggatgctgtcgcca aagat
2R18_1_4 66-485	gccagttcgtccgcctt tgg
2R19_1_4 35-454	ccagcgggctgacatc cggt
2R20_1_4 05-424	gcacatccagtcgata gccca
2R21_1_3 88-409	agccattggaaatccc cttgac
2R22_1_3 58-377	agatcactgtccgcaat gga
2R23_1_3 41-361	tggaaatccaccagcg gtctt
2R24_1_3 28-347	cggtcttcggccagctt gac
2R25_1_2 84-303	accgtctcatccacga tgg
2R26_1_2 69-289	cgatgggcacgtaaaa gcacc
2R27_1_1 88-210	atccgggccacgcagt tcgacct
2R28_1_1 68-188	tgacgctcgaccgaaa catcc
2R29_1_1 52-171	atcccagacctgcacat gtt
2R30_1_1 39-158	acatgttccttgaggtg acg
2R31_1_1 24-143	tgacggtagtcagcctc cac
2R32_1_1 11-131	gcctccacacttcaaaa gacg
2R33_1_9 2-114	gacggtgggcagcaa catgcggt
2R34_1_7 7-99	catgcggtgttagacc gtgtagg
2R35_1_4 5-67	ctgcggcctccacacc atcgga
2R36_1_2 -21	cgacggagagataag cgcca
3F1_1_65- 84	cagttaaatgggggtgc aacg
3F2_1_10 7-128	tgtacattcctctgtgact ttgg
3F3_1_17 8-198	gatgttcagtgagac gacaa
3F4_1_20 1-221	agaaattacaggaccc gatgc
3F5_1_24 7-269	cctcgagatttatcaaa gcttgc
3F6_1_39 1-410	ttatgggcacaaagggg ttgc
3F7_1_43 4-456	aaattatcgaacccgat gtctcc
3F8_1_57 4-594	atcgtctcccgtaccgg ttgg
3F9_1_58 8-609	cggttggtctagtgaac taggt
3F10_1_6	gagatggatctaaggg

26-647	gaacca
3F11_1_7 28-750	gtggaatgttatcgta catgca
3F12_1_9 01-923	gagcctttaactggacc taacac
3F13_1_9 73-994	tctgcagtttattcccca agac
3F14_1_1 040-1060	agattggaacagtgcct gagg
3R1_1_10 39-1060	cctcaagcactgttcca atctc
3R2_1_89 7-917	gggccagttaaaggctc acaa
3R3_1_73 9-758	tcagcatctgcatggta cga
3R4_1_66 7-686	ccgtaaagtittccagct gc
3R5_1_63 0-650	agttgggtccccttagat cca
3R6_1_56 1-582	ggagacgattaaagga atgcca
3R7_1_44 7-469	gtaattgaagagggga gacatcg
3R8_1_39 8-419	gaattgattgaacccc ttgtg
3R9_1_25 8-278	tgaccaactgcaagctt tgat
3R10_1_2 09-228	ttttgtagcatcggtcc tg
3R11_1_1 85-206	atttctactgtcgctcc actg
3R12_1_1 14-134	ggagatccaaagtcac gagga
4F1_1_15- 36	gttaatttctccagaac ccgt
4F2_1_46- 65	ttcagcagccgggttg agga
4F3_1_60- 81	tgaggactgtggtgtaa aagca
4F4_1_10 6-126	ctgcctttgatcttccgc agc
4F5_1_12 6-146	cctggaagaggattact ggca
4F6_1_15 2-171	gtgaatccgtacaggct tgg
4F7_1_17 1-190	ggatgttctctgccaga gac
4F8_1_18 8-207	gacaggttgagatcac cgga
4F9_1_22 7-246	tccagttaatgactccc cgt
4F10_1_2 39-258	ctccccgtgacctgtca cag
4F11_1_2 78-297	tttatgctccgtctgtg ac
4F12_1_3 48-367	ccactggtggccttcaa tcg
4F13_1_4 36-455	gaacctgatatttgccc gct

4F14_1_4 52-471	cgcttcagttcaggg acca
4F15_1_4 77-497	cgaagaactactttcc gggt
4F16_1_5 36-558	gcaggacatttaactac caggga
4F17_1_5 71-590	gtggcaagatccggtt ggag
4F18_1_5 85-604	ttggagcaaacagggt gggt
4F19_1_6 66-687	aggtgaagacctgaat gttgg
4F20_1_6 91-712	ggctgtcctaacttgatt gagc
4F21_1_7 66-786	ccgtttgaatgcggtctt gat
4F22_1_8 00-819	tggacgcagacattga aagc
4R1_1_80 3-822	aaggctttcaatgtctg cgt
4R2_1_77 5-796	tgacatactgatcaaga ccgca
4R3_1_76 3-782	agaccgcattcaaacg gcgt
4R4_1_74 8-768	cggcgttgatcgtatc ccat
4R5_1_71 2-733	acagcagcccgtctc aattcg
4R6_1_69 2-713	cgctcaatcaagttagg acagc
4R7_1_67 5-694	agcccggaaccaacatt cagg
4R8_1_66 1-682	cattcaggtcttcacctt ttgc
4R9_1_62 3-645	gagttgtcctccaact ccgcgt
4R10_1_5 85-607	cgaaaccacctgtttg ctccaa
4R11_1_5 65-584	ccggatcttgccacca acat
4R12_1_5 50-571	ccaacatctttgtccct ggta
4R13_1_5 34-556	cctggtagttaaatgtc ctgctg
4R14_1_4 74-495	ccggaaaagtagttctt cggct
4R15_1_4 61-480	ttcggttttggtccctg aa
4R16_1_4 20-440	ggttctgtaacaaggg catcc
4R17_1_3 99-418	atccaaaaccagttgcc agc
4R18_1_3 65-387	gagcattacatctgaat cagcga
4R19_1_3 53-372	atcagcgattgaaagc cacc
4R20_1_2 89-308	cctccggtttcgtcaca gag
4R21_1_2	tccaagctctgcctgtg

50-270	acag
4R22_1_2 37-257	tgtgacaggtcacggg gagtc
4R23_1_1 90-209	ggtcgggtgatctcaac ctg
4R24_1_1 77-197	tcaacctgtctctggca ggaa
4R25_1_1 56-175	catcccagacctgtac ggat
4R26_1_1 41-161	acggattcacataagtg ccag
4R27_1_1 23-143	cagtaatcctctccag gctg
4R28_1_9 4-116	atcaaaggcaggagc atgtgggt
4R29_1_5 1-71	ccacagtcctcaaccc ggctg
4R30_1_2 7-47	aaaggagttgaacggg ttctg
4R31_1_2 -24	ggaaattaacagttcag gacgca
5F1_1_68- 87	gcgtcaagggctatac ggtc
5F2_1_92- 111	accacatgctgctggc aacg
5F3_1_11 3-133	tgtttgacagcctggaa tacg
5F4_1_13 1-151	acgactgcgcccattc gaagg
5F5_1_14 4-163	tctgaaggagcatgtcc agg
5F6_1_23 2-251	atcagcccgcgcgaca tgga
5F7_1_28 0-299	gtgccgacggtcgatc acaa
5F8_1_34 1-360	ggatcattactggctg tcg
5F9_1_35 7-377	gtcgcttccgatggc gacct
5F10_1_3 83-403	agtttgctgctgggatc gcca
5F11_1_3 96-415	gatcgccattgccaag ggct
5F12_1_4 30-449	acggagcccgatgtct cacc
5F13_1_4 64-483	gcccgaaggcggatg acctt
5F14_1_4 83-504	tatggcccgcgtgttcg gcgaa
5F15_1_5 04-523	aacgggtgcgcgatatc cggt
5F16_1_5 32-551	tacaagcgtctggactt cca
5F17_1_5 74-593	cgctccgctgtgtcca agca
5F18_1_5 87-606	ccaagcaggcgcgctt tgag
5F19_1_6 09-630	ctatgtcgaaggttcgg atctg

