

**UNIVERSIDADE ESTADUAL DE SANTA CRUZ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E
BIOLOGIA MOLECULAR**



**APLICAÇÃO DA METAGENÔMICA DE SOLOS VISANDO À
IDENTIFICAÇÃO DE NOVOS AGENTES ANTIMICROBIANOS.**

JAIME HENRIQUE AMORIM SANTOS

ILHÉUS – BAHIA – BRASIL

Fevereiro de 2008

JAIME HENRIQUE AMORIM SANTOS

APLICAÇÃO DA METAGENÔMICA DE SOLOS VISANDO À
IDENTIFICAÇÃO DE NOVOS AGENTES ANTIMICROBIANOS

Dissertação apresentada à
Universidade Estadual de Santa Cruz,
como parte das exigências para obtenção
do título de Mestre em Genética e
Biologia Molecular.

ILHÉUS – BAHIA – BRASIL

Fevereiro de 2008

JAIME HENRIQUE AMORIM SANTOS

APLICAÇÃO DA METAGENÔMICA DE SOLOS VISANDO À
IDENTIFICAÇÃO DE NOVOS AGENTES ANTIMICROBIANOS

Dissertação apresentada à
Universidade Estadual de Santa Cruz,
como parte das exigências para obtenção
do título de Mestre em Genética e
Biologia Molecular.

Aprovada:

Dr. Marcos Mota do Carmo Costa
CENARGEN/EMBRAPA

Prof. Dr. Martin Brendel
UESC

Prof^a Dr^a Aline Silva
UESC

Prof. Dr. Júlio César de Mattos
Cascardo
UESC – Orientador

Dedicatória

Dedico este trabalho e tudo alcançado na minha
vida aos meus pais Jaime e Elza.

AGRADECIMENTOS

A todos os meus familiares, em especial aos meus pais, por estarem sempre ao meu lado apoiando e mostrando-me o melhor caminho a ser seguido, nunca medindo esforços para o meu sucesso pessoal, assim como, às minhas irmãs, pelo apoio e por suportar os meus dias difíceis.

À minha noiva Paloma, por todo apoio emocional.

À Universidade Estadual de Santa Cruz (UESC) e à Coordenação do Programa de Pós Graduação em Genética e Biologia Molecular, pela oportunidade da realização do curso e pelo apoio logístico.

À Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB), pela concessão da bolsa.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), projeto PADCT, pelo apoio financeiro parcial para a compra de materiais e reagentes, assim como, apresentação dos resultados no Congresso Brasileiro de Genética.

À RENORBIO (Rede Nordeste de Biotecnologia), pelo suporte financeiro parcial para a compra de equipamentos, materiais e reagentes utilizados nos trabalhos moleculares.

Desejo expressar minha homenagem e admiração aos cientistas que contribuíram de forma direta ao êxito deste trabalho:

- ao Prof. Dr. Júlio Cezar de Mattos Cascardo, por depositar inteira confiança na minha pessoa; pela orientação segura no decorrer do trabalho, atenção constante e amizade. Uma pessoa admirável. Essencial no meu aprendizado em Biotecnologia e na minha formação científica.

- ao Prof. Dr. João Carlos Teixeira Dias, professor e amigo, por todo profissionalismo e dedicação fundamentais ao sucesso desse trabalho. Essencial no meu aprendizado sobre Microbiologia Ambiental.
- Aos Professores Dr. Martin Brendel, Dr^a Rachel Passos Rezende e Ms. Carlos Priminho Pirovani, pela disposição e atenção, como também, pelas sugestões e críticas construtivas que contribuíram na minha formação científica.

À Tharcilla, colega de profissão e amiga, pela dedicação ao me ensinar os primeiros passos nas técnicas de Biologia Molecular, pelas sugestões e parceria.

Ao Robson (Robinho), pela inestimável contribuição no processo de seqüenciamento do DNA neste trabalho.

À professora Fernanda Gaiotto, pela solicitude em me ceder os primers que contribuíram para o sucesso deste trabalho.

Ao Dr. Marcos Costa e ao Ramon Vidal, pela fantástica contribuição na parte de bioinformática.

Ao Cristiano, pela amizade e contribuição científica em interessantes discussões envolvendo genômica.

À Cristiane, técnica do Laboratório de monitoramento ambiental, pelo apoio constante no laboratório de Monitoramento Ambiental.

A todos os colegas do mestrado e laboratório, pela troca de experiências, ajuda, convívio e descontração.

À secretaria do Programa de Pós-Graduação (Luciana), pela eficiência e disponibilidade.

Espero não ter esquecido de ninguém, contudo saibam que expressei minha gratidão a todos direta e indiretamente ligados ao desenvolvimento e êxito deste trabalho. Mesmo que seu nome não conste aqui, saiba que sou muito grato também.

ÍNDICE

LISTA DE FIGURAS	VIII
EXTRATO	X
ABSTRACT	XI
1. INTRODUÇÃO	12
2. REVISÃO BIBLIOGRÁFICA	14
3. CAPÍTULO 1- UM NOVO E EFICIENTE MÉTODO PARA EXTRAÇÃO DE DNA METAGENÔMICO DE SOLO DE MATA ATLÂNTICA	21
Resumo	21
3.1. Introdução	22
3.2. Material e métodos	24
3.2.1. Coleta dos Solos	24
3.2.2. Análise dos Solos.....	24
3.2.3. Métodos de extração ex situ de DNA utilizados	24
3.2.4. Capacidade de amplificação via PCR das formas de purificação do método B.....	26
3.2.5. Comparação do rendimento de DNA	27
3.2.6. Capacidade de digestão por <i>EcoRI</i> e <i>BamHI</i>	28
3.3. Resultados e discussão.....	28
3.4. Conclusão	32
3.5. Referências bibliográficas	33
4. CAPÍTULO 2 – UMA RÁPIDA ANÁLISE DO PERFIL TAXONÔMICO E METABÓLICO DE METAGENOMAS DE SOLOS	36
Resumo	36
Introdução	37

Metodologia	40
Resultados e discussão	46
Conclusão	61
Referências	62
5. CAPÍTULO 3- METAGENOMA DE MATA ATLÂNTICA COMO FONTE DE SUBSTÂNCIAS ANTIMICROBIANA.	68
Resumo	68
Introdução.....	69
Metodologia.....	70
Resultados e Discussão.....	72
Referências	81
6. CONCLUSÕES GERAIS	84
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	85
8. APÊNDICE.....	93

Lista de figuras

Figura 3.1- Amplificação do gene 16S de DNAs provenientes de três metodologias de purificação no método A de extração	29
Figura 3.2-Digestão do eDNAs oriundos dos métodos A e B com enzimas de restrição ...	32
Figura 4.1 - Otimização da amplificação do eDNA de lanfarm utilizando-se primers randomicos	47
Figura 4.2- RAPD metagenômico com os primers D13, OPC01, OPK08, OPB13, 7OPS0, OPS10, T17, Q08, OPJ04 e T20.....	48
Figura 4.3- RAPD metagenômico com os primers D13, OPC01, OPI07, OPN18, OPN19, OPK20, OPB13, OPS07, OPS10 e Q08.....	49
Figura 4.4- Diagnóstico de clonagem de DNA metagenômico de MA em pTZ57R/T através de PCR de colônia.....	50
Figura 4.5- Comparação da representatividade de grupos taxonômicos entre Land1 (435 reads) e Land2 (570 reads)..	52
Figura 4.6- Comparação da representatividade em nível de função biológica entre as bibliotecas Land1 e Land2.....	53
Figura 4.7- Comparação da diversidade taxonômica entre os solos de landfarm e Mata Atlântica (Land2 vs Mata1).....	55
Figura 4.8- Comparação da diversidade entre Mata (Mata Total) e Land2.	57
Figura 4.9- Comparação da diversidade metabólica e funcional entre os solos de landfarm e Mata Atlântica.	58
Figura 4.10- Diversidade taxonômica geral do solo de Mata Atlântica.....	59
Figura 4.11- Diversidade taxonômica geral do solo de landfarm.....	60
Figura 5.1- Bioensaio do clone B3:MTII PL08 contra <i>Bacillus subtilis</i> ATCC.....	73
Figura 5.2- Bioensaio do clone B3:MTII PL08 contra <i>Staphylococcus aureus</i> ATCC.....	73
Figura 5.3- Bioensaio de clones contra <i>E. coli</i> CCT.....	74

Figura 5.4- Eletroforese do produto de PCR de colônia do clone Mata II PL08 B3 utilizando o primer M13	74
Figura 5.5- Contigs F (1) e R (1) e as ORFs com homologia com peptídeo sinal e glutamato sintetase, respectivamente.	76
Figura 5.6- Busca por ORFs dentro do consenso F, com 793 pares de bases do clone MTII PL08 B3.....	77
Figura 5.7- Busca por ORFs dentro do consenso R, com 515 pares de base do clone MTII PL08 B3.....	78
Figura 5.8- Bioensaio contra <i>B. subtilis</i> para verificar indução de expressão do antimicrobiano por IPTG e a faixa de seu tamanho em KDa.....	79
Figura 5.9- SDS-PAGE para avaliação de expressão diferencial de proteínas com indução por IPTG.....	80

EXTRATO

AMORIM, Jaime Henrique, M.S., Universidade Estadual de Santa Cruz, Ilhéus, fevereiro de 2007. **APLICAÇÃO DA METAGNÔMICA DE SOLOS VISANDO À IDENTIFICAÇÃO DE NOVOS AGENTES ANTIMICROBIANOS**. Orientador: Júlio Cézar de Mattos Cascardo. Co-orientador: João Carlos Teixeira Dias.

Um grama de solo de floresta contém, em estimativa, 4×10^7 células procarióticas. Baseado na reassociação cinética de DNA isolado de vários solos, o número de genomas procariotos distintos tem sido estimado no intervalo de 2.000 a 18.000 por grama de solo, tornando este um dos mais desafiadores ambientes naturais para os microbiologistas. Partindo-se da premissa de que menos de 1% dos microorganismos presentes nos solos são cultiváveis, a metagenômica, uma análise genômica independente de cultivo aparece como a melhor ferramenta para entender e explorar toda essa diversidade. O termo metagenômica é derivado do conceito estatístico de meta-análise (o processo de combinar estatisticamente análises distintas) e genômica (a análise compreensiva do material genético de um organismo). Essa tecnologia possibilita acessar informações em nível de filogenia como também isolar e clonar novos genes responsáveis pelas mais diversas funções. Entretanto, alguns aspectos em metagenômica convencional limitam o potencial de estudos filogenéticos e exploração biotecnológica dos solos, tais como dificuldades encontradas nos processos de clonagem em função da co-extração de substâncias interferentes juntamente com o DNA metagenômico e perda de eficiência na representatividade genética e funcional induzido pelas metodologias de extração de DNA utilizadas até o momento. Mas neste trabalho é apresentada uma alternativa para diminuir consideravelmente a perda de eficiência na representatividade bem como facilitar o processo de clonagem de DNA de solo, baseada numa metodologia simplificada de extração direta e na clonagem de seqüências metagenômicas amplificadas ao acaso através de reações de RAPD em vetores T/A. Desta forma, foi possível representar de forma mais ampla e fiel a microbiota contida nas amostras de solo. Além disso, o potencial biotecnológico desta técnica foi comprovado através da identificação de um clone com atividade antimicrobiana contra *B. subtilis* e *S. aureus*.

Palavras-chave: Metagenômica de solos, RAPD, antimicrobiano, bioinformática

ABSTRACT

AMORIM, Jaime Henrique, M.S., Universidade Estadual de Santa Cruz, Ilhéus, fevereiro de 2007. SOIL METAGENOMICS AIMING IDENTIFICATION OF NEW COMPOUNDS WITH ANTIMICROBIAL ACTIVITY Advisor: Júlio César de Mattos Cascardo. Co-advisor: João Carlos Teixeira Dias.

One gram of soil forest contains, in estimate, 4×10^7 procariotic cells. Based on kinetics reassociation of DNA isolated from various soils, the number of distinct procariotic genomes has been estimated in the range of 2.000 to 18.000 per gram of soil, making this one of the most challenging natural environments for microbiologists. Based on the premise that less than 1% of microorganisms present in soil are cultivable, metagenomics, an independent genomic analysis of cultivation, appears as a tool to better understanding and exploiting all this diversity. The term metagenomics is derived from the concept of statistical meta-analysis (the process of combining different statistical analyses) and genomics (a comprehensive analysis of the genetic material of an organism). This technology allows access to information on phylogenetic and functional levels, by isolating and cloning new genes responsible for many different functions. However, some conventional metagenomic aspects have limited the phylogenetic studies potential and soil biotechnological exploitation, such as difficulties when cloning, according to contaminant co-extraction with metagenomic DNA and representativeness bias in genetics and functional diversity induced by DNA extraction protocols used until now. Here, we present an alternative to considerably reduce representation bias and facilitate the soil DNA cloning process, based on a simplified method of direct extraction and the RAPD metagenomic amplified sequences cloning in a T/A vector. Thus, it was possible to represent in a more broad and reliable way the microbiota contained in soil samples. Furthermore, the biotechnological potential of this technique has been proven through the identification of a clone with antimicrobial activity against *B. subtilis* and *S. aureus*.

Key words: Soil metagenomic, RAPD, antimicrobial, bioinformatics.

1. INTRODUÇÃO

Menos de 5% da microbiota dos solos é cultivável. E a metagenômica de solos tem sido utilizada para entender e determinar o perfil da microbiota destes ambientes, como também para explorar seu potencial biotecnológico de maneira cada vez mais ampla, visto que se trata de uma análise independente de cultivo, onde o DNA de microrganismos presentes no solo é clonado em hospedeiros bem caracterizados, como a *Escherichia coli* ou é utilizado como molde para amplificar, via PCR (Polymerase chain reaction), seqüências específicas, as servem de base para investigações filogenéticas (VIGDIS & OVREAS, 2002; SCHLOSS & HANDELSMAN, 2003).

As técnicas utilizadas para acessar a diversidade do solo em níveis genético e funcional têm sido intensamente desenvolvidas com atenção especial na representatividade da microbiota dos solos, pois para que o ambiente seja estudado de maneira eficiente é necessário representar de forma também eficiente os grupos taxonômicos ali existentes (DANIEL, 2004). Isso permite traçar perfis filogenéticos com grande confiabilidade como também explorar de maneira mais ampla o potencial de identificar novas substâncias, tais como novos fármacos e enzimas. Entretanto, a co-extração de contaminantes juntamente com o DNA metagenômico de solo a ser clonado tende a inibir reações enzimáticas que são essenciais no próprio processo de clonagem, diminuindo as chances de êxito. Pois as técnicas utilizadas até então para eliminar essa limitação são baseadas em processos que diminuem a representatividade ou mesmo deixam de representar grupos taxonômicos pertencentes à microbiota do solo em questão.