5F20_1_6 85-704	gcgggctgccccaatg cgat
5F21_1_7 25-746	tgctgagctatggcaat gacat
5F22_1_7 54-774	gcaaacacgccgtatg aatgc
5F23_1_7 90-809	tgcaattgccggaag acta
5F24_1_8 08-828	tacattggcaagtcgca actg
5F25_1_8 58-878	tcagatccggccgctg gtgat
5F26_1_9 33-952	ccggggttgagaggtc ggct
5F27_1_9 53-972	cggccattttctcacctg aa
5F28_1_9 78-997	ggatgaacgccgcgata ggca
5F29_1_1 022-1041	cgggcactgcgttgga agtt
5F30_1_1 074-1093	tcaggagaagttctgg cgggt
5R1_1_10 75-1094	caccgccagaactctc ctg
5R2_1_10 28-1049	gggtgtctcaactccaa cgcag
5R3_1_10 10-1029	agtgcggcgcccca atgag
5R4_1_98 2-1001	accatgcctatcgcg cggt
5R5_1_95 1-972	ttcaggtgagaaaatg gccgag
5R6_1_88 6-905	tccaggcagggcgagg atctg
5R7_1_87 0-889	tctgcccgccgatcac cagc
5R8_1_78 9-808	agtcctccggcggaattg cag
5R9_1_75 9-779	agcccgcatcatcacg gcgtg
5R10_1_7 30-752	cgggtcatgtcattgcc atagct
5R11_1_6 95-714	gatccgctcgatcgcat tgg
5R12_1_6 68-687	cgcccgacgttcagg tctt
5R13_1_6 52-671	tcttcaccagcgccga acaa
5R14_1_6 17-638	ggcatccccagatccg aacctt
5R15_1_5 35-554	ccctggaagtccagac gctt
5R16_1_5 23-542	agacgctttagcgga agaa
5R17_1_5 01-520	ggatatcgcgaccgtt tcg
5R18_1_4 80-499	cgaacacgcgggcca taagg
5R19_1_4	cataaggtcatccgct

67-486	tcg
5R20_1_4 27-446	gagacatcgggctccg tgat
5R21_1_3 95-416	aagcccttggaatgg cgatcc
5R22_1_3 49-369	atcggaagcgacag ccagta
5R23_1_3 28-347	tgatccgcgccagttt cac
5R24_1_2 78-299	ttgtgatcgaccgtcg cacat
5R25_1_2 53-275	tagcactggctgtctgc catctt
5R26_1_2 00-219	acgcagcgcatcggg acctt
5R27_1_1 70-189	ctggcgctcgaggcg acat
5R28_1_1 51-171	atccagacctggaca tgctc
5R29_1_1 33-152	tccttcagatggcgca gtc
5R30_1_1 16-135	gtcgtattccaggctgt caa
5R31_1_1 03-122	ctgtcaaacaccgttgc cag
5R32_1_8 2-101	agcatgtggttagac cgt
5R33_1_6 9-88	agaccgtatagcccttg acg
5R34_1_2 -23	cgggaggggaagatg cttgcca
6F1_1_44- 63	tcctgcgtctcgaccat caa
6F2_1_98- 117	atgccgatggcaacatt cgt
6F3_1_11 1-133	cattcgtgcgaagtcc ggcgac
6F4_1_14 9-168	acctgccccacatgctc gtc
6F5_1_25 8-280	tggacacccgcagga gcgcatcg
6F6_1_28 7-307	aggcagccaatggcc gtctgc
6F7_1_31 7-337	aaccgctccagcccgc attcg
6F8_1_35 5-375	ggatcatgtgtgcccgt agct
6F9_1_39 1-410	aaccgcgtcgatcccg ttgg
6F10_1_4 72-491	gcccagatcgctacctt cca
6F11_1_5 10-529	cctgccgggaatagcc cgag
6F12_1_5 81-601	aaccgcaggctgacc gcatcc
6F13_1_6 10-629	caccgctcgaccaaa tcat
6F14_1_6 55-674	gggcgagacatcggg ctctt

6F15_1_7 23-743	ggagcagatcactgtc ggcga
6F16_1_7 55-774	aaccggctcctcggaca gctt
6F17_1_7 72-791	cttcagcgccacggga tcgt
6F18_1_8 30-851	agacactgccccggca acatcc
6F19_1_8 68-890	ggtcagaagctgcatc aactgc
6F20_1_9 65-984	tggtaatcctcctcgac gga
6F21_1_9 82-1003	ggactcgaacacagtc ggcagc
6F22_1_1 004-1023	agcatgtggttagac cgt
6F23_1_1 050-1069	cccttcggagaacgg cgtg
6F24_1_1 088-1107	agggagagcaaaaggc gtcat
6R1_1_95 4-975	gaggattaccacatct caagc
6R2_1_86 8-888	aagttgatgcagcttct gacc
6R3_1_84 3-862	ggacctgcgcgggat gttgc
6R4_1_82 9-848	tgttgcggggcagtg ctg
6R5_1_81 5-836	agtgtctgtatgtccg atggt
6R6_1_72 6-745	gatcgccgacagtgat ctgc
6R7_1_69 8-717	gtgaaggcgctggcca atgc
6R8_1_66 1-681	ctggtcgaagagcccg atgtc
6R9_1_61 7-636	ctgcgatgatttgg cga
6R10_1_5 90-612	gtgttcgcgcgggatg cggtcag
6R11_1_5 17-538	cgttcgcgctcgggc tattcc
6R12_1_4 75-494	acgtggaaggttagcga tctg
6R13_1_4 30-449	tcgaggcggtgtctga tctc
6R14_1_3 98-417	ggctgccccaacggg atcga
6R15_1_3 65-384	gggctattgagctacg gcaa
6R16_1_3 47-367	caacgacatgaccgc tgga
6R17_1_2 57-279	gatgcgtcctcgggg tgtccat
6R18_1_1 07-129	ccggacttcgacacga atgttgc
6R19_1_8 3-105	atcggtatggtcgga tgacca
6R20_1_5	cccgacaccgaattg

7-76	atgg
6R21_1_3 7-56	tcgagacgcaggagg ggctg
7F1_1_20- 39	cgcacctctcggggca tcag
7F2_1_38- 59	aggccatcgggcgtct ggacat
7F3_1_66- 85	gcgtgcccggatccca atgg
7F4_1_10 4-123	atggccaccccttcgct gta
7F5_1_13 3-153	ataggcgcgaggagt caactg
7F6_1_19 7-217	caactgggcaggacat ccgca
7F7_1_24 6-265	ccggaccttcttggcc acc
7F8_1_30 1-320	ggtggcggtgttcag aacc
7F9_1_31 9-338	ccggcccaaccgcgat tcgt
7F10_1_3 65-384	tgcgtagctgagcaa tcc
7F11_1_3 95-415	cgctcgatattatggg gcaa
7F12_1_4 41-463	cggcctccatcagcgc gttccac
7F13_1_4 58-479	ttccacagcggcatcc cgacat
7F14_1_4 86-506	cctccacatagatctcg aacc
7F15_1_5 21-541	tagccagaacgcgcaa tgacc
7F16_1_5 49-569	ggccctgaaagtcgaa ccagc
7F17_1_6 07-626	gaagaccccgccatc agct
7F18_1_6 46-665	cacggccaggggtga cacat
7F19_1_6 66-685	cgggttcacgaccaat acg
7F20_1_7 21-743	caacaataatcactgt ccgcga
7F21_1_7 44-764	tggagatccaccaacg gtcct
7F22_1_7 59-778	ggctctcgccagtttc agg
7F23_1_8 07-828	tctcatccacgacggg catgta
7F24_1_8 66-885	ggggtaagcatctgca ccag
7F25_1_9 17-937	tgccgtccaccgaga catcc
7F26_1_9 53-973	gatttcaggtgggcgta atcc
7F27_1_9 78-997	ccactgtacgaaacag gggtg
7F28_1_9 94-1013	ggtgggcaggagcat ccggt