O presente trabalho visou contribuir para uma evolução na representatividade de microbiotas de solos em bibliotecas metagenômicas bem como identificar novos agentes antimicrobianos em metagenomas de solos. E os objetivos específicos foram:

1. Desenvolver uma técnica de extração de DNA metagenômico de solos onde a exposição da população de células à pressão de lise fosse a mais homogênea possível;

2. Desenvolver uma técnica para clonar DNA metagenômico totalmente livre de substâncias interferentes;
3. Desenvolver tecnologia para acessar o perfil taxonômico e metabólico de solos através de técnicas de extração, clonagem e ferramentas de bioinformática;
4. Realizar buscas por atividades enzimática e antimicrobiana nas bibliotecas metagenômicas construídas com a tecnologia desenvolvida.

2. REVISÃO BIBLIOGRÁFICA

A biosfera é dominada por microorganismos, embora a maioria dos presentes na natureza ainda não tenha sido estudada (RONDON et al., 2000). O solo aparece como um dos principais reservatórios da diversidade genética microbiana e pode ser considerado como um ambiente complexo (CURTIS et al., 2002). Essa extrema complexidade resulta de múltiplos parâmetros relacionados, incluindo textura e estrutura do solo, umidade, pH, variações climáticas e atividade biótica.

Partindo-se da premissa de que menos de 1% dos microorganismos presentes nos ambientes são cultiváveis (YUN et al., 2004), a metagenômica, uma análise genômica independente de cultivo aparece como a melhor ferramenta para entender e explorar toda essa diversidade. O termo metagenômica é derivado do conceito estatístico de meta-análise (o processo de combinar estatisticamente análises distintas) e genômica (a análise abrangente do material genético de um organismo (SCHLOSS & HANDELSMAN, 2003). Essa tecnologia possibilita acessar informações em nível de filogenia e o isolamento e clonagem de novos genes da comunidade microbiana responsáveis pelas mais diversas funções (STREIT & SCHMITZ, 2004).

Todo esse universo a ser compreendido e explorado torna a metagenômica uma linha de pesquisa com enorme potencial científico e tecnológico, na medida em que cada vez mais ambientes são estudados com aplicação da metagenômica, tendo como objetivo, por exemplo, traçar o perfil microbiano de um ambiente. Neste empreendimento, pesquisadores têm obtido resultados que corroboram cada vez mais o fato de que cerca de 95% dos microorganismos do solo não são cultiváveis. A base desse tipo de análise está na amplificação de seqüências específicas de DNA, como o DNAr 16S, direto da amostra ambiental (PACE et al. 1985;

SCHLOSS & HANDELSMAN, 2003). Como essas seqüências são homólogas, é possível estudar perfis filogenéticos e até mesmo traçar distâncias filogenéticas entre os grupos taxonômicos que compõem a microbiota do ambiente em questão. Este tipo de análise é possível através da utilização de ferramentas de bioinformática como BLAST, DOTUR, Bellerophon e “Check_Chimera at the RPD”. A primeira compara as seqüências geradas com as depositadas no National Center for Biotechnology e serve de base para a utilização de várias outras ferramentas. A segunda é utilizada para estimar a diversidade microbiana da amostra e determinar unidades taxonômicas operacionais e a terceira e a quarta são utilizadas para remover seqüências quiméricas da seqüência gerada com o objetivo de aumentar a confiabilidade da análise a ser feita em seguida (ALTSCHUL et al., 1990; MAIDAK et al., 1996; HUBER et al., 2004; SCHLOSS & Handelsman, 2005). Também podemos citar o J-LIBSHUFF, desenvolvido para aumentar a acurácia da comparação de bibliotecas 16S, contribuindo para estudos onde se busca determinar diferenças na composição da microbiota de ambientes distintos, os quais exigem seqüências de alta qualidade (SCHLOSS et al., 2004; PAPAS et al., 2005). Mas já existem alternativas para determinar a diversidade de uma amostra ambiental sem a utilização de seqüências específicas. Exemplificando, podemos citar o programa MEGAN, onde é possível computar e explorar o conteúdo taxonômico de uma amostra ambiental a partir de seqüências não específicas (HUSON et al., 2007). Primeiramente, essas seqüências devem ser submetidas ao algoritmo BLAST ou outra ferramenta de comparação que utilize a taxonomia do NCBI; e o resultado dessas comparações servirá de entrada para a análise no Megan, o qual é utilizado para computar e explorar o conteúdo taxonômico do banco de seqüências (HUSON et al., 2007).

Entre as análises de bioinformática também podemos citar as que predizem genes ou ORFs dentro das seqüências geradas através de algoritmos específicos. A maioria desses é baseada nos modelos de Hidden Markov e são capazes de agrupar seqüências por função biológica ou taxonomia (MENCONI & MARANGONI, 2006).

No que diz respeito às aplicações biotecnológicas da metagenômica, novas substâncias com aplicações industriais e farmacológicas são as mais procuradas (DUBEY et al., 2006). Novas enzimas que possam agilizar e/ou baratear processos industriais têm sido procuradas intensivamente aplicando-se a metagenômica (KIM et al., 2005). Novos agentes antitumorais e antimicrobianos também têm ocupado destaque entre os fármacos procurados utilizando-se a

metagenômica. E a seleção de novos antimicrobianos desperta especial interesse por estar relacionada com uma problemática de saúde pública em nível global; a emergência, cada vez mais freqüente de cepas de microorganismos patogênicos resistentes a antibióticos. Um problema grave, em que os antimicrobianos disponíveis atualmente vão se tornando ineficientes na medida em que servem como seletores de cepas que expressam genes específicos de resistência. Essa situação é agravada quando pacientes submetidos à antibioticoterapia interrompem o tratamento, aumentando a probabilidade de êxito dos microorganismos resistentes em sobreviverem e agravarem a infecção, dificultando um segundo tratamento. E como se isso não bastasse, emergem os chamados *superbugs*, microorganismos resistentes a múltiplos antibióticos que compõem um dos maiores desafios da medicina moderna (WRIGHT, 2007). E a utilização da metagenômica para seleção de novos antimicrobianos pode ajudar nesse desafio.

Com o objetivo de promover a bioprospecção de novos genes, alguns sistemas têm sido utilizados para clonar e expressar DNA metagenômico. Entre os que são aptos a carregar grandes seqüências podemos destacar os que utilizam vetores BAC (bacterial artificial chromosome). Baseados em plasmídeos-F, são capazes de carregar seqüências de DNA de até 300 kb (SHIZUYA et al., 1992). E em metagenômica, quanto maior o tamanho da seqüência clonada, maior a probabilidade de se encontrar uma nova substância, visto que a clonagem de seqüências grandes permite a expressão tanto de produtos diretos dos genes (proteínas, enzimas e ribozimas) como produtos secundários, provenientes de vias metabólicas clonadas de maneira completa (RONDON et al., 2000).

Embora a clonagem de grandes fragmentos seja preferencial, algumas limitações na manipulação de sistemas envolvendo BACs, tais como poucas cópias do vetor por célula e dificuldade de ligação de grandes fragmentos no vetor, têm levado pesquisadores a buscar novas alternativas para clonagem de DNA metagenômico (LAMMLE et al., 2007). Desta forma, alguns trabalhos têm sido bem sucedidos utilizando vetores como o pUC19 e clonando fragmentos de até 8 kb (LAMMLE et al., 2007). O menor tamanho do inserto parece não ser suficiente para eliminar as chances de sucesso, pois novas enzimas entre amilases, celulasas e proteases foram encontradas em trabalhos recentes, reforçando o valor da alternativa (YUN et al., 2004). Além disso, a utilização de vetores sensíveis à indução de expressão por IPTG pode compensar a perda do promotor natural na seqüência clonada.

No caso da clonagem e expressão de genes provenientes de DNA metagenômico de solos, a situação é um pouco mais complicada, pois todas as reações enzimáticas envolvidas neste contexto sofrem interferência de substâncias co-extraídas juntamente com o DNA a ser clonado (ROBE et al., 2003). Seja em estudos filogenéticos ou aplicações biotecnológicas, as etapas de extração de DNA metagenômico determinam se o trabalho será ou não bem sucedido. O rDNA 16S, utilizado em estudos filogenéticos é amplificado através de uma reação enzimática; DNA metagenômico clonado para fins biotecnológicos geralmente é digerido por enzimas de restrição antes de ser ligado ao vetor, processo que também envolve uma reação enzimática. E se houverem substâncias interferentes co-extraídas com o DNA capazes de inibir todas essas reações, a clonagem do DNA metagenômico de solo torna-se inviável (ROBE et al., 2003; ROH et al., 2006).

Um dos principais contaminantes co-extraídos com DNA de solo são os ácidos húmicos, que se ligam ao DNA e impedem ou dificultam o trabalho de qualquer enzima que utilize o mesmo como substrato (ROBE et al., 2003). Os ácidos húmicos, são, portanto, capazes de inibir digestões, ampliações, hibridização DNA-DNA e até mesmo a transformação de células competentes. Mas não são os únicos interferentes co-extraídos juntamente com o DNA metagenômico. E a ligação ao DNA também não é o único mecanismo pelo qual reações enzimáticas necessárias no processo de clonagem são inibidas. Alguns compostos têm a capacidade de seqüestrar importantes cofatores das enzimas, os quais têm de estar em uma concentração específica para que a enzima trabalhe eficientemente e caso essa exigência não seja satisfeita, a reação em questão é inibida (WILSON, 1997).

Além dessas dificuldades impostas pelos contaminantes, relacionados às etapas de extração e purificação de DNA metagenômico, quando o objetivo do trabalho é bioprospecção de novos genes, é importante atentar para a representatividade da diversidade microbiana na amostra. Isso porque, logicamente, quanto maior a representatividade da amostra, maior a probabilidade de êxito. E a maioria dos métodos utilizados para extração de DNA metagenômico de solos apresenta deficiências no que diz respeito a manter representatividade da amostra (DANIEL, 2004; ROBE et al., 2003), uma vez que, na tentativa de livrar o DNA do contaminantes, realizam-se algumas etapas de purificação e/ou lavagens que resultam em perda de DNA, introduzindo uma ineficiência em nível de representatividade no trabalho.

Entre os métodos convencionais de extração, existem basicamente duas vertentes; a extração direta ou *in situ* e a extração indireta ou *ex situ*, ambas com diferentes métodos compartilhando o mesmo princípio.

Na extração direta, primeiro realiza-se a lise *in situ* dos microrganismos, os quais encontram-se aderidos às partículas de solo, liberando o DNA metagenômico que fica exposto aos componentes do solo (PARK et al., 2006). Depois realiza-se a separação do DNA das partículas do solo. Todos os métodos baseados nesse princípio derivam do trabalho de OGRAM e colaboradores, de 1987 (ROBE et al., 2003). Esse princípio proporciona maior rendimento de DNA em comparação com o princípio da extração *ex situ*. Neste último, primeiro as células são separadas das partículas de solo e depois sofrem lise para expor o DNA em um meio contendo menos contaminantes em potencial em relação à extração direta (ROH et al., 2006; ROBE et al., 2003).

A extração direta proporciona maior rendimento de DNA, mas a qualidade é questionável porque expõe o DNA a uma alta concentração de contaminantes. Se existe uma tentativa de purificar esse DNA, é introduzida uma interferência na representatividade da microbiota do ambiente em questão. E a extração indireta proporciona um DNA mais livre de contaminantes, mas as etapas de liberação das células das partículas de solo também introduzem interferência na representatividade (ROH et al., 2006; STREIT & SCHMITZ, 2004). Ainda que muitos trabalhos de metagenômica utilizando esses dois princípios de extração de DNA tenham sido bem sucedidos no sentido de identificarem novas substâncias, existe a possibilidade de evolução no que diz respeito à representatividade, pois é importante lembrar que menos de 1% da microbiota de solos é cultivável. Existem muitos microrganismos ainda desconhecidos, genes que não foram acessados, ou seja, existe ainda um universo a ser explorado e compreendido (KIM et al., 2005). Se não houver uma busca por evolução no princípio de representatividade em metagenômica todo esse universo será explorado de forma limitada e não será tão compreendido de maneira ampla (WHITEHOUSE & HOTTEL, 2006). Mas como manter representatividade da amostra sem co-extrair contaminantes que limitarão a clonagem do DNA metagenômico? Como unir o que há de melhor nos princípios de extração direta e indireta? A resposta pode estar em técnicas simples e rotineiramente utilizadas nos laboratórios de Biologia Molecular.

Uma simples fervura da amostra de solo pode promover uma lise homogênea das células ali contidas. Isso resultaria, em teoria, em uma extração com boa representatividade da microbiota. Mas como eliminar o problema das substâncias interferentes?

A técnica de RAPD (random amplified polymorphic DNA), permite a amplificação aleatória de seqüências de DNA através da utilização de primers não específicos, pois nessa técnica, utiliza-se apenas um oligonucleotídeo de 10 bases iniciador que possui seqüência arbitrária, e portanto sua seqüência alvo é desconhecida (ARAÚJO et al., 2003), ao contrário da PCR (polymerase chain reaction). A amplificação de DNA metagenômico via RAPD com vários primers diferentes em reações diferentes poderia permitir uma cobertura do metagenoma. Aliando essa idéia a uma simples fervura da amostra de solo, seria possível unir o que existe de melhor nos métodos de extração direta e indireta, visto que a extração baseada em uma simples fervura da amostra de solo aliada à amplificação via RAPD com vários primers diferentes manteria a representatividade da amostra. E a própria amplificação geraria seqüências sintetizadas *in vitro*, livres de substâncias interferentes.

Em reações de RAPD utiliza-se *Taq*-DNA polimerase, a qual possui atividade de exonuclease 5'→3'. Isso significa dizer que às seqüências amplificadas são adicionados “overhangs” dATP (desoxiadenosina-trifosfato). E existem vetores tipo T/A, como o ptz57R/T especialmente desenhados para serem ligados a esse tipo de produto de reação de amplificação (<http://www.fermentas.com/profiles/kits/pdf/instaclone1213.pdf>), o que significa dizer que a clonagem desses produtos de RAPD em vetores tipo T/A seria vantajosa do ponto de vista prático, pois seria a clonagem de seqüências metagenômicas sintetizadas *in vitro*, livres de contaminação e facilmente ligáveis a vetores comerciais como o ptz57R/T. Tudo isso possibilitaria a clonagem em larga escala de DNA metagenômico de alta qualidade com alta representatividade da microbiota ambiental, aumentando consideravelmente as chances de bioprospecção de novas substâncias.

Além do potencial de bioprospecção de novos genes, seria possível também traçar um perfil metabólico bastante interessante da microbiota, uma vez que a clonagem de seqüências aleatórias engloba genes das mais variadas famílias e funções (BERTRAND et al., 2005). Alguns programas de bioinformática como o Megan (HUSON et al., 2007) permitem agrupar seqüências clonadas por função biológica, como também estudar a diversidade biológica na amostra. Todas essas análises podem ser feitas baseadas em seqüências não específicas. Dessa

forma, seria possível, além da tentativa de bioprospecção de novos genes, determinar o perfil microbiano e metabólico de uma amostra ambiental através dessa nova metodologia em metagenômica.

3. Capítulo 1- UM NOVO E EFICIENTE MÉTODO PARA EXTRAÇÃO DE DNA METAGENÔMICO DE SOLO DE MATA ATLÂNTICA

AMORIM, JH¹; SILVA-MACENA, TN²; LACERDA-Jr³, GV; REZENDE, RP⁴; DIAS, JCT⁴; CASCARDO, JCM⁴.

- 1- *Estudante do Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Estadual de Santa Cruz (UESC), Ilhéus, Bahia- Brasil;*
- 2- *Mestre em Genética e Biologia Molecular;*
- 3- *Estudante de Graduação no Bacharelado em Ciência Biológicas da UESC;*
- 4- *Docentes do Departamento de Ciência Biológicas da UESC.*

Resumo

A necessidade e o potencial de identificação e bioprospecção de novos biocompostos é uma realidade. Técnicas moleculares introduzidas nas últimas décadas nos permitem ter acesso ao potencial metabólico, independente de cultivo, no entanto dependente do isolamento de DNA diretamente da amostra ambiental . A Mata Atlântica, bioma com grande diversidade biológica, possui um solo rico em matéria orgânica, com componentes que interferem nas reações necessárias à exploração de seu potencial biotecnológico. Aqui, buscamos uma otimização do tratamento da amostra antes da extração (*ex situ*) do DNA através da aplicação de uma nova metodologia empregada em amostras de solo de Mata Atlântica e pela sua comparação com outra metodologia pré-existente. A nova metodologia, mais eficiente, permitiu extrair DNA em maior quantidade e com maior capacidade de amplificação e digestão em comparação com a pré-existente, tornando possível uma melhor exploração do solo de Mata Atlântica.

Palavras-chave: Extração, Purificação, solo, Mata Atlântica.

3.1. Introdução

Os microrganismos são os organismos com a maior distribuição no planeta desempenhando papéis extremamente importantes nos ciclos biogeoquímicos, representando um enorme reservatório de moléculas aplicáveis na saúde e na indústria (ROBE et al., 2003; YUN et al., 2004). Técnicas baseadas em cultivo têm sido tradicionalmente as ferramentas primárias usadas para o estudo destes organismos de diferentes ambientes (CHANDLER et al., 1997), embora tenham sido aprimoradas, sabe-se que apenas uma pequena fração da diversidade microbiana na natureza, menos de 1%, pode ser hoje cultivada em laboratório com as tecnologias existentes (WARD et al., 1990; AMANN et al., 1995). Estas metodologias, não fornecem necessariamente informação abrangente sobre a diversidade e abundância de comunidades microbianas (van HAMME et al., 2003), impedindo a plena exploração do seu conteúdo genético para fins biotecnológicos.

Metodologias moleculares desenvolvidas nas últimas décadas (extração de ácidos nucléicos, amplificação por PCR, clonagem e seqüenciamento de DNA) têm sido otimizadas e adaptadas para superar as limitações, evitando o isolamento e cultivo dos microrganismos. A utilização destas metodologias vem permitindo uma avaliação mais precisa da diversidade microbiana no ambiente e a descoberta de novos grupos de organismos, nunca antes cultivados (AMANN et al., 1995; HUGENHOLTZ et al., 1998; SCHLOSS & HANDELSMAN, 2003).

A Mata Atlântica (MA), bioma com grande diversidade biológica, pode ter um enorme potencial como fonte de novas moléculas, sobretudo com atividade antimicrobianas, uma vez que a competição por espaço e nutrientes neste ambiente é intensa, favorecendo os microrganismos produtores de substâncias inibidoras do crescimento de seus competidores (COSTA & SIQUEIRA, 2004). A MA entretanto, possui um solo rico em matéria orgânica,

com componentes que interferem nas reações necessárias à exploração de seu potencial biotecnológico, especialmente o ácido húmico, composto por elementos de alto peso molecular, contendo nitrogênio em formas cíclicas e anéis aromáticos formados por poli condensação durante a decomposição da matéria orgânica no solo (PAUL & CLARK, 1989). A interferência que os contaminantes provocam nas moléculas de DNA, inviabilizam reações enzimáticas que são determinantes na construção de uma biblioteca metagenômica de alto peso molecular, tais como digestão, ligação e amplificação. O ácido húmico desnatura o DNA pela ligação de seus grupos fenólicos a amidas ou por serem oxidados e formarem quinonas que se ligam covalentemente ao DNA (ROBE et al., 2003; WHITEHOUSE & HOTTEL, 2006).

Em estudos com fins biotecnológicos, como identificação de novas substâncias com aplicações médicas e industriais, além da necessidade de um DNA livre de contaminantes, o método de extração tem que proporcionar uma quantidade e qualidade de DNA suficiente para as reações. E nestas situações, logicamente, quanto mais DNA, melhor, visto que a representatividade da amostra aumenta, elevando a possibilidade de se encontrar um novo gene e conseqüentemente, uma nova substância (LAMMLE et al., 2007).

Na tentativa de tornar possível a exploração biotecnológica do solo de Mata Atlântica, diversos métodos de extração de DNA *ex situ* têm sido empregados e adaptados. Neste trabalho propõe-se uma otimização do tratamento da amostra antes da extração do DNA através da aplicação de uma nova metodologia empregada em amostras de solo de Mata Atlântica e pela sua comparação com outra metodologia pré-existente.

3.2. Material e métodos

3.2.1. Coleta dos Solos

As amostragens na área sob Mata Atlântica foram coletadas na Fazenda Caimbi (14°51'37" 39°02'04") em Ilhéus, Bahia-Brasil por meio de transectos retirando-se 20 sub-amostras para formar uma amostra composta. As amostras foram retiradas na profundidade de 0-10 cm, com a utilização de uma espátula estéril, sendo imediatamente colocadas em caixa refrigerada até a chegada no laboratório. A seguir, a amostra composta foi peneirada (malha 2 mm) e resíduos de plantas e raízes removidos do solo cuidadosamente.

3.2.2. Análise dos Solos

As análises orgânica e físico-química do solo de Mata Atlântica foram realizadas pelo Laboratório de Solos e Nutrição de Plantas da Embrapa de Cruz das Almas, Bahia.

3.2.3. Métodos de extração *ex situ* de DNA utilizados

Método A para extração *ex situ* de DNA de solo de Mata Atlântica.

Quatro gramas de solo foram introduzidas em frasco Erlenmeyer de 250 mL, contendo a 50 mL de Tampão Fosfato de Sódio (pH 7,0) e Tween 80 a 0,1%. A mistura foi incubada sob agitação constante à temperatura ambiente durante 18 horas. O sobrenadante da mistura foi distribuído em tubos Falcon de 15 mL e centrifugado a 5000 r.p.m por 10 minutos. O precipitado foi lavado 4 vezes em Tampão TE 50/50 mM (Tris-HCl e EDTA), pH 8,1, e centrifugados por 3 minutos a 5000 r.p.m.

A lise física das células foi feita por maceração do precipitado em nitrogênio líquido com auxílio de Gral e Pistilo. O macerado foi suspenso em 1 mL de TE 50/50 e então adicionou-se 1 mL de fenol/clorofórmio/ álcool isoamílico (25:24:1 pH 7,3). A mistura foi centrifugada a 5000 rpm por 10 minutos e a fase superior recolhida. O DNA foi precipitado pela adição de isopropanol e acetato de sódio a 0,3 M, e deixado por 2 horas em freezer a -20 °C. Após foi centrifugado 5000 rpm por 10 minutos descartando o sobrenadante em seguida o precipitado foi lavado em etanol 70% e depois com etanol a 95%. O DNA foi suspenso em 200 µL de tampão TE (Tris-HCl) e Purificado em coluna de Sephadex G-200 de acordo com Maciel (2007).

Método B para extração *ex situ* de DNA de solo de Mata Atlântica.

Quatro gramas de solo foram introduzidas em frasco Erlenmeyer de 250 mL, contendo a 50 mL de Tampão Fosfato de Sódio (pH 7,0) e Tween 80 a 0,1%. A mistura foi incubada sob agitação constante à temperatura ambiente durante 18 horas. Após decantação, o sobrenadante da mistura foi distribuído em tubos cônicos de 15 mL e centrifugados a 5000 r.p.m por 10 minutos. Adicionou-se Glass Beads ao precipitado em proporção de igualdade. O precipitado foi lavado 4 vezes em Tampão TE 50/50 mM (Tris-HCl e EDTA), pH 8,1, com centrifugação por 3 minutos a 5000 r.p.m. Antes de cada centrifugação, realizou-se agitação intensa da suspensão em vortex por um minuto, a fim de promover auxílio da desfloculação pelas Glass Beads. Após a última lavagem, adicionou-se 200 µL de SDS a 20% e incubou-se a temperatura ambiente por 20 minutos.

A lise física das células foi feita por choque térmico do precipitado alternando exposição do tubo fechado ao nitrogênio líquido por 1 minuto seguida de exposição do mesmo à água em fervura por 4 minutos em 3 ciclos. O DNA foi extraído pela adição de fenol/clorofórmio/ álcool isoamílico (25:24:1 pH 7,3) na proporção de 1:1. A mistura foi centrifugada a 5000 rpm por 10 minutos e a fase superior recolhida. O DNA foi precipitado pela adição de isopropanol e acetato de sódio a 0,3 M, e deixado por 2 horas em freezer a -20 °C. A seguir, foi centrifugado a 5000 rpm por 10 minutos descartando o sobrenadante com posterior lavagens do precipitado em etanol 70% e etanol a 95% na sequência. O DNA foi suspenso em 200 µL de TE clássico (pH 8,0). O procedimento de purificação deste segundo método de extração de DNA teve três variações:

Purificação 1- Uso de coluna de Sephadex G-200 após a extração.

Purificação 2- Uso de PVPP (Polivinilpolipirrolidona) antes da lise em associação com uso de coluna de Sephadex G-200 após extração.

Purificação 3- Após extração, eletroforese em gel de agarose 0,8%, corte da banda de interesse (48 kb) e purificação em coluna GFX (Amersham Biosciences).

3.2.4. Capacidade de amplificação via PCR das formas de purificação do método B

Os DNAs recuperados das três formas de purificação do método B foram submetidos à reação da cadeia da polimerase (PCR). Os recuperados nas purificações 1 e 2 foram diluídos 100 vezes. O recuperado da purificação 3 não foi diluído devido a sua grande perda de DNA (dados não mostrados) durante o procedimento. A região amplificada foi a codificante de

RNA ribossômico 16S. Os primers utilizados foram 27f-[5'- AGA GTT TGA TCM TGG CTC AG- 3'] e 1525r- [5'- AAG GAG GTG WTC CAR CC- 3'].

A PCR foi realizada em termociclador MJ Research PTC200 (Eppendorf). Na reação foram utilizadas as seguintes condições de tempo e temperatura: 1 ciclo 94°C por 5 minutos; 34 ciclos de 94°C por 1 minuto, 59°C por 1 minuto e 72°C por 2 minutos, finalizando com 1 ciclo de 72°C por 10 minutos.

A PCR foi feita contendo as seguintes proporções de reagentes: 2,5 µL de tampão 10X da reação para a *Taq* DNA polimerase (100 mM Tris-HCl, pH 8,8 a 25°C, 500 mM KCl, 0,8%) concentração final de 1X; 3,7 mM de cloreto de magnésio, 0,4 pmol/µL de deoxirribonucleotídeos (dATP, dTTP, dCTP e dGTP), 0,4 pmol/µL de cada iniciador, 0,4 mM de BSA (Soro Albumina Bovina), 3 unidades de enzima *Taq* DNA polimerase (Fermentas) e água ultra pura estéril para um volume final de 25 µL. Como molde para a reação foram utilizadas alíquotas de 150 ng do DNA total das amostras de solo diluídas na proporção 1:1000.

3.2.5. Comparação do rendimento de DNA

Os métodos A e B foram comparados a fim de indicar aquele com maior rendimento de DNA, sob as mesmas condições iniciais. As extrações foram feitas a partir de frações de 4g de solo de Mata Atlântica de amostra coletada conforme item 1. Foi realizada uma análise espectrofotométrica em Genequant (GE healthcare) para quantificar o rendimento de DNA de A e B. Também foi realizada uma eletroforese para comparar visualmente este rendimento.

3.2.6. Capacidade de digestão por *EcoRI* e *BamHI*

O DNA extraído com o método B, em suas formas de purificação 1 e 2, foi comparado com o extraído pelo método A quanto à capacidade de digestão pelas enzimas *EcoRI* e *BamHI*. O teste foi feito pela exposição de 110 ng de DNA a 1 unidade de enzima por uma hora.

3.3. Resultados e discussão

Análise do solo

O resultado da análise do solo encontra-se na Tabela 1.

Table 3.1- Característica físicas e químicas do solo se Mata Atlântica coletado.

Sample	pH	P	K	Ca	Mg	Ca+Mg	Al	Na	H+Al	S	CTC	V	Organic matter
	In water	mg/dm ³	cmol _c /dm ³									%	g/kg
Atlantic Forest	6.0	39	0.82	11.8	0.8	12.6	0.1	0.30	4.73	1372	18.45	74	68,82

Granulometry (g/Kg) with NaOH dispersion

Sand silt clay	silt	argyle	Textural classification
496	232	272	sandy clay soil

Capacidade de amplificação via PCR do DNA obtido pelas metodologias de purificação do método de extração B

Entre os três tipos de purificação 1 e 2 mostraram-se eficientes quanto à amplificação (Figura 3.1). A forma 3, mesmo com resultados de leitura em espectrofotômetro nas razões 260 nm/230 nm e 260 nm/280 nm mostrando nível aceitável de contaminação para eDNA (enviromental DNA), não se mostrou eficiente quanto à amplificação.

A amplificação via PCR é uma boa estratégia de avaliação do nível de pureza do DNA, uma vez que envolve sucessivas reações enzimáticas (RIESENFELD et al., 2004). A acoplagem das DNA polimerases à molécula de DNA exige sítios livres de contaminação (ROH et al., 2006). Sendo assim, eDNA não amplificável em condições ótimas de reação provavelmente tem sua amplificação inibida por contaminantes, como os ácidos húmicos, que impedem a ação da Taq-DNA polimerase (ROBE et al., 2003; KOZDROJ & van ELSAS, 2001; ROH et al., 2006.).