7F29_1_1 009-1030	ccggttatagatcgtgt aggct
7F30_1_1 031-1050	tttaccgccggacgcctc cac
7R1_1_10 85-1104	atggccctgatctctcc gtc
7R2_1_10 40-1060	ttctgatgggggtggag gcgtc
7R3_1_10 19-1041	tccggggtaaaagcct acacgat
7R4_1_97 7-996	accctgtttcgtacagt gga
7R5_1_95 6-975	gaggattacgccacc tgaa
7R6_1_94 0-961	cctgaaatccgccgtg caactc
7R7_1_87 2-891	gggcgtctgggtgcaga tgct
7R8_1_84 3-862	cgacctgcgcggcatg ttgc
7R9_1_81 7-838	ccaatgctattacatgc ccgtc
7R10_1_8 05-824	tgcccgctcgtgatga gacg
7R11_1_7 69-788	atcccggtgtcctgaaa ctg
7R12_1_7 54-774	aaactggccgaggac cgttgg
7R13_1_7 30-749	tctccatcgcggacagt gat
7R14_1_6 86-705	gcgcagggctaccgtc tgga
7R15_1_6 68-688	ggacgtattggtcgatg aacc
7R16_1_6 53-672	gaaccgatgtgtcac ccct
7R17_1_6 20-639	cccaaggccgatgag ctgat
7R18_1_5 94-613	gggtcttcggcgacagc atcc
7R19_1_5 40-559	ctttcaggggccacgata tgg
7R20_1_5 21-541	ggtcattgcgcgttctg gcta
7R21_1_4 83-502	cgagatctatgtggag ggct
7R22_1_4 45-464	tgtggaacgcgctgat ggag
7R23_1_4 21-441	gggcgcgatctggatg tccac
7R24_1_3 93-414	tgccccaataatcga gcgga
7R25_1_3 65-384	ggattgctcagctacg gcaa
7R26_1_3 27-348	aacacgccccacgaat gcgggt
7R27_1_3 08-327	ttgggccggttctgcaa cac
7R28_1_2	ggccaaggaaggtcc

43-262	gggtgc
7R29_1_1 93-214	ggatgtcctgccagtt gcgac
7R30_1_1 67-186	tggccgctgatggcgg gcaa
7R31_1_1 44-165	aagcgcgtggggcgag ttgacct
7R32_1_1 13-135	tatagccccgactaca gcgaagg
7R33_1_4 3-62	tgcattgtccagacgcc cgat
7R34_1_1 9-38	tgatgcccgcagaggt gcgt
8F1_1_22- 42	caccagagcagtacgc attcc
8F2_1_35- 54	cgattccatccgggtg ctc
8F3_1_47- 67	gggtctccacctctag ctgc
8F4_1_77- 96	tcccaatcgcgtcatctg ac
8F5_1_98- 117	atgccgatcgcaacatt cgt
8F6_1_11 3-133	ttcgtggcatgatccgg cgac
8F7_1_18 1-200	cggccagaattgatca cagg
8F8_1_20 5-225	caccgcttcgccggaa atctc
8F9_1_22 8-247	tcgggcggtatctgctg cacc
8F10_1_2 85-304	cgatacaccaatggc cgtg
8F11_1_2 99-318	gccgtgtgggtgttaca gaa
8F12_1_3 15-337	agaaccgcccaaggc cgattcgc
8F13_1_3 54-375	gggtcatatcattgccg tagct
8F14_1_3 69-390	cgtagctcagaagacc gtcttc
8F15_1_3 87-406	tctcgatccgctcaatg agg
8F16_1_4 30-449	gagatctgcgccagcc tcaa
8F17_1_4 42-461	agcctcaaacagacgg ttcc
8F18_1_4 74-495	caagatcgccgccttc gacgta
8F19_1_4 89-508	cgacgtagatctcgaa ccca
8F20_1_5 12-531	tgctttgagtagcccga cct
8F21_1_5 29-551	ccttgccaccaccagat cgcgtc
8F22_1_5 53-572	ttcgaaactggtagacgc caa
8F23_1_5	gcgcacagtctcgccg aagac

8F24_1_6 13-633	acgcgccatcaagtc tctgc
8F25_1_6 50-670	gcaagcggcgaaaca tccggt
8F26_1_6 65-684	tccggttcatccaccag cac
8F27_1_6 77-696	accagcacatcaaggc gcca
8F28_1_7 04-723	gcgatcccttgaccc agta
8F29_1_7 28-747	agatcactgtccgcaat gga
8F30_1_7 40-759	gcaatggagatccacc aacg
8F31_1_7 80-800	ctacgggatcattgagc atcc
8F32_1_7 98-817	tcccgccggtctcatcg acg
8F33_1_8 25-845	catagtaacattgaccg ggca
8F34_1_9 13-932	cacctggcgttcgcag gcga
8F35_1_9 35-957	tccagacttgacatg acgctt
8F36_1_9 64-983	gtggtagtcggcttcga cac
8F37_1_9 76-997	ttcgacacttcaaaca ccgtg
8F38_1_1 027-1046	gcctttgactccagcgg ctt
8F39_1_1 078-1099	gcgccgagagaggga aatggca
8F40_1_1 095-1116	tggcaggtttcagaac ggatcat
8F41_1_1 123-1145	ccctttcataccatgagt tgagg
8R1_1_11 32-1151	gctcgccctcaactcat gggt
8R2_1_11 05-1124	ggtttgccatgaccgtt ctg
8R3_1_10 90-1110	gttctgaaacctgccatt tcc
8R4_1_10 62-1081	gcgcctgcgtcgaaacc cctt
8R5_1_10 19-1039	tggagtcaaaggctac accgt
8R6_1_93 7-956	agcgtcatgtgcaagtc tgg
8R7_1_90 7-927	tgcgaacgccaggtgg aactg
8R8_1_87 5-894	gccggacgcctgatgc agat
8R9_1_84 8-867	ccgcgcgatctgcgtg gcat
8R10_1_8 31-851	gcattgatgcccggtca atgtt
8R11_1_7 91-811	tgagaccggcgggat gctcaa
8R12_1_7	ccgtaaagctcgaga

60-779	ggat
8R13_1_7 30-749	tctccattgctggacagt gat
8R14_1_7 16-737	acagtgatctgtgtact gggt
8R15_1_6 99-721	ctgggtcaaggggatc gcaaacg
8R16_1_6 85-704	caaacggctggcgctt tgat
8R17_1_6 43-665	atgtttcgccgttgcg gtgcag
8R18_1_6 18-637	caaggcagaggacttg atgg
8R19_1_5 97-617	cgcgtgtcttcggcga gactg
8R20_1_5 59-578	tccgctttggcgcttac cag
8R21_1_5 39-561	cagttcgaaggacgcg atctggt
8R22_1_5 16-536	tggcaaggtcgggcta ctcaa
8R23_1_5 01-520	ctcaagcaggggtggg ttcg
8R24_1_4 71-493	cgtcgaaggcggcgat cttgga
8R25_1_4 04-423	cgtcggggtgtccca acct
8R26_1_3 86-405	ctcattgagcggatcga gag
8R27_1_3 65-384	ggcttctgagctacgg caa
8R28_1_3 51-370	cggcaatgatatgacc gacg
8R29_1_2 80-299	ccattgggtgtatcggg cgt
8R30_1_2 36-255	gaaggcccgtgcag cagat
8R31_1_2 11-230	cgatcgagatttcggc gaa
8R32_1_1 79-198	tgtgatcaattctggcc gct
8R33_1_1 62-183	ccgctcgttgcaaatgg gcgtc
8R34_1_1 09-129	ccggatcatgccacga atgtt
8R35_1_7 9-98	tggtaaatgacgcat tgg
8R36_1_5 3-72	gggacgcagctagag gtgga
8R37_1_3 0-50	caccggatggaatgcg tactg
8R38_1_1 6-35	gtactgctctggtgcgc gaa
9F1_1_46- 65	cttttcggcgaccatcg cct
9F2_1_70- 89	cgtgatcagcaatccg gcga
9F3_1_13 6-155	gatgattccggccttta cca

9F4_1_17 5-195	ctgcgcgtcgaacccg tggtt
9F5_1_19 6-216	ggcgtcgtctgttcca ggag
9F6_1_27 9-300	gcaacgcgcgccgga cgatgtt
9F7_1_30 4-323	gccggccttctggtcgt cat
9F8_1_32 8-347	ggacttcaggctgccg agtt
9F9_1_35 2-373	gcttcgtagagcagc gtgacg
9F10_1_4 30-449	gtcgaccgggtccttgc gct
9F11_1_4 77-496	cgaccgggatgatggc aacg
9F12_1_5 09-529	ttggccaccggttctcg cagc
9F13_1_5 25-546	gcagcttttctgtcgt agtc
9F14_1_5 48-567	gagggtggttctctgat ctg
9F15_1_6 15-636	cgccgctgatgacctg ggtgtt
9F16_1_6 35-654	ttgtcttctcgatccgg ac
9F17_1_6 56-678	ttcttggcggtgccgag catgtc
9F18_1_6 85-707	cacgctctgaacttag tgccga
9F19_1_7 01-720	gtgccgagatcctcgct gat
9F20_1_7 22-741	acttggccaccggtca ggat
9F21_1_7 92-811	tgaccgcagccacctt gagg
9F22_1_8 14-833	gccgcgcagcttgitg acga
9F23_1_8 59-879	gacgtctcggccacg atgag
9F24_1_9 15-934	acggcaggagggcct gaacc
9F25_1_9 41-960	agcttctgtcgtgcag gag
9F26_1_9 53-974	tgcaggaggatgtagg gttcat
9F27_1_9 95-1014	tcggcgttggtgacga agta
9F28_1_1 059-1078	cgagttcggtctgcag cgac
9F29_1_1 095-1116	tgatcacgcctcgttg ccgac
9F30_1_1 124-1143	atggccttcgcatcat ctc
9F31_1_1 147-1167	gatctcgggtcgtcat tagc
9F32_1_1 162-1181	attagcggagatggtc ccga
9F33_1_1	gacctgggcgatctcg