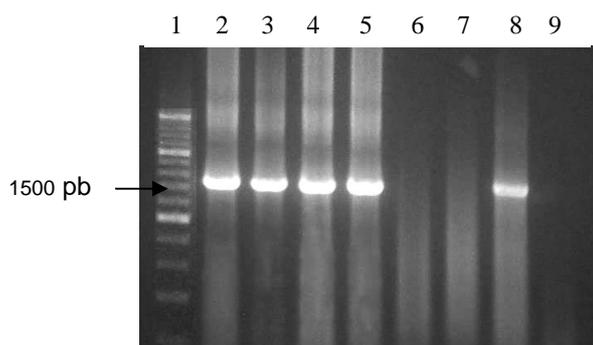


Figura 3.1- Amplificação do gene 16S de DNAs provenientes de três metodologias de purificação no método A de extração: Canaleta 1- marcador 200- 3000 pb – Canaleta 2- DNA purificação 1 (1 μ L) amplificado –Canaleta 3- DNA purificação 1 (3 μ L) amplificado- Canaleta 4- DNA purificação 2 (1 μ L) amplificado- Canaleta 5- DNA purificação 2 (3 μ L) amplificado- Canaleta 6- DNA purificação 3 (1 μ L) não amplificado- Canaleta 7- DNA purificação 3 (3 μ L) não amplificado- Canaleta 8- Controle positivo- Canaleta 9- Controle negativo.

Rendimento e qualidade de DNA e capacidade de digestão por *EcoRI* e *BamHI*.

Com o método B, obteve-se um rendimento de, em média, 1000 ng.µL⁻¹, enquanto com o método A, obteve-se rendimento de, em média, 310 ng.µL⁻¹. O resultado dessas leituras em espectrofotômetro mostram que o método B tem um rendimento quase uma vez maior que o do método A. Além disso, tendo um rendimento de 50 µg/g de amostra de solo, o método B mostrou-se mais eficiente do que o melhor método de extração de DNA de solo (método 6) do trabalho de Roh e colaboradores, em 2006, cujo rendimento foi de 8,3 microgramas de DNA por grama de amostra de solo.

O melhor rendimento do método B também pode ser observado visualmente pela análise de bandas após eletroforese das amostras comparando-se com o método A (Figura 3.2). As principais diferenças entre os dois métodos de extração estão na adição de *Glass Beads* no método B, enquanto no método A isso não acontece, e na etapa de lise, na qual no método B, utiliza-se choque térmico e no método A, maceração. Estas duas diferenças podem ser a chave da explicação do maior rendimento de DNA no método B. Por se tratar de um método de extração indireto, onde as células são desligadas das partículas de solo antes da lise (PARK et al., 2006), a adição de *Glass Beads* provocou aumento no rendimento de DNA por também aumentar o número de células desligadas das partículas de solo, livres para sofrer uma lise eficiente de maneira a liberar seu material genético. Em metodologias indiretas de extração de DNA de solo, as células que permanecem aderidas às partículas do solo são perdidas, uma vez que a própria metodologia força a remoção dessas partículas (ROBE et al., 2003; STREIT & SCHMITZ, 2004). Isso também provoca um viés na representatividade da diversidade, o que pode resultar em subestimação da diversidade

da microbiota de um ambiente (DANIEL, 2004). Proporcionalmente às limitações que um viés na representatividade da diversidade pode provocar, também são provocadas limitações quando o objetivo do trabalho é a bioprospecção de novos genes, uma vez que a microbiota contida na amostra ambiental não é representada de forma eficiente.

A lise por choque térmico também contribui para um maior rendimento uma vez que, estando o material exposto a diferenças bruscas de temperatura, todas as células presentes sofrerão esta pressão, aumentando a eficiência de lise. Já a maceração, pode não expor tantas células à sua pressão de lise quanto o choque térmico, especialmente quando se leva em conta a habilidade manual no procedimento (DUBEY et al., 2006; ROBE et al., 2003).

Além disso, o resultado de leitura na razão 260 nm/230 nm para DNA extraído com o método B, foi, em média, 1,9. Para DNA extraído com o método A, foi, em média, 1,25. Isso mostra que o método B também rende um DNA mais livre de ácidos húmicos que o do método A, visto que admite-se resultado igual ou maior que 2 na razão 260 nm/230 nm para DNA livre de contaminação por ácidos húmicos (ROH et al., 2006).

O resultado de leitura na razão 260 nm/280 nm para DNA extraído com o método B foi, em média, 0,35 e 0,5 para DNA extraído com o método A. Isso mostra que ambos os DNAs estão livres de contaminação por proteína, tomando com referência o limite máximo de 1,7 (ROH et al., 2006).

Essa diferença no nível de pureza entre os DNAs acabou por interferir na digestão com as enzimas *EcoRI* e *BamHI*, onde o DNA extraído com o método B em suas duas formas de purificação foi perfeitamente digerido tanto por *EcoRI* quanto por *BamHI*. Já o DNA extraído com o método A foi apenas parcialmente digerido (Figura 3.2).

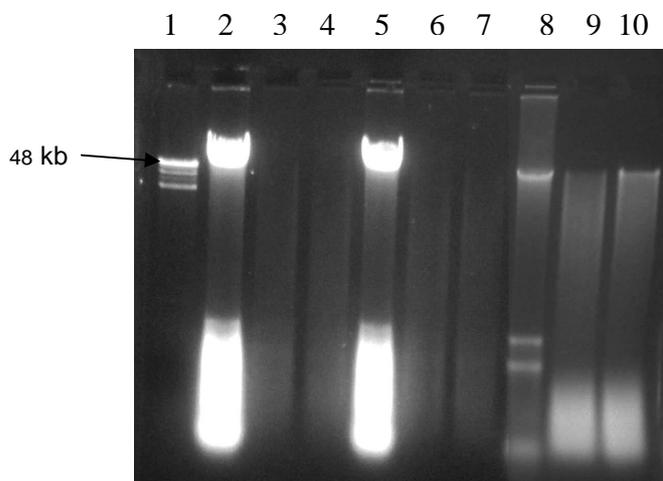


Figura 3.2-Digestão do eDNAs oriundos dos métodos A e B com enzimas de restrição: Canaleta 1- marcador de alto peso molecular 20-40 kb. Canaleta 2- DNA extraído pelo método B, purificação 1 sem digestão. Canaleta 3- DNA extraído pelo método B purificação 1 digerido por *EcoRI*. Canaleta 4- DNA extraído pelo método B purificação 2 digerido por *BamHI*. Canaleta 5- DNA extraído pelo método B purificação 2 sem digestão. Canaleta 6- DNA extraído pelo método B purificação 2 digerido por *EcoRI*. Canaleta 7- DNA extraído pelo método B purificação 2 digerido por *BamHI*. Canaleta 8- DNA extraído pelo método A sem digestão. Canaleta 9- DNA extraído pelo método A parcialmente digerido por *EcoRI*. Canaleta 10- DNA extraído pelo método A parcialmente digerido por *BamHI*.

3.4. Conclusão

Conseguimos otimizar o tratamento da amostra de solo de Mata Atlântica antes da extração do DNA tornando possível a sua extração em quantidade e nível de pureza adequados a construção de uma biblioteca metagenômica. O método B mostrou-se superior ao método A tanto na quantidade de DNA extraído, quanto em sua qualidade. O resultado de leitura em espectrofotômetro do DNA extraído pelo método B, foi praticamente o suficiente para determiná-lo como livre de contaminação por ácido húmico. Além disso, duas das três formas de purificação do método B foram amplificáveis com PCR, mostrando sua adequação para estudos filogenéticos. Ainda devido ao baixo nível de contaminação, o

DNA extraído pelo método B mostrou uma maior facilidade em ser digerido pelas enzimas *EcoRI* e *BamHI*. Todos estes resultados mostram que o método B para extração de DNA de solo de Mata Atlântica pode ser perfeitamente empregado para estudos moleculares. Esta dificuldade superada abre caminho para a análise do perfil filogenético da microbiota do solo da Mata Atlântica, bem como sua exploração biotecnológica. A exploração de um ambiente tão rico em biodiversidade pode resultar, assim, em descobertas com potencial de melhorar a qualidade de vida das pessoas.

3.5. Referências bibliográficas

AMANN, R. I.; LUDWIG, W. e SCHLEIFER, K. **Microbiology Reviews**. Vol. 59, p. 143-169, 1995.

COSTA, J.L. e SIQUEIRA, E.R. Análise de DNA dos solos e atividade enzimática como bioindicadores de diversidade microbiana em sistemas de restauração florestal da Mata Atlântica. **II Seminário de Pesquisa FAP-SE, Embrapa, SE**. 2004.

DANIEL, R. The soil metagenome—a rich resource for the discovery of novel natural products. **Current Opinion in Biotechnology**. Vol. 15. p.199–204. 2004.

DUBEY, S.K.; TRIPATHI, A.K. e UPADHYAY, S.N . Exploration of soil bacterial communities for their potential as bioresource. **Bioresource Technology**. Vol. 5. p. 2217–2224, 2006.

HUGENHOLTZ, P.; GOEBEL, B.M. e PACE, N.R.. **Journal of Bacteriology**. Vol.180. p. 4765-4774, 1998.

KIM, Y.J.; CHOI, G.; KIM, S.; YOON, G.; KIM, Y.S.; RYU, Y. Screening and characterization of a novel esterase from a metagenomic library. **Protein Expression and Purification**. Vol. 45. p. 315–323, 2006

KOZDROJ, J. e van ELSAS, J.D. Structural diversity of microbial communities in arable soils of a heavily industrialised area determined by PCR-DGGE fingerprinting and FAME profiling. **Applied Soil Ecology**. Vol. 17. p. 31–42, 2001

LAMMLE, K.; ZIPPER, H.; BREUER, M.; HAUER, B.; BUTA, C.; BRUNNER, H. e RUPP, S. Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. **Journal of Biotechnology**. Vol. 127. p.575–592. ,2007

MACIEL, B.M ; DIAS, J.T ; SANTOS, A. C. F. ; ARGOLO FILHO, R. C. ; FONTANA, R. ; LOGUERCIO, L. L. ; REZENDE, R.P. Microbial surfactant activities from a petrochemical landfarm in a humid tropical region of Brazil. **Canadian Journal of Microbiology**. Vol. 53. p. 937-943, 2007.

PAUL, E.A. and CLARK, F.E. Soil Microbiology and Biochemistry. **New York: Academic Press**; 1989.

RIESENFELD, C.S.; SCHLOSS, P.D. and, HANDELSMAN, J. METAGENOMICS: Genomic Analysis of Microbial Communities. **Annual Review of Genetics**. Vol. 38 p. 525-552, 2004.

ROBE, P; NALIN, R.; CAPELLANO, C.; VOGEL, T.M.; SIMONET, P. Extraction of DNA from soil. **European Journal of Soil Biology** . Vol. 39. p. 183-190, 2003

ROH, C.; VILLATTE, F.; KIM, B. e SCHMID, R.D. Comparative Study of Methods for Extraction and Purification of Enviromental DNA From Soil and Sludge Samples. **Applied Biochemistry and biotechnology**. Vol. 134 p. 97-112, 2006

RONDON, M.R.; AUGUST, P.R.; BETTERMANN, A.D.; BRADY, S.F.; GROSSMAN, T.H.; LILES, M.R.; LOICONO, K.A.; LYNCH, B.A.; MACNEIL, I.A.; MINOR, C.; TIONG, C.L.; GILMAN, M.; OSBURNE, M.S.; CLARDY, J.; HANDELSMAN, J.; GOODMAN, R.M. Cloning the Soil Metagenome: a Strategy for Accessing the Genetic e Functional Diversity of Uncultured Microorganisms. **Applied and Enviromental Microbiology**. Vol. 18 p. 2541–2547, 2000.

SCHLOSS, P. and HANDELSMAN, J. Biotechnological prospects from metagenomics. **Current Opinion in Biotechnology**. Vol. 14. p. 303–310, 2003

STREIT, Wolfgang R. e SCHMITZ, Ruth A. 2004. Metagenomics – the key to the uncultured microbes. **Current Opinion in Microbiology**. Vol 7. p. 492–498, 2004

van HAMME, J.D.; SINGH, A. e WARD, O.P.. **Microbiol. Mol. Biol. Rev.** Vol. 67. p.503-549, 2003.

VOGET, S.; STEELE, H.L.; STREIT, W.R. Characterization of a metagenome-derived halotolerant cellulase. **Journal of Biotechnology**, 2006.

WARD, D.M.; WELLER, R. & BATESON, M.M. (1990). Microbial molecular taxonomy. **Nature**. Vol. 345. p. 63-65, 1990.

WHITEHOUSE, C.A and HOTTEL, H.E. Comparison of five commercial DNA extraction kits for the recovery of *Francisella tularensis* DNA from spiked soil samples. **Molecular and Cellular Probes**. Vol. 21. p. 92–96, 2007

YUN, J.; KANG, S.; PARK, S.; YOON, H.; KIM, M.; HEU, S. and RYU, S. Characterization of a Novel Amylolytic Enzyme Encoded by a Gene from a Soil-Derived Metagenomic Library. **Applied and Environmental Microbiology**. Vol. 11 p. 7229–7235, 2004

4. Capítulo 2 – Uma rápida análise do perfil taxonômico e metabólico de metagenomas de solos

AMORIM, J.H¹; SILVA-MACENA, T.N²; COSTA, M.M.C³; VIDAL, R.O²; VILELA-DIAS, C.²,
LACERDA-Jr⁴, G.V; de OLIVEIRA, J.M⁴; REZENDE, R.P⁵; DIAS, J.C.T⁵; CASCARDO, J.C.M⁵

1. Estudante do Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Estadual de Santa Cruz (UESC), Ilhéus, Bahia- Brasil; 2. Mestre em Genética e Biologia Molecular; 3. Pesquisador do CENARGEN/EMBRAPA; 4. Estudante de Graduação no Bacharelado em Ciências Biológicas da UESC; 5. Docentes do Departamento de Ciências Biológicas da UESC.