180-1200	gcgtt
9F34_1_1 214-1233	ttggcgacccgctgga tgtc
9F35_1_1 267-1286	gcgtttcagatccatcg gggt
9F36_1_1 387-1408	ggcgacttcttcacca tctgc
9F37_1_1 435-1455	gagctcgatctccttgg cgac
9F38_1_1 462-1483	gccgtccttggatgcc gggga
9F39_1_1 480-1499	gggagcgccgtacgat ttgg
9F40_1_1 494-1514	atttggccaggatcacg ttgc
9F41_1_1 526-1547	cccagcgtcaccttga cggcgt
9F42_1_1 560-1581	cgatgccgtgcagcat acggtc
9F43_1_1 586-1605	gcgtccgtgtgaattt gac
9R1_1_15 88-1607	acgtcaaattcagcac ggac
9R2_1_15 60-1579	ccgatgtgcacggc atcg
9R3_1_15 36-1558	tattctgcgaacgcgg tcaagg
9R4_1_15 03-1522	gaaggggcgcaacgt gatcc
9R5_1_14 67-1486	cgctccccgatcacc aagg
9R6_1_14 42-1461	gtcaccgtcgccaagg agat
9R7_1_14 22-1441	cgagctctcgataagt tcg
9R8_1_13 35-1355	cgaccaccgaacggt tctgg
9R9_1_13 22-1341	gttctggcccacgccat cgt
9R10_1_1 307-1326	atcgtgcgcgaaggcg ccaa
9R11_1_1 273-1292	gcctgaacccgatgga tctg
9R12_1_1 240-1259	acatggccgtggaagc ggtc
9R13_1_1 220-1239	gtcgcgacatccagc gggt
9R14_1_1 190-1209	gtctcgaccaacgcgg agat
9R15_1_1 158-1178	ggaccatctccgcta gacg
9R16_1_1 019-1038	ttcgaacggggctacct gtc
9R17_1_1 001-1022	tgtcgcctacttctgca ccaa
9R18_1_9 68-987	atctgcgagttggatga acc
9R19_1_8 65-884	cgctgctcatcgtggcc gag

9R20_1_8 27-848	cgctggcgacactggt cgtcaa
9R21_1_7 91-812	gcctcaagggtgctgc ggtcaa
9R22_1_7 44-765	aaggccatgctggagg acatcg
9R23_1_6 95-715	cgaggatctcggcact aagtt
9R24_1_6 71-692	agagcgtgaccctcga catgct
9R25_1_6 53-672	ctcggcaccgccaaga aggt
9R26_1_6 19-641	aggacaacaccacggt catcgac
9R27_1_5 73-592	ggcccgatgcacgca gatcc
9R28_1_5 52-571	cgcgcagatcgagga aacca
9R29_1_5 34-553	cacctccgactacgac aagg
9R30_1_4 79-498	ggcgttgcatcatccg ggt
9R31_1_4 54-473	gtgccaccgaggtcga actg
9R32_1_4 34-456	ctgaaggagcgcaag gaccgggt
9R33_1_3 79-400	cgaggaaggcatcggt gccgga
9R34_1_3 59-380	gaggcggcgctcacgt gctcta
9R35_1_3 28-347	aactcggcagcctgaa gtcc
9R36_1_2 49-268	ccggcagatcttcgtga acg
9R37_1_2 09-229	catcgtcggcaagctc ctgga
9R38_1_1 83-202	cgacgcccaaccacgg gttcg
9R39_1_1 42-161	tcgacatggtaaaggc cgga
9R40_1_7 0-91	aatcgccggattgctga tcacg
9R41_1_4 9-70	gacggaggcgaatggt cgccgaa
10F1_1_4 6-65	cttttcggcgaccatcg cct
10F2_1_7 0-89	cgtgatcagcaatccg gcga
10F3_1_1 36-155	gatgattccggccttta cca
10F4_1_1 75-195	ctgcgcgtcgaacccg tggtt
10F5_1_1 96-216	ggcgtcgtctgttcca ggag
10F6_1_2 79-300	gcaacgcgcgccgga cgatgtt
10F7_1_3 04-323	gccggccttctggtcgt cat
10F8_1_3	ggacttcaggctgccg

28-347	agtt
10F9_1_3 52-373	gcttcgtagagcagc gtgacg
10F10_1_ 430-449	gtcgaccggctcctgc gct
10F11_1_ 477-496	cgaccgggatgatggc aacg
10F12_1_ 509-529	ttggccaccggttctcg cagc
10F13_1_ 525-546	gcagcttttctgtcgt agtc
10F14_1_ 548-567	gaggtggttctcgtat ctg
10F15_1_ 615-636	cgcgcgtcgatgaccgt gggtgtt
10F16_1_ 635-654	ttgtccttctcgatccgg ac
10F17_1_ 656-678	ttcttgccggtgccgag catgtc
10F18_1_ 685-707	cacgtctcgaacttag tgccga
10F19_1_ 701-720	gtgccgagatcctcgct gat
10F20_1_ 722-741	acttgccaccgggtca ggat
10F21_1_ 792-811	tgaccgcagccactt gagg
10F22_1_ 814-833	gccgcgcagcttgttg acga
10F23_1_ 859-879	gacgtctcggccacg atgag
10F24_1_ 915-934	acggcaggagggcct gaacc
10F25_1_ 941-960	agcttctgtcgtgcag gag
10F26_1_ 953-974	gtcaggaggatgtagg gttcat
10F27_1_ 995-1014	tcggcgttggtgacga agta
10F28_1_ 1059-1078	cgagttcgggtctgcag cgac
10F29_1_ 1095-1116	tgatcacgccctcgttg ccgac
10F30_1_ 1124-1143	atggccttcgcgatcat ctc
10F31_1_ 1147-1167	gatctcgggtcgtcat tagc
10F32_1_ 1162-1181	attagcggagatggtc ccga
10F33_1_ 1180-1200	gacctgggcgatctcg gcgtt
10F34_1_ 1214-1233	ttggcgaccgcgtgga tgtc
10F35_1_ 1267-1286	gcgtttcagatccatcg ggt
10F36_1_ 1387-1408	ggcgacttcttcacca tctgc
10F37_1_ 1435-1455	gagctcgatctccttgg cgac

10F38_1_ 1462-1483	gccgtccttggtgatcc gggga
10F39_1_ 1480-1499	gggagcgccgtacgat ttgg
10F40_1_ 1494-1514	atttgccaggatcacg ttgc
10F41_1_ 1526-1547	cccagcgtcaccttga cggcgt
10F42_1_ 1560-1581	cgatgccgtgcagcat acggtc
10F43_1_ 1586-1605	gcgtccgtgctgaatt gac
10R1_1_1 588-1607	acgtcaaattcagcac ggac
10R2_1_1 560-1579	ccgtatgctgcacggc atcg
10R3_1_1 536-1558	tattctgccaacgccg tcaagg
10R4_1_1 503-1522	gaagggccgcaacgt gatcc
10R5_1_1 467-1486	cgtccccggatcacc aagg
10R6_1_1 442-1461	gtcaccgtcgccaagg agat
10R7_1_1 422-1441	cgagctctcgataagt tcg
10R8_1_1 335-1355	cgaccaccgcaacggt tctgg
10R9_1_1 322-1341	gttctggcccacgccat cgt
10R10_1_ 1307-1326	atcgtgcgcgaaggcg ccaa
10R11_1_ 1273-1292	gcctgaacccgatgga tctg
10R12_1_ 1240-1259	acatggccgtggaagc ggtc
10R13_1_ 1220-1239	gtcgccgacatccagc gggt
10R14_1_ 1190-1209	gtctcgaccaacgccg agat
10R15_1_ 1158-1178	ggaccatctccgcta gacg
10R16_1_ 1019-1038	ttcgaacggggctacct gtc
10R17_1_ 1001-1022	tgtcgccctacttcgtca ccaa
10R18_1_ 968-987	atctgcgagttggatga acc
10R19_1_ 865-884	cgctgctcatcgtggcc gag
10R20_1_ 827-848	cgctggcgacactggt cgtcaa
10R21_1_ 791-812	gcctcaagggtggtgc ggtcaa
10R22_1_ 744-765	aaggccatgctggagg acatcg
10R23_1_ 695-715	cgaggatctcggcact aagtt
10R24_1_ 	agagcgtgaccctcga

671-692	catgct
10R25_1_653-672	ctcggcaccgccaagaaggt
10R26_1_619-641	aggacaaccacacgggtcatcgac
10R27_1_573-592	ggcccgatgcacgcatcc
10R28_1_552-571	cgcgcagatcgaggaacca
10R29_1_534-553	cacctccgactacgacaagg
10R30_1_479-498	ggcgttgccatcatccgggt
10R31_1_454-473	gtgccaccgaggtcgaactg
10R32_1_434-456	ctgaaggagcgcaaggaccgggt
10R33_1_379-400	cgaggaaggcatcggtgccgga
10R34_1_359-380	gaggcggcgtcacgctgctcta
10R35_1_328-347	aactcggcagcctgaagtcc
10R36_1_249-268	ccggcagatcttctgtaacg
10R37_1_209-229	catcgctggcaagctcctgga
10R38_1_183-202	cgacgccaaccacgggttcg
10R39_1_142-161	tcgacatggtaaaggccgga
10R40_1_70-91	aatcgccggattgctgatacgc
10R41_1_49-70	gacggaggcgtggtcgccgaa
11F1_1_15-36	gttaattcctccagaaccgt
11F2_1_106-126	ctgccttgatcttccgcagc
11F3_1_126-146	cctggaagaggattactggca
11F4_1_152-171	gtgaatccgtacaggtctgg
11F5_1_171-190	ggatgttctgccagagac
11F6_1_227-246	tccagttaatgactcccgt
11F7_1_239-258	ctccccgtgacctgtcacag
11F8_1_278-297	tttatgctccgctctgtgac
11F9_1_348-367	ccactgggtggctttcaatcg
11F10_1_436-455	gaacctgatatttgccgct
11F11_1_452-471	cgcttgagttcagggacca
11F12_1_477-497	agaagaactactttccgggt

11F13_1_571-590	gtggcaagatccgggtggag
11F14_1_585-604	ttggagcaaacagggtggtt
11F15_1_678-697	gaatgttggtccgggctgtc
11F16_1_691-712	ggctgtcctaacttgattgagc
11F17_1_764-786	cgccgttgtaatgcgggtcttgaa
11F18_1_805-826	gcagacattgaaagcctgagtc
11F19_1_934-956	gggaaaattacttccgacacctg
11F20_1_964-985	cgatattcagccttctcgcat
11F21_1_1002-1021	caaacaccttgagacgctc
11R1_1_974-994	acctcgcaatgccagggaagg
11R2_1_949-970	aatatcgaggagaccagggtgtc
11R3_1_917-937	tcccaaccacattactgccga
11R4_1_803-822	caggcttcaatgtctgcgt
11R5_1_775-797	ctgacatactttcaagaccgca
11R6_1_763-782	agaccgcattcaaacggcgt
11R7_1_748-768	cggcgttgatcgtatccat
11R8_1_716-735	agacagcagcccgtttcaa
11R9_1_692-713	cgctcaatcaagttaggacagc
11R10_1_667-689	ggaccaacattcaagcttcacc
11R11_1_623-645	gagttgtcctcccaactccgcat
11R12_1_590-611	atttcaaaaccacctgtttgc
11R13_1_565-584	ccggatcttgccaccaacat
11R14_1_550-571	ccaacatcttggttccctggta
11R15_1_530-552	gtagttaatgtactgctgcgga
11R16_1_475-495	ccgggaaagtagttcttctgc
11R17_1_458-477	tgctttgtgtccctgaactg
11R18_1_420-440	ggttctgtaacaagggcatcc
11R19_1_399-418	atccaaaaccagttgccagc
11R20_1_365-387	gagcattacatctgaatcagcga
11R21_1_365-387	atcagcgattgaaagc

353-372	cacc
11R22_1_339-358	gccaccagtgggtgtgtgcta
11R23_1_289-308	cctccgggttctgtcacagag
11R24_1_250-270	tccaagctctgcctgtgacag
11R25_1_237-257	tgtgacaggtcacggggagtc
11R26_1_194-214	tgtccggccccgggtgactcaa
11R27_1_177-197	tcaactgtctctggcaggaa
11R28_1_156-175	catcccagacctgtacggat
11R29_1_141-161	acggattcacataagtgccag
11R30_1_123-143	cagtaatcctcttccaggctg
11R31_1_94-116	atcaaaggcaggagcatgtggt
11R32_1_51-70	cacaggcctcaaccccgctg
11R33_1_27-47	aaaggagttgaacgggttctg
11R34_1_2-24	ggaaattaacagttcagacgca
12F1_1_10-31	gaaattctcatagacggtcgca
12F2_1_34-56	gcgcattccgtccgggtgtctcaa
12F3_1_100-121	gccgatggcgacattggtgta
12F4_1_18-139	ggtatgatcgggcgacatgtg
12F5_1_32-151	accatgtggccgaggacacc
12F6_1_205-224	caccgcgtcgccctcaatgg
12F7_1_228-247	tggggcggatctgctgcacc
12F8_1_271-292	aagggcgtcgcgccgatacag
12F9_1_299-318	gccgatgggtgttgagaa
12F10_1_326-345	agcccgcatctatgcggcgt
12F11_1_356-378	gtcatatcgttgccgtactcag
12F12_1_394-416	ccgctcatcaggttcggacagc
12F13_1_415-434	gcctgcccgcacctccagat
12F14_1_440-460	ccgcctcaaacagccgatgc
12F15_1_487-509	ttcgacataaatctcgacccac
12F16_1_552-571	cctcgaaactgataaacgcca

12F17_1_630-651	cggcctttggtcctga attgc
12F18_1_653-672	agcggtagacatccg gctc
12F19_1_665-684	tccggctcatccacca gcac
12F20_1_696-717	agccattggcgatgcc cttcac
12F21_1_728-747	agatcactgtccgaat gga
12F22_1_743-765	atggagatccgccagc ggctctc
12F23_1_760-782	gtctcggccagcttca ccgcta
12F24_1_802-821	gccggctctcatccacg atgg
12F25_1_827-846	tagtagcactgaccgg gcat
12F26_1_937-956	ccagacctgcacatgg cggt
12F27_1_961-980	gtggtgataatccgcct caa
12F28_1_984-1004	tctcgaacacggttgga agca
12F29_1_1050-1069	caccctccgaaaacgg ggta
12F30_1_1081-1100	gcgggacagggaaatt ctgg
12F31_1_1096-1115	tctgggtgaaacactg cca
12R1_1_1095-1114	ggcagtgtttcaaccca gaa
12R2_1_1050-1069	taccccgtttcggagg gtg
12R3_1_1019-1038	gggtgcaaaggctaca ccgt
12R4_1_964-983	gcgttgaggcggattat cac
12R5_1_943-962	acctcaaacgccatgt gcag
12R6_1_856-877	aatgctgacccgcgct gatctg
12R7_1_830-851	gcatgatcccgggtca gtgcta
12R8_1_806-825	gtgcccacgtggatg agac
12R9_1_776-795	ctgaatgatccggtagc gggt
12R10_1_743-762	gaccgctggcggatct ccat
12R11_1_730-749	tctccattgcggacagt gat
12R12_1_696-717	gtgaagggcacgcga atggct
12R13_1_656-675	gatgagccggatgtct cacc
12R14_1_640-659	caccgctcgaattca ggga
12R15_1_630-651	gaccaaaggccgagg