Resumo

O solo é um dos mais desafiadores ambientes naturais para os microbiologistas por conter uma diversidade estimada no intervalo de 2.000 a 18.000 diferentes genomas por grama de solo. Mas toda essa diversidade genética e funcional tende a ser subestimada pelas técnicas de metagenômica aplicadas até o momento, principalmente em função das metodologias de extração de DNA. Aqui nós apresentamos uma alternativa baseada na simplificação do método de extração de DNA, utilizando uma amplificação do DNA com primers aleatórios tal qual na técnica de RAPD (**R**andom **A**mplified **P**olymorphic **D**N^A), aliado a uma clonagem em vetores A/T, com o propósito de acessar a diversidade genética e funcional de metagenomas de solo, com rápida aplicação e baixo custo. A nova metodologia mostrou eficiência em representar de maneira fiel às microbiotas dos solos de landfarm e Mata Atlântica, em níveis taxonômico e funcional.

Palavras-chave: Metagenômica, RAPD, diversidade, perfil funcional..

Introdução

O solo é um dos principais reservatórios da diversidade genética microbiana e pode ser considerado como um ambiente complexo (CURTIS et al., 2002). Essa extrema complexidade resulta de múltiplos parâmetros relacionados, incluindo textura e estrutura do solo, umidade, pH, variações climáticas e atividade biótica. Um grama de solo de floresta contém, em estimativa, 4×10^7 células procarióticas, enquanto um grama de solos cultiváveis contém 2×10^9 células procarióticas. Baseado na reassociação cinética de DNA isolado de vários solos, o número de genomas procariotos distintos tem sido estimado no intervalo de 2.000 a 18.000 por grama de solo (DANIEL, 2004), tornando este um dos mais desafiadores ambientes naturais para os microbiologistas. Entretanto, mesmo que a diversidade de genomas procariotos presentes em apenas um grama de solo possa exceder a de procariotos já conhecidos e catalogados, existe uma tendência em subestimar essa diversidade, uma vez que parte dos genomas são excluídos nas primeiras etapas do processo de clonagem de metagenomas e conseqüentemente não são representados nas análises (DANIEL, 2004.; ROH et al., 2006).

Embora a clonagem de grandes fragmentos seja preferencial por permitir a expressão tanto de produtos diretos dos genes (proteínas, enzimas e ribozimas) como produtos secundários, provenientes de vias metabólicas clonadas de maneira completa (RONDON et al., 2000; TRINGE et al., 2005)., algumas limitações na manipulação de sistemas envolvendo BACs (SHIZUYA et al., 1992), tais como poucas cópias do vetor por célula e dificuldade de ligação de grandes fragmentos no vetor, têm levado pesquisadores a buscar novas alternativas para clonagem de DNA metagenômico (LAMMLE et al., 2007). Desta forma, alguns trabalhos têm sido bem sucedidos utilizando vetores como o pUC19 e clonando fragmentos de até 8 kb. O menor tamanho do inserto parece não ser suficiente para eliminar as chances de sucesso, pois novas enzimas entre amilases, celulasas e

proteases foram encontradas em trabalhos recentes, reforçando o valor da alternativa (YUN et al., 2004)

Contaminantes co-extraídos com DNA de solo como os ácidos húmicos, que se ligam ao DNA e impedem ou dificultam o trabalho de qualquer enzima que utilize o mesmo como substrato (ROBE et al., 2003). Os ácidos húmicos, são, portanto, capazes de inibir digestões, ampliações, hibridização DNA-DNA e até mesmo a transformação de células competentes (WILSON, 1997).

Além dessas dificuldades impostas pelos contaminantes, relacionados às etapas de extração e purificação de DNA metagenômico, quando o objetivo do trabalho é bioprospecção de novos genes, é importante atentar para a representatividade da diversidade microbiana na amostra. Isso porque, logicamente, quanto maior a representatividade da amostra, maior a probabilidade de êxito. E a maioria dos métodos utilizados para extração de DNA metagenômico de solos apresenta deficiências no que diz respeito a manter representatividade da amostra (DANIEL, 2004 & ROBE et al., 2003), uma vez que, na tentativa de livrar o DNA do contaminantes, realiza-se algumas etapas de purificação e/ou lavagens que resultam em perda de DNA, introduzindo ineficiência em nível de representatividade no trabalho.

Entre os métodos convencionais de extração, existem basicamente duas vertentes; a extração direta ou *in situ* e a extração indireta ou *ex situ*, ambas com diferentes métodos compartilhando o mesmo princípio.

A extração direta proporciona maior rendimento de DNA, mas a qualidade é questionável porque expõe o DNA a uma alta concentração de contaminantes. Se existe uma tentativa de purificar esse DNA, é introduzida uma interferência na representatividade da microbiota do ambiente em questão. E a extração indireta proporciona um DNA mais livre de contaminantes, mas as etapas de liberação das células das partículas de solo também introduzem viés na representatividade (ROH et al., 2006 & STREIT e SCHMITZ, 2004). Ainda que muitos trabalhos de metagenômica utilizando esses dois princípios de extração de DNA tenham sido bem sucedidos no sentido de identificarem novas substâncias, existe a possibilidade de evolução no que diz respeito à representatividade, pois é importante lembrar que menos de 5% da microbiota de solos é cultivável. Existem muitos microrganismos ainda desconhecidos, genes que não foram acessados, ou seja,

existe ainda um universo a ser explorado e compreendido (KIM et al., 2005). Se não houver uma busca por evolução no princípio de representatividade em metagenômica todo esse universo será explorado de forma limitada e não será tão compreendido de maneira ampla (WHITEHOUSE & HOTTEL, 2007).

Uma simples fervura da amostra de solo pode promover uma lise homogênea das células ali contidas. Isso resultaria, em teoria, em uma extração com boa representatividade da microbiota. Mas como eliminar o problema dos contaminantes?

A técnica de RAPD (random amplified polymorphic DNA) (VROH et al., 1996; KANT & TRIPLET, 2002), permite a amplificação aleatória de seqüências de DNA através da utilização de primers não específicos, pois nessa técnica, utiliza-se apenas um oligonucleotídeo de 10 bases iniciador que possui seqüência arbitrária, e portanto sua seqüência alvo é desconhecida (ARAÚJO et al., 2003), ao contrário da PCR (polymerase chain reaction). Essa técnica foi aplicada para gerar perfis de amplificação visando estudos de biodiversidade (HARRY et al., 2001; OH et al., 2003; OJO, 2007) A amplificação de DNA metagenômico via RAPD com vários primers diferentes em reações diferentes poderia permitir uma cobertura mais ampla do metagenoma. Aliando essa idéia a uma simples fervura da amostra de solo, seria possível unir os métodos de extração direta e indireta, visto que a extração baseada em uma simples fervura da amostra de solo aliada à amplificação via RAPD com vários primers diferentes manteria a representatividade da amostra. E a própria amplificação geraria seqüências sintetizadas *in vitro*, livres de contaminação.

Em reações de RAPD utiliza-se Taq-DNA polimerase, a qual possui atividade de exonuclease 5'→3'. Isso significa dizer que às seqüências amplificadas são adicionados “overhangs” dATP (desoxiadenosina-trifosfato). E existem vetores tipo T/A, como o ptz57R/T especialmente desenhados para serem ligados a esse tipo de produto de reação de amplificação (<http://www.fermentas.com/profiles/kits/pdf/instaclone1213.pdf>), o que significa dizer que a clonagem desses produtos de RAPD em vetores tipo T/A pode ser amplamente vantajosa do ponto de vista prático, pois trata-se da clonagem de seqüências metagenômicas sintetizadas *in vitro*, livres de contaminação e facilmente ligáveis a vetores comerciais como o pTZ57R/T. Tudo isso tende a possibilitar a clonagem em larga escala de

DNA metagenômico de alta qualidade com alta representatividade da microbiota ambiental, aumentando consideravelmente as chances de bioprospecção de novas substâncias.

Além do potencial de bioprospecção de novos genes, também é possível traçar o perfil metabólico da microbiota, uma vez que a clonagem de seqüências aleatórias engloba genes das mais variadas famílias e funções (BERTRAND et al., 2005). Mas por se tratar de clonagem de seqüências aleatórias, existe a necessidade de desenvolvimento de novas estratégias em bioinformática, visto que os algoritmos utilizados na determinação de diversidade e perfis filogenéticos de bibliotecas de seqüências específicas, como as do tipo 16S, não podem ser aplicadas nesse caso (SCHLOSS & HANDELSMAN, 2003). O Megan, um programa utilizado para computar e explorar o conteúdo taxonômico de bibliotecas metagenômicas compostas por seqüências aleatórias com base na taxonomia do NCBI pode servir de modelo neste caso (HUSON et al., 2007). Para determinar também o conteúdo funcional e metabólico de bibliotecas é necessário desenvolver abordagens baseadas em outras ontologias, como Gene Ontology (Gene Ontology <http://www.geneontology.org/>) e KEGG (Kyoto Encyclopedia of Genes and Genome <http://www.genome.jp/kegg/>).

Aqui, apresentamos uma tentativa de bioprospecção de novos genes, determinação do perfil microbiano e metabólico de duas amostras ambientais bem contrastantes (uma impactada por petróleo (landfarm) e outra de um solo de Mata Atlântica) por meio dessa nova abordagem de análise metagenômica.

Metodologia

Coleta dos solos. As amostragens na área sob Mata Atlântica foram coletadas na Fazenda Caimbi (14°51'37" 39°02'04") em Ilhéus, Bahia-Brasil por meio de transectos retirando-se 20 sub-amostras para formar uma amostra composta. As amostras foram retiradas na profundidade de 0-10 cm, com a utilização de uma pá, sendo imediatamente colocadas em caixa refrigerada até a chegada no laboratório. A seguir essa amostra composta foi peneirada (malha 2 mm) e resíduos de plantas e raízes removidos do solo cuidadosamente.

As amostras de solo de *landfarm* da Refinaria Landulpho Alves, São Francisco do Conde, Bahia-Brasil estudadas no presente trabalho foram coletadas, segundo metodologia descrita por SANTOS em 2007. Até a extração do DNA as amostras foram mantidas em freezer a -20°C.

Extração convencional de DNA de Landfarm e diluições para RAPD metagenômico.

Em Erlenmeyer de 250 mL, adicionou-se 4 g de solo a 50 mL de Tampão Fosfato de Sódio e Twen 80 a 0,1% e incubou-se em *shaker* à temperatura ambiente overnight. Distribuiu-se os sobrenadantes em tubos falcon de 15 mL. Centrifugou-se a 5000 r.p.m por 10 minutos e descartou-se os sobrenadantes. Foram adicionadas Glass Beads ao precipitado em proporção de igualdade e o solo foi lavado em tampão PBS (Phosphate-buffered saline) 1x pH 7 por 4 vezes, com centrifugação de 3 minutos a 5000 r.p.m, descarte do sobrenadante e ressuspensão do precipitado em 3 mL de tampão em cada lavagem. Em seguida, o solo foi novamente lavado, agora, em tampão TE 50/50 mM (Tris-HCl e EDTA) pH 8,1 por quatro vezes, também com centrifugação de 3 minutos a 5000 r.p.m, sendo o sobrenadante descartado e o precipitado ressuspensão em 3 mL de tampão em cada lavagem.

Após a última lavagem, foram adicionados 200 µL de SDS a 20% à suspensão, a qual foi incubada a temperatura ambiente por 20 minutos. A lise das células foi induzida submetendo-se o precipitado a choque térmico alternando-se exposição do tubo fechado ao nitrogênio líquido por 1 minuto seguida de exposição do mesmo à água em fervura por 4 min em 3 ciclos. Adicionou-se fenol/clorofórmio/ álcool isoamílico (25:24:1 pH 7,3) na proporção de 1:1. A fase superior foi coletada e o DNA foi precipitado com isopropanol e acetato de sódio 0,3M. Em seguida, centrifugou-se a 5000 rpm por 10 minutos e lavou-se o precipitado com etanol a 70% e depois com etanol a 95%. Resuspendeu-se o DNA em 200 µL de TE clássico e purificou-se em coluna de sephadex G-200, segundo metodologia descrita por SANTOS, em 2007. Após a purificação, determinou-se rendimento e qualidade do DNA em espectrofotômetro Genequant (Amershan Biosciences). Em seguida, o DNA foi diluído nas razões de 1:10, 1:100, 1:200 e 1:1000.

RAPD metagenômico e extração baseada em fervura de DNA de Mata Atlântica. O DNA de Landfarm extraído pelo método convencional foi diluído nas razões 1:10; 1:100:

1:200 e 1:1000 e submetido a reações de amplificação aleatória (RAPD) utilizando-se o primer de 10 bases D13 (Operon Kit- Gibco). Em seguida, a diluição de razão 1:200 foi submetida a amplificações por RAPD (Tabela 1), em reações distintas, com os primers OPC01, OPK08, OPI07, OPN18, OPN19, OPK20, OPB13, OPS07, OPS10, D13, T17, R08, Q08, OPJ04 E T20 (Operon Kit- Gibco) e os produtos foram unidos para dar origem ao Mix I. Em seguida, 1 g de amostra de solo de *landfarm* foi colocado em micro tubo Eppendorf de 1,5 mL e suspenso em 1 mL de água ultra-pura estéril. O tubo contendo a suspensão foi submetido à fervura por 20 minutos e após isso, centrifugado por 10 min a uma rotação de 13.400 rpm. O sobrenadante foi diluído na razão de 1:200. O suposto DNA contido nessa diluição foi submetido a reações de amplificação aleatória utilizando-se os mesmos primers utilizados para gerar o Mix I, gerando, então, o Mix II. Realizou-se o mesmo procedimento utilizando-se 1 g de solo de Mata Atlântica, gerando o Mix III. A diluição de DNA de solo de Mata Atlântica na razão 1:200 de DNA ainda foi submetida a mais duas rodadas de RAPD com dois grupos de primers distintos, em reações distintas, respectivamente; OPC19, OPQ14, OPD11, OPI17, OPI09, OPC17, OPM06, OPK19, OPD16, OPI19, OPI15, OPM04, OPK13, OPQ06, OPN17, OPC19, OPD17, OPB06, OPC15 e OPI06 (Operon Kit- Gibco) para gerar o Mix IV e OPB03, OPM03, OPI04, OPM13, OPQ11, OPQ15, OPD19, OPC03, OPJ16, OPK12, OPB14, OPN14, OPJ14, OPN09, OPQ10, OPM05, OPB07, OPD07, OPI18 e OPJ06 (Operon Kit- Gibco) para gerar o Mix V (Tabela 1). Todos os produtos de reação foram analisados em gel de agarose a 0,8%.

Tabela 1- Reação de RAPD utilizada para amplificação de DNA metagenômico em todas as situações do experimento.

Reagentes	Volume (µL)
Tampão IX (Fermentas)	1,25
MgCL ₂ 3,5 mM (Fermentas)	1,85
dNTPS 0,48 mM (Fermentas)	0,6
Primer 1 ng/µL	2,5
H ₂ O	0,5
<i>Taq DNA pol.</i> 0,25 U/µL (Fermentas)	3,9
BSA	1,5
DNA	0,4
Volume final	12,5

Clonagem metagenômica e construção de bibliotecas. Os Mix I, II, III e IV, respectivamente, foram quantificados em Genequant (Amersham Biosciences) e forneceram DNA metagenômico para reações de ligação ao vetor ptz57R/T (Tabela 2), conforme instruções do fabricante (Instaclone PCR cloning Kit -Fermentas). Utilizou-se 5 µL de cada reação de ligação, referente a seu respectivo Mix, para transformar células quimicamente competentes de *E. coli* TOP10 contidas em 250 µL de solução de CaCl₂, por meio de choque térmico (5 min a 0° C e 45 s a 42° C). Em seguida adicionou-se 800 µL de meio LB (Luria-Bertani) ao tubo de transformação e incubou-se a 37° C com agitação por 1 hora. Inoculou-se 100 µL por placa contendo LB com ampicilina (100 mg/mL), IPTG 50 mM e X-gal (7,5 mg/mL). Selecionou-se as colônias brancas com auxílio de catadores de colônias, transferindo-as para placas de poliestireno de 96 poços. Os Mix I, II, III, IV e V, respectivamente, deram origem às bibliotecas Land1, com 768 clones; Land2, com 768 clones; MataI, com 480 clones; Mata II, com 960 clones e Mata III, com 864 clones.

Tabela 2- Reação de ligação dos produtos de RAPD metagenômico ao vetor TA ptz57R/T (Instaclone PCR Cloning Kit- Fermentas).

Reagentes		Volume (µL)
H ₂ O		8
ptz57R/T	0,0055 µg/µL	2
Tampão	1x	4
DNA ligase	0,25 u/µL	1
DNA	40 ng/ µL	5
Volume final		20

Diagnóstico de transformação via PCR. Uma amostragem de colônias brancas de cada transformação foi realizada para diagnóstico de transformação via PCR utilizando-se os primers M13 F (5'-CGC CAG GGT TTT CCC AGT CAC GAC- 3') e M13 R (5'-TCA CAC AGG AAA CAG CTA TGA C- 3'), na concentração de 3,2 µM, em reações de 12,5 µL (volume final).

Sequenciamento e tratamento das seqüências. As amostras foram seqüenciadas em MegaBace 1000 (GE Healthcare), na Universidade Estadual de Santa Cruz, Bahia- Brasil. Em seguida, as seqüências geradas foram submetidas a um tratamento contendo as seguintes etapas; tratamento pelo Phred, Crossmatch (GREEN, 1999), Lucy e RepeatMasker (<http://www.repeatmasker.org>). Foram aceitas seqüências com extensão mínima de 100 pares de base e qualidade mínima (Phred \geq 10).

Blast e seleção das seqüências para análises de bioinformática. As seqüências aceitas após tratamento foram submetidas à comparação contra os bancos de dados do NCBI-NR, KEGG e G.O. (<http://www.geneontology.org/>). Após o Blast, aplicou-se um “cut off” -e 0.0000001. As seqüências selecionadas após este procedimento foram submetidas à análise de bioinformática.

Análises de bioinformática. A base da análise está no MEGAN (HUSON et al., 2007). , um programa cujos arquivos de entrada são resultados do Blast, e que pode ser utilizado para determinar o conteúdo taxonômico de um banco de seqüências metagenômicas utilizando a taxonomia do NCBI (National Center for Biotechnology Information <http://www.ncbi.nlm.nih.gov/>) para ordenar os resultados. Neste trabalho, nós adequamos essa metodologia ao uso da taxonomia do G.O (Gene Ontology <http://www.geneontology.org/>) e do KEGG (Kyoto Encyclopedia of Genes and Genome <http://www.genome.jp/kegg/>). Desta forma o resultado do Blast contra o banco de dados GO foi utilizado diretamente para classificação de acordo com esta ontologia. E o resultado do Blast contra o banco de dados NCBI_NR foi utilizado para classificação de acordo com a taxonomia do NCBI, utilizando o MEGAN, e da ontologia KEGG, através de um programa desenvolvido por nós que utiliza o “accession number” (código de acesso do banco de dados NCBI_NR) do “hit” para agrupamento de acordo com as vias metabólicas, baseando no “ko number” (código de acesso do KEGG) correspondente. Também foram desenvolvidos programas para formatação de arquivos baseados nos agrupamentos descritos acima, para entrada e execução do programa TREEVIEW (<http://jtreeview.sourceforge.net/>) e utilizados seus os gráficos de saída. Além disso, foi

comparada a diversidade taxonômica e metabólica entre bibliotecas para fins específicos, como descritos a seguir.

Determinação do impacto da metodologia da extração de DNA na representatividade da diversidade taxonômica. Após os tratamentos citados, as seqüências das bibliotecas Land1 e Land2 foram segregadas por grupo taxonômico (taxonomia do NCBI-NR) e comparadas quanto à capacidade de representar a diversidade da amostra. Para cada grupo taxonômico, a comparação foi feita baseada em três estatísticas: Stekel, Audic-Claverie e Fish.

Comparação da diversidade funcional entre as bibliotecas Land1 e Land2. Após os tratamentos, as seqüências das bibliotecas Land1 e Land2 foram segregadas por grupo funcional (função biológica com base no G.O) e comparadas quanto à capacidade de representar a diversidade da amostra. Para cada grupo, a comparação foi feita baseada em três estatísticas: Stekel, Audic-Claverie e Fish, exigindo-se um mínimo de cinco reads.

Comparação das características taxonômica e funcional entre os solos de *landfarm* e Mata Atlântica (Land2 vs Mata1). Por terem sido construídas com o mesmo conjunto de primers, as bibliotecas Land2 e Mata1 foram escolhidas para relalizar as comparações das características taxonômica e funcional entre os solos de *landfarm* e Mata Atlântica. Após os tratamentos, as seqüências das respectivas bibliotecas foram segregadas por grupo taxonômico (NCBI) e funcional (função biológica com base no G.O) e comparadas quanto às características taxonômicas e funcionais. Para cada grupo, a comparação foi feita baseada em três estatísticas: Stekel, Audic-Claverie e Fish, exigindo-se um mínimo de cinco reads.

Determinação do impacto de número de seqüências e números de primers utilizados na comparação da diversidade taxonômica e funcional entre a biblioteca Land2 e o solo de Mata Atlântica (Land2 vs Mata Total). Após os tratamentos, as seqüências das bibliotecas Land2 e o conjunto Mata Total (bibliotecas Mata1, Mata2 e Mata3) foram segregadas por grupo taxonômico (NCBI) e funcional (função biológica com base no G.O)

e comparados quanto à capacidade de representar a diversidade da amostra, em nível taxonômico e funcional. Para cada grupo, a comparação foi feita baseada em três estatísticas: Stekel, Audic-Claverie e Fish, exigindo-se um mínimo de cinco reads.

Determinação do impacto de número de primers na comparação da diversidade taxonômica e funcional entre os solos de *landfarm* e Mata Atlântica (Land Total vs Mata Total). Após os tratamentos, as seqüências das bibliotecas dos conjuntos Land Total (Land1 e Land2) e Mata Total (bibliotecas Mata1, Mata2 e Mata3) foram segregadas por grupo taxonômico (NCBI) e funcional (função biológica com base no G.O) e comparados quanto à capacidade de representar a diversidade da amostra, em nível taxonômico e funcional. Para cada grupo, a comparação foi feita baseada em três estatísticas: Stekel, Audic-Claverie e Fish, exigindo-se um mínimo de cinco reads.

Predição do número de ORFs potenciais nas bibliotecas de Mata Atlântica e *Landfarm*.

A determinação da predição de ORFS foi executada utilizando-se o programa getorf da família de programas EMBOSS. (<http://emboss.sourceforge.net/download/>), versão 5.0.0 com restrição para predição de ORFs com, no mínimo, 50 aminoácidos em Mata Atlântica Total e Landfarm Total. As seqüências FASTA com mais de 50 aminoácidos foram submetidas a uma análise de BLAST-P e foram consideradas aceitas apenas seqüências que apresentaram hit no GenBank com $e \text{ value} \leq e^{-7}$.

Resultados e discussão

Rendimento de DNA através do método de extração convencional. Obteve-se DNA na concentração de 800 ng/ μ L utilizando-se o método convencional de extração de DNA de solo desenvolvido no Laboratório de Monitoramento Ambiental da UESC. Além disso, o DNA extraído apresentou média 2,0 na razão 260 nm/230 nm de leitura em

espectrofotômetro Genequant (Amershan Biosciences), revelando alto rendimento e qualidade (ROH et al., 2006).

Padronização do RAPD metagenômico para DNA de *landfarm* extraído através do método convencional e DNA de Mata Atlântica extraído através do método especial. A diluição de razão 1:200 do DNA de *landfarm* extraído através do método convencional mostrou a melhor amplificação (Figura 4.1). Essa diluição é necessária para fornecer uma quantidade de DNA molde em equilíbrio com os outros reagentes da reação de RAPD; uma quantidade exagerada de DNA pode saturar a reação bloqueando a ação dos primers através de seu consumo total nos ciclos iniciais. Isso poderia impedir a amplificação das seqüências metagenômicas. Todas as diluições do DNA de *landfarm* utilizadas neste trabalho apresentaram boa amplificação via RAPD, mas a diluição na razão 1:200 apresentou uma maior quantidade de bandas quando analisadas após eletroforese em gel de agarose a 0,8%, o que é mais interessante do ponto de vista biotecnológico. A mesma razão de diluição foi escolhida para amplificar DNA de Mata Atlântica extraído através da metodologia especial. E essa escolha teve como premissa um DNA, supostamente, com a concentração no intervalo de 1 a 10 ng/ μ L após diluído.

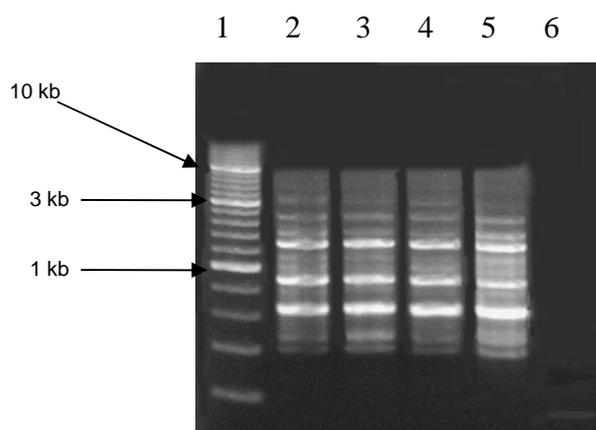


Figura 4.1 - Otimização da amplificação do eDNA de lanfarm utilizando-se primers randômicos; 1. Marcador 1Kb (Fermentas); 2. RAPD da diluição 1:10; 3. RAPD da diluição 1:100; 4. RAPD da diluição 1:1000, 5. RAPD da diluição 1:200 e 6. controle negativo.

Amplificação aleatória de DNA metagenômico de *landfarm* extraído através de metodologia convencional. Dos quinze primers de RAPD utilizados para amplificar DNA de *landfarm* extraído através de método tradicional, dez foram considerados eficientes; D13, OPC01, OPK08, OPB13, OPS07, OPS10, T17, Q08, OPJ04 E T20 (Figura 4.2). Sequências com mais de 1 kb foram amplificadas.

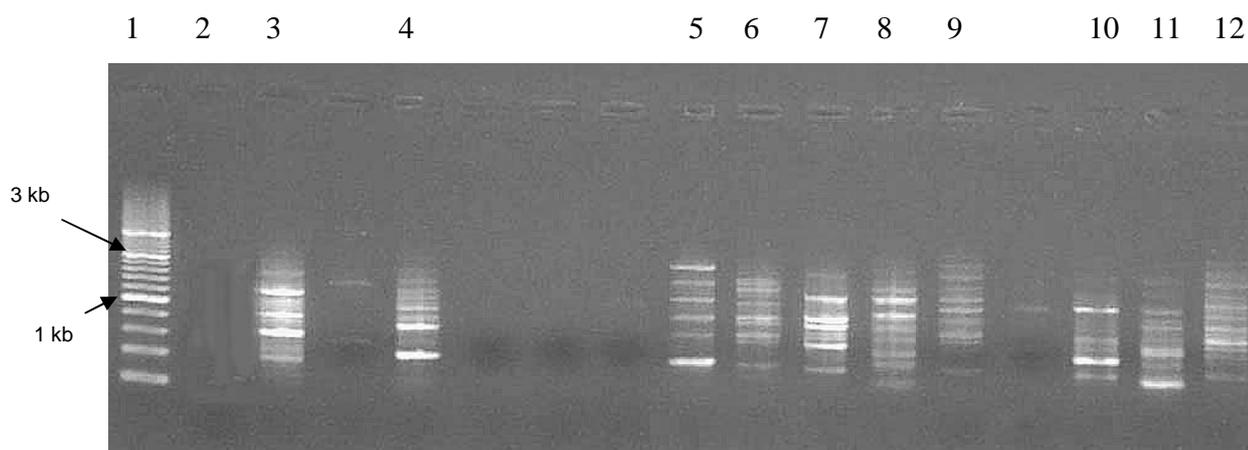


Figura 4.2- Da esquerda para a direita; 1-Marcador 1Kb (Fermentas); 2-controle negativo utilizando o primer D13; reação eficiente com o primer, respectivamente; 3-D13, 4-OPC01, 5-OPK08, 6-OPB13, 7-OPS07, 8-OPS10, 9-T17, 10-Q08, 11-OPJ04 e 12-T20

Amplificação aleatória de DNA metagenômico de *landfarm* extraído através de metodologia baseada em fervura. Dos quinze primers de RAPD utilizados para amplificar DNA de *landfarm* extraído através de método baseado em fervura, da mesma forma que no experimento anterior, dez foram considerados eficientes; D13, OPC01, OPI07, OPN18, OPN19, OPK20, OPB13, OPS07, OPS10 e Q08 (Figura 4.3). Sequências de mais de 1 kb foram amplificadas.

Embora a quantidade de primers considerada eficiente na amplificação de DNA de *landfarm* extraído pelas duas metodologias tenha sido a mesma, houve diferença quanto à eficiência dos primers; para DNA de *landfarm* extraído através de metodologia convencional, os primers OPJ04, OPK08, OPS10, T17 e T20 foram exclusivos em mostrar eficiência em amplificar DNA metagenômico. E para DNA extraído através de metodologia especial, os primers OPN19, OPS10, OPI07, OPN18 e OPK20 foram exclusivos.