622-641	atctg
12R16_1_595-615	cgcgtctttggcgaga ccgtg
12R17_1_559-581	tctttcgctttggcgttta tcag
12R18_1_539-561	cagttcgaggggcgc gatctggt
12R19_1_501-520	ctcgaaacagggtggg ttcg
12R20_1_481-502	cgagatttatgtcgaag gcagc
12R21_1_445-465	ctctggcatcggtgttt gag
12R22_1_404-423	cgggcaggctgtccga acct
12R23_1_383-405	ctgatcgagcggatcg aaagtgg
12R24_1_365-384	ggcctgctgagttacg gcaa
12R25_1_308-327	ctgggcccgttctgcaa cac
12R26_1_236-255	gaaggcccgggtgcag cagat
12R27_1_170-189	ttctggccgctctatgc gaa
12R28_1_108-130	gcccgatcataccacc aatgtcg
12R29_1_31-53	agacaccggacggaa tgccct
12R30_1_16-35	gccctgcgaccgtctat gag
13F1_1_3-22	ggcagtcattacaccg aacc
13F2_1_3-254	gcaaaaccccgtttact ccgggt
13F3_1_96-116	ctgtcttgcctacagt ctt
13F4_1_125-144	tcgaggctgatgctgc gcat
13F5_1_174-195	ctcgtgtgagagacaa gttctg
13F6_1_191-210	ttctgatcaaaggccc gat
13F7_1_216-236	gcgactgatgaaaatg ctcag
13F8_1_241-260	cgcgatatggacaaga tgca
13F9_1_253-274	aagatgcaggcggatc agtgtt
13F10_1_277-296	tatgtgccgatcgttga cca
13F11_1_293-314	accaaaggcgggtat gctgaa
13F12_1_317-336	acccggtggcgatcaa actg
13F13_1_337-356	gcagatgaccattactg gct
13F14_1_349-369	tactggctttccgtcgct gat

13F15_1_364-384	gctgatggagatctttg gcaa
13F16_1_382-401	caatttgcctaggcat cgt
13F17_1_415-437	ctggatgtcgaagtttc cgaacc
13F18_1_434-453	aaccagacgtgtctcc gcta
13F19_1_474-495	agatgatttgatggcac gggta
13F20_1_491-513	gggtatttgagatgat gtcagg
13F21_1_563-584	ttgtcatgcccgatcc ggctg
13F22_1_579-598	cggctggtcgaacaa gggtg
13F23_1_612-631	tgtcgaaggagcggat tacg
13F24_1_652-671	ttgttcgcaaatgggcg cga
13F25_1_681-701	acgtgcgggttgcccc aatgg
13F26_1_707-726	aacggatcgagagtgg ccta
13F27_1_759-780	cactccttatgaatcg gtctt
13F28_1_787-806	ttttgcaattctccgctg ga
13F29_1_807-826	ttacatcggttaaggccg ctt
13F30_1_821-840	ccgcttggcggatca ggca
13F31_1_856-877	agacaaatccgtccgc tgacta
13F32_1_897-916	gtgcaccgaatcttggc ctt
13F33_1_911-930	ggcctttgatggctgag gggt
13F34_1_945-966	agtgccctctgctgcat attcg
13F35_1_964-985	tcgcccaggtttgaggt gaacg
13F36_1_1021-1040	cccgcacgcggatc gagtt
13F37_1_1044-1063	gactgaacacggatg cgat
13F38_1_1072-1093	gtacaagagaaattctg gcgct
13R1_1_1070-1091	cgccagaatttctctgt acga
13R2_1_1049-1071	gacggcccatcgcata ccgtgtt
13R3_1_1036-1056	accgtgttcagtcac actc
13R4_1_1008-1027	tgtcgggtgcccgaatg cgac
13R5_1_982-1001	accatccctatggcga cggt
13R6_1_982-1001	aactcgggcgaatatg

55-974	cagc
13R7_1_9 38-957	agcagaggccacttgc ccca
13R8_1_8 99-920	atcaaaggccaagatt cgggtgc
13R9_1_8 66-885	gccctcgatagtcagc ggac
13R10_1_ 854-873	cagcggacggatttgt ctgg
13R11_1_ 811-831	cgccaaagcggcctta ccgat
13R12_1_ 798-817	taccgatgtaatccagc gga
13R13_1_ 775-795	attgcaaaagcgctca agacc
13R14_1_ 738-760	tgttctcgcgtcgcata ccgac
13R15_1_ 704-723	gccactctcgatccgtt caa
13R16_1_ 680-699	attggggcaaccgcga cgta
13R17_1_ 655-674	aggctcgcgccatttgc gaa
13R18_1_ 628-647	tcccaaagcggcattc cgta
13R19_1_ 611-631	cgtaatccgctccttcg acat
13R20_1_ 596-616	cgacatagatctcgaa cccac
13R21_1_ 576-595	ctgtttcgaccagccg gat
13R22_1_ 563-583	agccggatcgggcgat gacaa
13R23_1_ 529-548	aaggggcaaccttctgta tcg
13R24_1_ 485-504	atctccaaataccgtg cca
13R25_1_ 417-438	tgttccggaacttcga catcc
13R26_1_ 392-412	ccagggaagcacga tgccta
13R27_1_ 358-378	aagatctccatcagcg acgga
13R28_1_ 344-363	gacggaaagccagta atggg
13R29_1_ 321-340	ctgccagtttgatcgcc acc
13R30_1_ 288-307	taccgccattttggtcaa cg
13R31_1_ 253-275	taacactgatccgcctg catctt
13R32_1_ 217-237	gctgagcattttcatca gtcg
13R33_1_ 196-215	agcgcatccgggccttt gat
13R34_1_ 174-195	cagaactgtctctcac acgag
13R35_1_ 101-123	actttcgaagactgtag gcagaa

13R36_1_ 45-64	cggcttcgacacccgg agta
13R37_1_ 32-51	cggagtaaacgggggtt ttgc
13R38_1_ 7-29	acgcgtcgggtcgggtg aatgac
14F1_1_2 2-41	ggccgccttggcacgt tcac
14F2_1_3 7-57	ttcacgagagcttgat tgac
14F3_1_5 2-71	attgaccgattcgcgga agg
14F4_1_6 8-88	aagggcattctcgaga ctgtc
14F5_1_9 4-114	taccggctgtcctggaa tgtc
14F6_1_1 17-136	ggaattcgtcgggtag ccag
14F7_1_1 30-151	tagccaggccaccagt tcggta
14F8_1_1 47-166	cggataccagttctgcc gaa
14F9_1_1 60-179	tgccgaagagacggg cacat
14F10_1_ 197-219	gtgccgagttccggcg aatacca
14F11_1_ 213-233	aataccaggcgaggt gatcc
14F12_1_ 264-285	ctaccagccagaagtc cgggtgc
14F13_1_ 280-299	cgggtcgtagtcttcga tgg
14F14_1_ 295-315	gatgggcttgccatcg agctt
14F15_1_ 322-341	caccagttgggtgctg aacg
14F16_1_ 348-367	gaccggccttgaccag ttcc
14F17_1_ 382-403	ggcttctttaccgatgta gtcc
14F18_1_ 429-449	ggccaagattgcattg gaagg
14F19_1_ 515-536	gccgggtcgatgacca tcaggt
14F20_1_ 532-552	caggtgtattgttcgcc gac
14F21_1_ 558-577	gaatggcgttcagac ggtc
14F22_1_ 613-633	gtagcccttttcgccgg agaa
14F23_1_ 684-703	ggccgtaatagggcac ttcg
14F24_1_ 704-723	cgtaccgcttcgccga ccag
14F25_1_ 728-747	gccagcaggtctcgg actt
14F26_1_ 740-760	tcggacttaggtccttg aacc
14F27_1_ 	cgaagcgcgatgccatg