Em comum, os primers D13, OPB13, OPC01, OPS07 e Q08 foram eficientes na amplificação de DNA de *landfarm* extraído através das duas metodologias em questão. Isso acontece porque, embora extraído por metodologias diferentes, o DNA em questão é oriundo da mesma amostra de solo. Logo, esses oligonucleotídeos de dez bases devem encontrar seqüências homólogas comuns nos dois tratamentos.

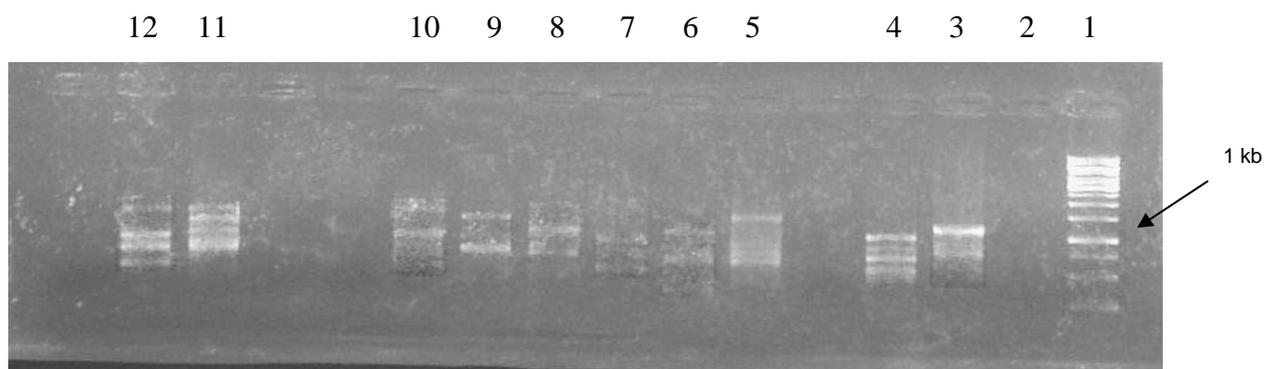


Figura 4.3- Da direita para a esquerda; 1-Marcador 1Kb (Fermentas); 2-controle negativo utilizando o primer D13; reação com o primer, respectivamente; 3-D13, 4-OPC01, 5-OPI07, 6-OPN18, 7-OPN19, 8-OPK20, 9-OPB13, 10-OPS07, 11-OPS10 e 12-Q08.

Amplificação aleatória de DNA metagenômico de Mata Atlântica extraído através de metodologia especial. Na amplificação aleatória de DNA de Mata Atlântica com a utilização do conjunto de primers referente ao Mix III, os eficientes foram D13, OPC01, OPK08, OPB13, OPS10, T17, Q08 e T20 (figura não mostrada).

Na amplificação aleatória de DNA de Mata Atlântica com a utilização do conjunto de primers referente ao Mix IV, os eficientes foram OPC19, OPQ14, OPD11, OPI17, OPI09, OPM06, OPK19, OPD16, OPQ06, OPB06, OPC15 e OPI06 (figura não mostrada).

Na amplificação aleatória de DNA de Mata Atlântica com a utilização do conjunto de primers referente ao Mix V, os eficientes foram OPM13, OPQ11, OPQ15, OPJ16, OPK12, OPN14, OPN09, OPQ10, OPB07, OPD07 e OPJ18 (figura não mostrada).

Confirmação de clonagem. Houve confirmação de clonagem em todas as amostragens de colônias brancas (seleção *blue-white*) de todas as transformações. Esses diagnósticos de transformação mostraram eficiência não apenas na clonagem em si, mas também no

tamanho das sequências clonadas. É possível visualizar *amplicons* sinalizando a clonagem de seqüências em torno de 1 kb e até mesmo seqüências com tamanho próximo a 2 kb (Figura 4.4), o que permitiu vislumbrar um potencial em bioprospecção de genes desses metagenomas (FERRER et al., 2007).

Esse diagnóstico é possível graças ao fato de os primers de sequenciamento M13 reconhecerem seqüências homólogas nas extremidades do múltiplo sítio de clonagem do vetor (pTZ57R/T) (Clark, 1998; <http://www.fermentas.com/catalog/kits/kitinstaclone.htm>, acessado aos 11.01.2008). Em função disso, quando o DNA de uma colônia de clones positivos é utilizado como molde para PCR, os primers anelam nessas regiões do vetor e terminam amplificando também o inserto, caso este esteja presente.

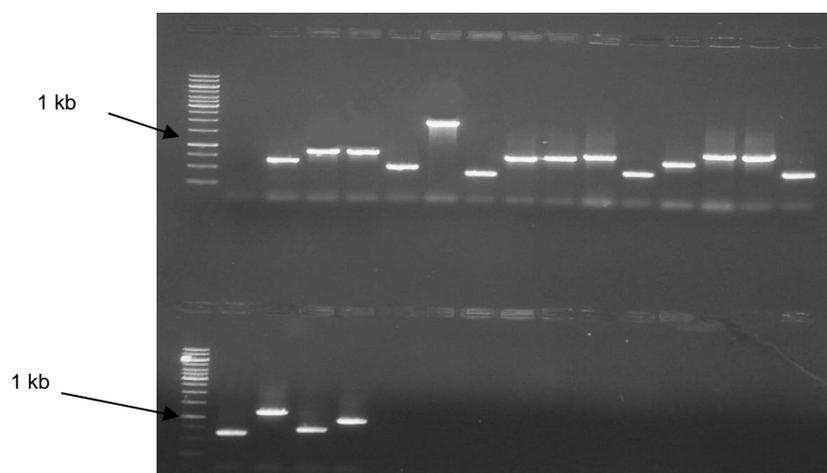


Figura 4.4- Diagnóstico de clonagem de DNA metagenômico de MA em pTZ57R/T através de PCR de colônia utilizando-se o primer M13- Na parte de cima, da esquerda para a direita; Marcador 1 kb (Fermentas); controle negativo; confirmação de clonagem em seqüência para todas as amostras analisadas. Na parte de baixo, da esquerda para a direita; Marcador 1 kb (Fermentas); confirmação de clonagem para mais quatro amostras.

Determinação do impacto da metodologia da extração de DNA na representatividade da diversidade taxonômica. Houve diferenças significativas na representatividade de 15 grupos taxonômicos entre Land1 (570 reads selecionadas) e Land2 (435 reads selecionadas).

Land2 foi superior em representar 8 grupos taxonômicos; Esfingobacteriáceas, grupo do *Bacillus cereus*, Hidrogenivirga, Sphingomonadales, *Bacillus sp.* SG-1, Rhodobactérias ainda não classificadas, Proteobactérias, Solibacteres e microorganismos

enquadrados nos “no-hits”. Nos grupos Sphingobacteriaceae, grupo do *Bacillus cereus*, Hydrogenivirga e Sphingomonadales a quantidade de “reads” em Land1 foi zero, enquanto em Land2 a quantidade foi significativamente alta. Isso mostra que o método de extração convencional deve levar a uma perda de DNA desses grupos taxonômicos (RIESENFELD et al., 2004). E a supressão de quaisquer grupos taxonômicos em bibliotecas metagenômicas limita seu potencial; primeiro porque mascara o perfil filogenético do ambiente, além de impedir a sinalização e caracterização de novos grupos taxonômicos; segundo porque em função do primeiro, a probabilidade de identificar novas substâncias diminui consideravelmente.

Mesmo assim, a biblioteca Land1 foi significativamente superior em representar 7 grupos taxonômicos (ver Figura 4.5). Mas para apenas um grupo (*Legionellales*) houve representatividade zero por parte de Land2. Essa comparação mostra um aparente equilíbrio na representatividade dos grupos taxonômicos por parte das duas bibliotecas, mas o fato de a biblioteca Land1 ter deixado de representar nove grupos taxonômicos mostra que construir uma biblioteca metagenômica utilizando um método convencional de extração de DNA introduz uma ineficiência muito grande em nível de representatividade, o que levaria a um erro em análises filogenéticas e, certamente também ocorreriam perdas sob o ponto de vista biotecnológico.

Cluster name	Land2 (570)	Land1 (435)	Stekel	Audic-Claverie	Fisher
1 Sphingobacteriaceae	53.0	0.0	-1.0	5.0150454291682714E-14	4.576181581650882E-13
2 Bacillus cereus group	28.0	0.0	-1.0	7.203690136455931E-8	1.8692845920734095E-7
3 Hydrogenivirga	23.0	0.0	-1.0	1.2274704732470594E-6	3.664331077662162E-6
4 Sphingomonadales	22.0	0.0	-1.0	2.164224255461823E-6	6.8840705643828574E-6
5 Bacillus sp. SG-1	0.0	11.0	-1.0	4.3239268918963E-5	1.0875827639795114E-4
6 Not assigned	6.0	22.0	7.2771326292318745	8.356993789144171E-5	3.163047173982589E-4
7 Legionellales	0.0	10.0	-1.0	9.989762129552204E-5	2.477542798662674E-4
8 unclassified Rhodobacterales	15.0	0.0	-1.0	1.1464119453349842E-4	2.7441215331711885E-4
9 Nitrosomonadales	0.0	8.0	-1.0	5.332228561182833E-4	0.0012905718310605102
10 Bacillus sp. NRRL B-14911	0.0	7.0	-1.0	0.0012319286675834906	0.0029511270087190043
11 Betaproteobacteria	0.0	7.0	-1.0	0.0012319286675834906	0.0029511270087190043
12 Acidobacteria (class)	10.0	0.0	-1.0	0.001953424961956904	0.006467648614090729
13 Bacteria	34.0	48.0	3.8395936048708474	0.0027917971899616596	0.010615068103361843
14 Bacillus	0.0	6.0	-1.0	0.002846180025106445	0.006756879033545369
15 Alteromonadales	9.0	0.0	-1.0	0.003441966434498408	0.012476931068326539
16 Proteobacteria	14.0	2.0	3.58596143401012	0.006058692572386619	0.018960959080629062
17 Sphingobacteria	0.0	5.0	-1.0	0.006575657299383542	0.01549026009730368
18 Bacillaceae	1.0	7.0	3.4147226954445307	0.00682157515930735	0.024548728642821524
19 Gammaproteobacteria	6.0	0.0	-1.0	0.01887824236538804	0.04003795071799761
20 Propionibacterineae	6.0	0.0	-1.0	0.01887824236538804	0.04003795071799761
21 Basidiomycota	5.0	0.0	-1.0	0.03328532206528786	0.07398224489977737
22 No hits	241.0	214.0	1.295172274966781	0.053065683723878404	0.1911222070672043
23 Bacteroidetes/Chlorobi group	5.0	9.0	1.247511546546185	0.0590791231214108	0.17299390038885112
24 Solibacteres	6.0	1.0	1.3692213193731104	0.07607649908441114	0.2494247661443111
25 Homo	3.0	6.0	0.9970726007673769	0.08339944441471772	0.18827620368199488
26 cellular organisms	23.0	25.0	0.7489823718943449	0.10841998824658877	0.23707576637440125
27 unclassified Bacteria	5.0	5.0	0.09104443974590981	0.32353542074135	0.753743469397794
28 Flavobacteriales	3.0	2.0	0.011054622756443422	0.476482612473364	0.9999999999997318

Figura 4.5- Comparação da representatividade de grupos taxonômicos entre Land1 (435 reads) e Land2 (570 reads). As estatísticas Stekel, Audic-Claverie e Fisher sinalizam em verde quando existe diferença significativa na representatividade.

Comparação da diversidade funcional entre as bibliotecas Land1 e Land2. A representatividade de função biológica nas bibliotecas Land1 e Land2 foi considerada equilibrada, como mostra a figura 4.6. Entretanto, como a representatividade taxonômica na biblioteca Land2 foi superior, o esperado era que a representatividade de funções biológicas também acompanhasse o resultado da primeira análise. Mas existe uma explicação molecular para isso; oligonucleotídeos de dez bases atuando sobre uma quantidade menor de genomas tende a anelar em um número maior de seqüências alvo dentro de uma mesmo genoma, e como resultado, explorando mais regiões e mostrando maiores diferenças na função biológica entre si do que quando atua sobre um conjunto maior de genomas (BIONDI et al., 2006). Isso porque a atuação sobre um conjunto menor de genomas permite os primers a anelarem em seqüências diferentes na medida em que se diminuiu a competição por sítios de anelamento mais abundantes (presentes numa mistura maior de genomas), resultando em maior representatividade de função biológica para os genomas que permanecem na mistura. Porém, isso não quer dizer, necessariamente, que a riqueza em nível de função biológica da biblioteca Land1 seria maior do que a Land2. Pois se fosse o

caso, a maior diversidade de função biológica estaria limitada a uma menor diversidade taxonômica (JOHNSON & SLATKIN, 2006). Isso, sem considerar os “no hits”.



Figura 4.6- Comparação da representatividade em nível de função biológica entre as bibliotecas Land1 e Land2. A abundância de genomas relacionados a determinada função biológica é diretamente proporcional à intensidade da cor vermelha.

Comparação das características taxonômica e funcional entre os solos de *landfarm* e Mata Atlântica (Land2 vs Mata1). Nessa comparação foram selecionadas 570 reads em Land2 e 221 em Mata1. Como esperado, existe uma diferença na diversidade taxonômica entre os solos de *landfarm* e Mata Atlântica. Dentre as diferenças encontradas, o solo de *landfarm*, por exemplo, é exclusivo em representar genomas da família Sphingobacteriacea (Audic-Claverie 2,068 e Fisher 1,130) e do gênero *Hydrogenivirga* (Audic-Claverie 3,843), mas o solo de Mata Atlântica possui mais genomas da classe Sphingobacteria, da qual a família Sphingobacteriaceae faz parte (Figura 4.7).

Outra particularidade interessante é o fato de o solo de *landfarm* ser exclusivo em genomas do grupo do *Bacillus cereus*. Neste grupo, como diz o próprio nome, está incluso o *Bacillus cereus*, um microrganismo patogênico oportunista capaz de causar β -hemólise e

intoxicações alimentares cujo quadro clínico inclui náuseas intensas e dores abdominais (KOTIRANTA et al., 2000; HOTON et al., 2005). Também no solo de *landfarm* foi encontrada uma quantidade consideravelmente maior de genomas de microrganismos pertencentes ao grupo da *Pseudomonas putida*; o qual é utilizado como biorremediador de solos impactados (GOMES et al., 2005). O solo de *landfarm* é um terreno que recebe resíduos (borra de petróleo, neste caso), onde também é aplicada a tecnologia ambiental *landfarming*, que serve para recuperar solos através da adição de microrganismos específicos (SANTOS, 2007).