798-819	attgac
14F28_1_ 831-850	agtaggccagggtccga atcg
14F29_1_ 912-933	ggccgtgttcgtcgca gagcac
14F30_1_ 935-954	acatagcgccccatcat gga
14F31_1_ 971-990	tcacgagtgcagacgt agtt
14F32_1_ 985-1004	gtagttgcagaaggcc tctg
14F33_1_ 1099-1118	atcctcggggcgcaca tagg
14F34_1_ 1124-1146	gggtgatacatgcgatt gtagac
14F35_1_ 1144-1163	gacggtggccttcag cagc
14F36_1_ 1169-1188	tcgacggagaggtgcc agta
14F37_1_ 1183-1202	ccagtagggagacttg cgaa
14F38_1_ 1199-1218	cgaacccgattggaga gcaa
14F39_1_ 1216-1235	caatagctgaatgcca gggt
14F40_1_ 1237-1258	gccactctgacggagg ttgacc
14F41_1_ 1279-1299	gtcgatggactgctgg aagtt
14F42_1_ 1322-1342	cactctcgtccttggt gat
14F43_1_ 1356-1375	gtcgaggagacgtatg cctt
14R1_1_1 349-1369	tacgtctctcgacacc cgct
14R2_1_1 330-1352	cgctcgaccatcagc caaggac
14R3_1_1 308-1327	ggagtgccgcgccatg acac
14R4_1_1 280-1300	caacttcagcagtcga tcga
14R5_1_1 231-1251	ctccgtcagagtggcg accct
14R6_1_1 202-1222	gctattgcttccaatcg ggg
14R7_1_1 190-1209	aatcgggttcgcaagtc tcc
14R8_1_1 172-1191	ccctactggcacctctc cgt
14R9_1_1 144-1163	gctgctggaagggcac cgtc
14R10_1_ 1097-1116	tatgtgcgccccgagg atgg
14R11_1_ 1014-1035	gagcggccattcgcg tcaagg
14R12_1_ 934-954	tccatgatggggcgct atgtg
14R13_1_ 919-939	tatgtggtgctctcgca cgaa

14R14_1_854-875	aggacgagttctggttc accat
14R15_1_827-848	attcggacctggcctac tggtt
14R16_1_740-760	ggttcaaggacctaagt ccga
14R17_1_693-712	agcgggtacgcgaagt ccct
14R18_1_637-656	acgtcttcatctcccag acc
14R19_1_611-633	ttctccggcgaaaagg gctacga
14R20_1_561-580	cgagaccgtctggaac gcca
14R21_1_518-537	aacctgatggtcatcgc acc
14R22_1_463-482	cctacgggcaggatct cgac
14R23_1_439-461	acgagaccaaacccctt ccaatgc
14R24_1_393-413	ccaaggaggcggaact acatcg
14R25_1_377-396	atcggtaaagaagccc tgga
14R26_1_359-378	gagcgcaccgcgggaa ctggt
14R27_1_289-308	atggcaagcccacga agac
14R28_1_275-296	tcgaagactacgcacc ggactt
14R29_1_234-253	cggcagcgaaccaat ggct
14R30_1_205-224	cgccctggtattcgccg gaa
14R31_1_186-207	gaactcggcaccaaca tcgcca
14R32_1_154-176	tgcccgctcttcggca gaactg
14R33_1_116-135	tggctacccgacgaatt cct
14R34_1_94-113	acattccaggacagcc ggta
14R35_1_80-99	ccggtagccgcgaca gtctg
14R36_1_68-87	acagtctgcgagatgc cctt
14R37_1_56-75	atgcccttcgcgaatc gggt
14R38_1_42-63	gaatcgggtcaatccaa gctctc
15F1_1_2-41	ggccgcttggcacgt tcac
15F2_1_3-7-57	ttcacgagagcttgat tgac
15F3_1_5-2-71	attgaccgattcgcgga agg
15F4_1_6-8-88	aagggcacatctgcaga ctgtc
15F5_1_9	taccggctgtcctggaa

4-114	tgctc
15F6_1_1-17-136	ggaattcgtcgggtag ccag
15F7_1_1-30-151	tagccaggccaccagt tcggta
15F8_1_1-47-166	cggtaccagttctgcc gaa
15F9_1_1-60-179	tgccgaagagacggg cacat
15F10_1_197-219	gtgccgagttccggcg aatacca
15F11_1_213-233	aatacaggcgaggt gatcc
15F12_1_264-285	ctaccagccagaagtc cggtgc
15F13_1_280-299	cggtgcgtagtctcga tgg
15F14_1_295-315	gatgggcttccatcg agctt
15F15_1_322-341	caccagttggtggctg aacg
15F16_1_348-367	gaccggccttgaccag ttcc
15F17_1_382-403	ggcttcttaccgatga gtcc
15F18_1_429-449	ggccaagattgcattg gaagg
15F19_1_515-536	gccggtcgatgacca tcaggt
15F20_1_532-552	caggtgtattgttcgcc gac
15F21_1_558-577	gaatggcgttccagac ggctc
15F22_1_613-633	gtagcccttttcgccg agaa
15F23_1_684-703	ggccgtaatagggcac ttcg
15F24_1_704-723	cgtaccgcttcgccga ccag
15F25_1_728-747	gccagcaggctcctcg actt
15F26_1_740-760	tcggacttaggtccttg aacc
15F27_1_798-819	cgaagcgcagccatg attgac
15F28_1_831-850	agtaggccaggtccga atcg
15F29_1_912-933	ggccgtgttcgtcgca gagcac
15F30_1_935-954	acatagcggcccatcat gga
15F31_1_971-990	tcacgagtcagacgt agtt
15F32_1_985-1004	gtagttgcagaaggcc tctg
15F33_1_1099-1118	atcctcggggcgacaca tagg
15F34_1_1124-1146	gggtgatacatgcgatt gtagac

15F35_1_1144-1163	gacgggtggccttcag cagc
15F36_1_1169-1188	tcgacggagaggtgcc agta
15F37_1_1183-1202	ccagtaggagagacttg cgaa
15F38_1_1199-1218	cgaacccgattggaga gcaa
15F39_1_1216-1235	caatagctgaatgcca gggt
15F40_1_1237-1258	gccactctgacggagg ttgacc
15F41_1_1279-1299	gtcgtaggactgctgg aagtt
15F42_1_1322-1342	cactcctcgtccttggt gat
15F43_1_1356-1375	gtcaggagacgtatg cctt
15R1_1_1-349-1369	tacgtctctcgacacc cgct
15R2_1_1-330-1352	cgctcggaccatcagc caaggac
15R3_1_1-308-1327	ggagtgcgcgccatg acac
15R4_1_1-280-1300	caacttcagcagtcga tcga
15R5_1_1-231-1251	ctccgtcagagtgccg accct
15R6_1_1-202-1222	gctattgcttccaatcg gggt
15R7_1_1-190-1209	aatcgggttcgcaagtc tcc
15R8_1_1-172-1191	ccctactggcacctctc cgt
15R9_1_1-144-1163	gctgtggaaggccac cgctc
15R10_1_1097-1116	tatgtgcgccccgagg atgg
15R11_1_1014-1035	gagcggcccatcgcg tcaagg
15R12_1_938-960	gtgccatccatgatgg ggcgcta
15R13_1_919-939	tatgtggtgcttgcga cgaa
15R14_1_854-875	aggacgagttctggttc accat
15R15_1_827-848	attcggacctggcctac tggtt
15R16_1_740-760	ggttcaaggacctaagt ccga
15R17_1_693-712	agcgggtacgcgaagt ccct
15R18_1_637-656	acgtcttcatctcccag acc
15R19_1_611-633	ttctccggcgaaaagg gctacga
15R20_1_561-580	cgagaccgtctggaac gcca
15R21_1_518-537	aacctgatggtcatcgc acc

518-537	acc
15R22_1_463-482	cctacgggagggatctcgac
15R23_1_439-461	acgagaccaaccccttccaatgc
15R24_1_414-434	ttggccacatggtgccc aagg
15R25_1_393-413	ccaaggaggcggaactacatcg
15R26_1_377-396	atcggtaaagaagccc tgga
15R27_1_359-378	gagcgcacccgggaa ctggt
15R28_1_289-308	atggcaagcccacga agac
15R29_1_275-296	tcgaagactacgcacc ggactt
15R30_1_234-253	cggcagcgaacccaat ggct
15R31_1_205-224	cgccctggtattcgccg gaa
15R32_1_186-207	gaactcggcaccaaca tcgcca
15R33_1_154-176	tgcccgctcttcggca gaactg
15R34_1_116-135	tggctacccgacgaatt cct
15R35_1_94-113	acattccaggacagcc ggta
15R36_1_80-99	ccggtagccgcgaca gtctg
15R37_1_68-87	acagtctgcgagatgc cctt
15R38_1_56-75	atgccccccgcgaatc ggt
15R39_1_42-63	gaatcggtaaatccaa gctctc
16F1_1_68-87	gggtcaagggtatac cgtc
16F2_1_91-110	aatcacatgttgctgcc cac
16F3_1_128-149	aggccgattgcgcca tctgaa
16F4_1_232-251	atcagcccgcgcgaca tgga
16F5_1_255-274	gatggccgatgaccag tgtt
16F6_1_278-297	acgtgcccacggtcga tcat
16F7_1_341-360	ccgatcattactggctg tcg
16F8_1_369-388	tggcgacctgctgcaat tcg
16F9_1_4	gagatcgtcgaacccg

24-446	atgtctc
16F10_1_464-483	gaccagggccgacg atctg
16F11_1_481-503	ctgatggcgcggtctt tggcga
16F12_1_503-524	aggcgggtgcgcgatat ccgctt
16F13_1_540-559	gctggccttcaggga gtcg
16F14_1_574-594	cgctcaggctggtcga aacag
16F15_1_615-634	cgagggttcggaactg ggca
16F16_1_634-656	atgccgctgtggaacg cgctgtt
16F17_1_661-680	gccgggtcggaccta acgt
16F18_1_694-713	cccaacaatatcgagc gcgt
16F19_1_722-742	gggtgttgagctatggc aacg
16F20_1_752-771	gcgagaacacgccgt atgaa
16F21_1_784-803	aagttctgcaattcgcc cga
16F22_1_801-822	cgaggactatatcggc aaggca
16F23_1_816-837	caaggcagcactggc cgaacag
16F24_1_859-880	cagatccgggcactgg tgatcg
16F25_1_879-898	cggtggcgagattccg ccct
16F26_1_951-972	gtcagcgatccattccc ctgaa
16F27_1_997-1017	atggtggatcgcgacc attgg
16F28_1_1025-1047	gcaccggggtggaagt ggaaacg
16F29_1_1068-1088	tacgggtgcgcgaggg gttctg
16R1_1_1037-1058	atgccgtcgggcgttc cactt
16R2_1_999-1020	cgcccaatggctgcga tccacc
16R3_1_979-998	atgccgatcgcgacatt cac
16R4_1_967-986	acattcacgccgaattc agg
16R5_1_950-972	ttcaggggaatggatc gctgacc
16R6_1_899-918	cagcggccaggcatc ctgac
16R7_1_8	tgacagggcggaatct

82-902	cgcca
16R8_1_831-852	cggtccgttcttggcct gttcg
16R9_1_805-824	gctgccttgccgatata gtc
16R10_1_789-808	agtctcggggcgaatt gcag
16R11_1_759-780	caggccgcattcatac ggctg
16R12_1_730-751	gggtcatgtcgttgcca tagct
16R13_1_691-710	cgctcgatattgtggg gca
16R14_1_642-661	cggcaaacagcgcgtt ccac
16R15_1_611-630	cagttccgaacctcg acat
16R16_1_576-595	cctgtttcagaccgcct gag
16R17_1_556-575	cgcgccaccacaagct cgac
16R18_1_526-547	aggccagccgcttgta gcggaa
16R19_1_507-526	aaaagcggatatcgcg cacc
16R20_1_476-497	aagacccgcgccatca gatcgt
16R21_1_435-454	ccagcggcgagacat cgggt
16R22_1_378-398	atccccagcccgaatt gcagc
16R23_1_344-363	cagcgacagccagtaa tgat
16R24_1_328-347	tgatcggcgccagttt cac
16R25_1_285-304	cgccacgatgatcgac cgtg
16R26_1_261-283	gcacgtagtaacctg gtcacg
16R27_1_190-209	tcgggcccttgatgc tgac
16R28_1_152-171	gtcccagacctgcaca tgtt
16R29_1_135-154	gttccttcagatgggcg caa
16R30_1_108-127	gcaggctgtcgaacac cgtg
16R31_1_92-111	cgtgggcagcaacatg tgat
16R32_1_74-96	gtgattgtagacggat agccct

ANEXO II – LISTA DOS MELHORES INICIADORES ESPECÍFICOS TESTADOS *in silico*

2033984-2035078 ruegeria pomeroyi dss-3, complete genome			
1F3 128-149	aggccgattgcgcccattctgaa	1R18 526-547	aggccagccgctttagcgga
16F3 128-149	aggccgattgcgcccattctgaa	16R18 526-547	aggccagccgctttagcgga
2182262-2183365 roseobacter denitrificans och 114 chromosome, complete genom			
2F12 487-506	gcccgtatcttggcgacag	2R5 954-974	ctccaggcagcggaagtgacc
194538-195680 candidatus pelagibacter ubique htcc1016 a3qgdraft_scaffold_1.2, whole genome shotgun sequence			
3F3 178-198	gatgttgagtgagcgacaa	3R3 739-758	tcagcatctgcatggtacga
81660-82583 sar324 cluster bacterium scgc aaa240-j09 aaa240-j09_cont465, whole genome shotgun sequence			
4F17 571-590	gtggcaagatccggttgag	4R 803-822	aaggcttcaatgtctgcgt
37339-38433 phaeobacter gallaeciensis ang1 contig175, whole genome shotgun sequence			
5F3 113-133	tgtttgacagcctggaatacg	5R2 1028-1049	ggtgtctcaactccaacgcag
2452182-2453288 dinoroseobacter shibae dfl 12 chromosome, complete genome			
6F2 98-117	atgccgatggcaacattcgt	6R2 868-888	aagttgatgcagcttctgacc
2383058-2384161 jannaschia sp. ccs1 chromosome, complete genome			
7F10 365-384	ttgccgtagctgagcaatcc	7R13 730-749	tctccatcgcggaacagtgat
1542571-1543728 ruegeria sp. tm1040 chromosome, complete genome			
8F22 553-572	ttcgaactggtagacgcaa	8R2 1105-1124	ggtttgccatgaccgttctg
36281-37900 alpha proteobacterium bal199 1103548001519, whole genome shotgun sequence			
9F20 722-741	acttgccaccggtcaggat	9R3 1536-1558	tattctcgccaacgccgtcaagg
10F20 722-741	acttgccaccggtcaggat	10R19 865-884	cgctgctcatcgtggccgag
13113-14168 sar324 cluster bacterium scgc aaa001-c10 a1c10_cont214, whole genome shotgun sequence			
11F13 571-590	gtggcaagatccggttgag	11R5 775-797	ctgacatactttcaagaccgca
139-1254 ruegeria mobilis f1926 f1926_contig_3, whole genome shotgun sequence			
12F21 728-747	agatcactgtccgaatgga	12R4 964-983	gcgttgaggcggattatcac
500200-501294 ruegeria sp. tw15 contig00006, whole genome shotgun sequence			
13F16 382-401	caatttgccctaggcatcgt	13R12 798-817	taccgatgtaatccagcgga
773394-774821 halomonas elongata dsm 2581, complete genome			
14F10 197-219	gtgccgagttccggcgaatacca	14R20 561-580	cgagaccgtctggaacgcca
15F21 558-577	aaatggcggtccagacggtc	15R6 1202-1222	gctattgctctcaatcggt

ANEXO III – LISTA DOS INICIADORES DEGENERADOS
(MELHORES EM NEGRITO)

1aF 5'	GCTGGGCATACACCGTGtayaaymrnat	3'
2aF 5'	GCTGGGCATACACCGTGtayaaymrnatg	3'
1bF 5'	CCTGAAGAAGCACGTGCAGntntggraygt	3'
1cF 5'	CGAGAACGGCGGCATGntnaaygaycc	3'
2cF 5'	CGGCGGCATGCTGaaygayccnrt	3'
1dF 5'	GTGTGGCCCATGGCCrtnccarggncc	3'
1eF 5'	TGGTCCAAGCAGGGAggntwygarrrt	3'
2eF 5'	GGTCCAAGCAGGGAGGCTwygarrrnta	3'
3eF 5'	GTCCAAGCAGGGAGGCTTygarrrntayg	3'
1fF 5'	TGGACCACGAGAACCAACccnywysartg	3'
1aR 5'	GAACatnykrtrta	3'
2aR 5'	GArcatnykrtrt	3'
3aR 5'	Gwrcatnykrtrt	3'
1bR 5'	ACCTGCCGCTCGCAGshnacrttycca	3'
1cR 5'	GGCCAGCTTCACCACCAynggrtcrttna	3'
2cR 5'	GGCCAGCTTCACCACCAynggrtcrtt	3'
1eR 5'	TCCGAGTCGTGCACGTAGayytcwanc	3'
2eR 5'	GCTCCGAGTCGTGCACGtanayytcw	3'
3eR 5'	TTGAGCTCCGAGTCGTGCrcrtanayytc	3'
1fR 5'	GGCCCAGGCCGcaytsrwrngg	3'