No solo de Mata Atlântica, o gênero *Acinetobacter* foi representado de maneira exclusiva. Pertencentes à família *Moraxellaceae*, também exclusiva em Mata Atlântica neste trabalho, são importantes microrganismos de solo que contribuem para a mineralização de, por exemplo, compostos aromáticos. Também são patogênicos oportunistas, sendo resistentes a várias classes de antibióticos. E, além disso, possuem grande potencial biotecnológico como hospedeiros pela sua versatilidade metabólica e robustez (RAHAL, 2006; GERISHER, 2008).

Também no solo de Mata Atlântica, de maneira geral, existem mais genomas da ordem *Rhizobiales* do que no solo de *landfarm*. Isso acontece porque esses microrganismos são fixadores de nitrogênio e vivem em simbiose com raízes de plantas (GARCIA-FRAILE et al., 2007). Como a amostra de solo de Mata Atlântica provém de um ambiente mais rico em plantas, ao contrário do encontrado em solos de *landfarm*, mais uma vez as bibliotecas mostraram eficiência em representar a microbiota dos seus respectivos ambientes.



Figura 4.7- Comparação da diversidade taxonômica entre os solos de landfarm e Mata Atlântica (Land2 vs Mata1).

Na comparação funcional, houve diferença significativa em apenas dois itens; seqüências envolvidas com biosíntese de glutamato, mais presentes em Land2 (Stekel 5,226) e processos biológicos desconhecidos, mais presentes em Mata1 (Stekel 3,008). Essa maior quantidade de seqüências envolvidas com processos biológicos desconhecidos em Mata Atlântica nos leva a pensar num maior potencial para identificação de novas substâncias.

Determinação do impacto de número de seqüências e números de primers utilizados na comparação da diversidade taxonômica e funcional entre a biblioteca Land2 e o solo de Mata Atlântica (Land2 vs Mata Total) e Land Total e Mata Total. Nesta análise foram selecionadas 570 reads em Land2 e 1257 reads em Mata Total. O aumento do número de seqüências e primers em Mata Atlântica alteraram, como esperado, a análise. Na

representação de genomas de microrganismos do grupo do *Bacillus cereus*, por exemplo, o solo de Mata Atlântica passou a ser considerado superior (Stekel 3,958). Além disso, uma diferença na representação de genomas da ordem Enterobacteriales também passou a existir nessa nova análise, sendo o solo de Mata Atlântica muito superior nesse aspecto (Stekel 21,448; Audic-Claverie 9,918; Fisher 4,801).

Também em Mata Atlântica, uma grande quantidade de genomas pertencentes à ordem *Legionellales* foi adicionada, tornando a diferença na quantidade de genomas deste grupo taxonômico significativamente superior à de *landfarm* (Audic-Claverie 5,991; Fisher 1,177). Os “no hits” também foram encontrados em quantidade muito superior em Mata Atlântica (Stekel 5,057; Audic-Claverie 7,888) e de maneira geral, a diversidade taxonômica também cresceu (Figura 4.8).

Embora o aumento de seqüências e primers em Mata Atlântica tenha impactado a análise, a fidelidade na representatividade exclusiva do gênero *Hydrogenivirga* foi mantida. E houve significativa superioridade na quantidade de genomas da família *Sphingobacteriaceae*, presentes somente em *landfarm* na primeira análise entre os solos.

Na nova comparação da diversidade entre *landfarm* (1005 reads selecionados) e Mata Atlântica (1257 reads selecionados), onde a quantidade de reads é equilibrada, a diversidade em Mata Atlântica permaneceu maior. E a superioridade numérica em genomas de alguns grupos taxonômicos passou a ser considerada estatisticamente significativa em função do equilíbrio no número de seqüências entre os solos (dados não mostrados). Esse resultado deve-se ao fato de uma maior quantidade de primers distintos tender a amplificar seqüências mais diversas em nível taxonômico (YUTIN et al., 2005; McHARDY & RIGOTSOS, 2007) ou cobrir de maneira mais ampla o metagenoma em questão, resultando em maior diversidade taxonômica.



Figura 4.8- Comparação da diversidade entre Mata (Mata Total) e Land2. Quanto mais intenso a tonalidade da cor vermelha uma maior presença, enquanto que a preta a ausência. A tonalidade cinza avermelhada indica um baixa representatividade.

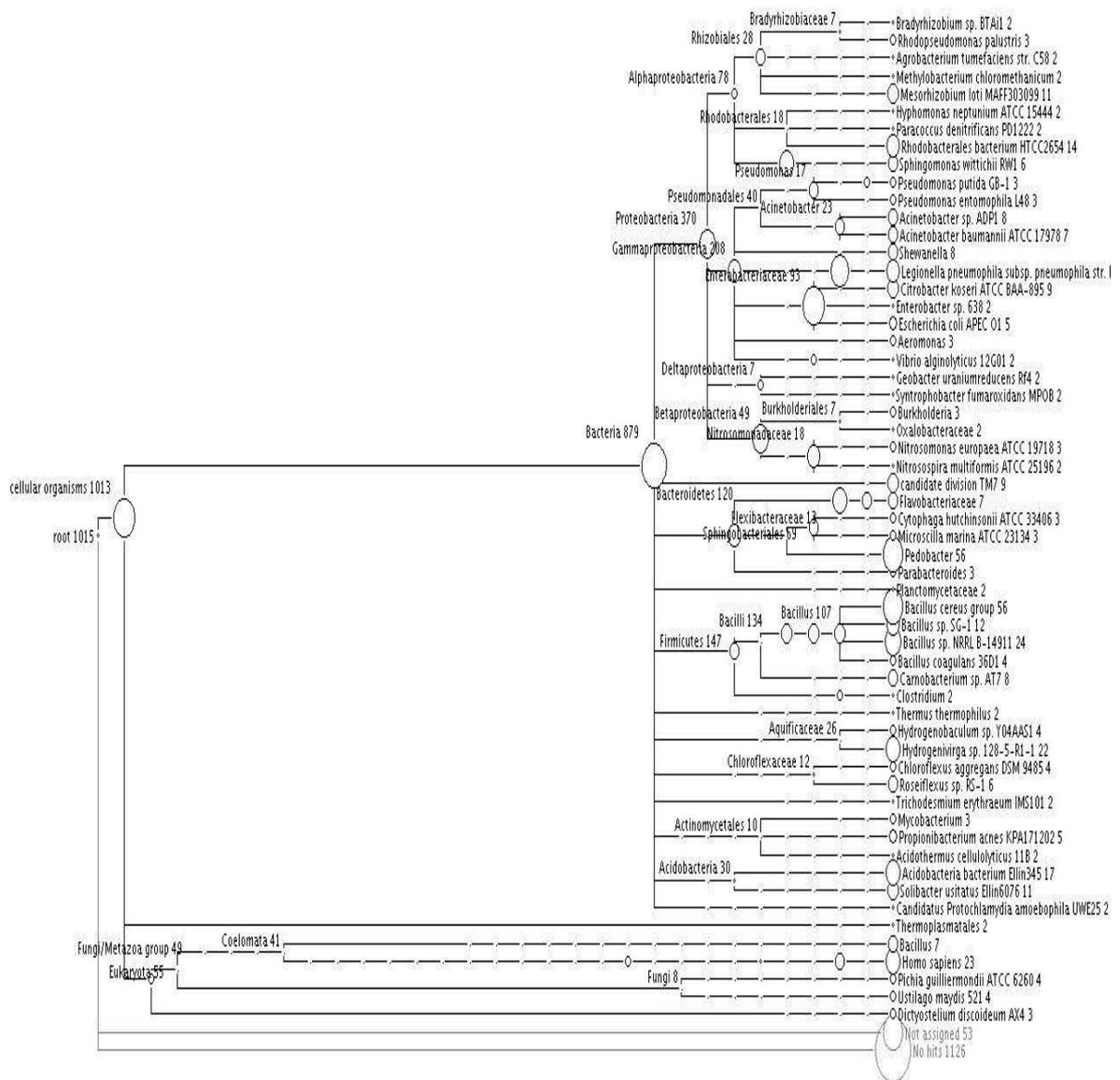


Figura 4.10- Diversidade taxonômica geral do solo de Mata Atlântica. As esferas próximas aos nomes dos grupos taxonômicos representam a abundância destes na amostra.

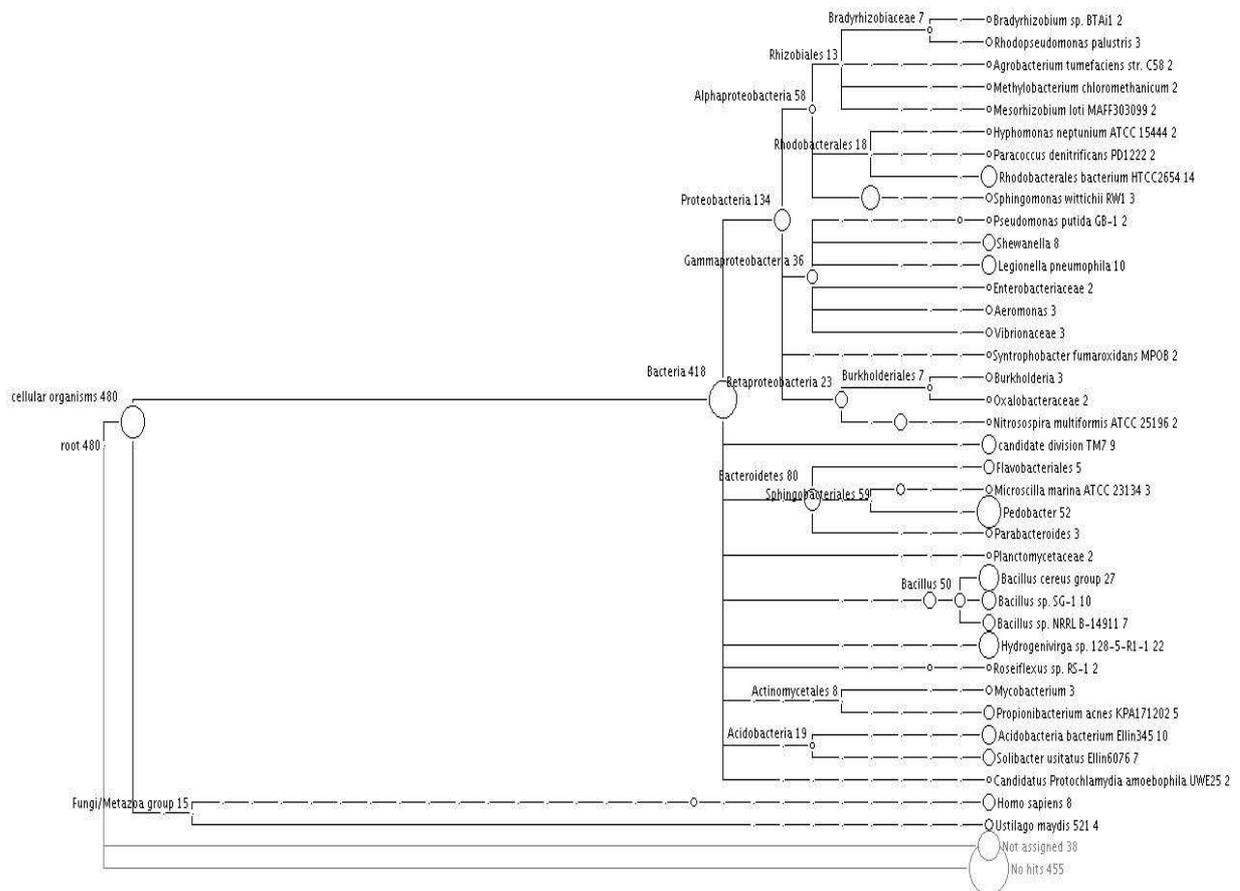


Figura 4.11- Diversidade taxonômica geral do solo de landfarm. As esferas próximas aos nomes dos grupos taxonômicos representam a abundância destes na amostra.

Tanto no caso de Mata Atlântica quanto de Landfarm, a representatividade e diversidade microbiana é superior a de trabalhos de estudos filogenéticos como o de Silva-Macena, em 2007, com 634 sequências 16S de solo de Mata Atlântica e a Santos, também em 2007, com 51 sequências 16S de *Landfarm*.

No caso de SILVA-MACENA, em 2007, com 634 sequências 16S de Mata Atlântica, foram representados os três filos; Firmicutes, Planctomicetos e Proteobacteria. Neste trabalho, foram representados os sete filos no mesmo solo; Firmicutes, Proteobacteria, Bacteroidetes *Deinococcus-thermus*, Aquificaceae, Chloroflexaceae, Acidobacteria e grupos taxonômicos ainda não classificados. Isso mostra que o método apresentado neste trabalho é superior em representar a microbiota ambiental, visto que se trata do mesmo solo e da mesma amostra. Pois, além de determinar os filos fornece informação sobre ORFs das

bactérias presentes, sendo possível traçar um perfil metabólico virtual dos microrganismos do solo.

No caso de SANTOS, em 2007, com 51 sequências de Landfarm, foram representados seis filos; Firmicutes, Bacteroidetes, Planctomicetos, Acidobacteria, Chloroflexiae e Proteobacteria. Neste trabalho, foram representados nove filos para o mesmo solo; Proteobacteria, Bacteroidetes, Planctomicetos, Firmicutes, Aquificaceae, Actinomicetos, Acidobacteria, uma divisão candidata e Candidatus Protochlamydia amoebophila, além de grupos ainda não classificados. Todos esses resultados e comparações reforçam a eficiência da técnica aqui apresentada.

Determinação do número de ORFs nas bibliotecas de Mata Atlântica e Landfarm.

Em Mata Atlântica, foram preditas 6481 ORFs potenciais, sendo que 1678 (26%) apresentaram hit contra o GenBank com e-value $\leq e^{-7}$ e em *Landfarm* foram previstas 3272, sendo que 524 (16%) apresentaram hit contra o GenBank com e-value $\leq e^{-7}$. Isto pode estar relacionado ao maior número total de sequências das bibliotecas de Mata Atlântica neste trabalho, sinalizando para um maior potencial em nível de bioprospecção de genes, neste momento (McHARDY & RIGOTSOS, 2007). Esses resultados mostram o potencial dessa metodologia em identificar genes. Deve-se ter em mente que certamente existem novos genes que foram detectados pela análise, mas que não encontraram hit no GenBank.

Conclusão

Foi possível amplificar sequências metagenômicas diretamente do sobrenadante da fervura do solo.

A associação do princípio direto de extração de DNA com a síntese *in vitro* de sequências metagenômicas livres de contaminação tornou a clonagem mais eficiente. E as bibliotecas metagenômicas construídas através desta nova metodologia neste trabalho representaram de maneira mais fiel seus respectivos ambientes. A nova metodologia permitiu também diminuir a interferência na representatividade dos grupos taxonômicos,

ampliando o potencial científico e tecnológico da metagenômica de solos. Agora, toda a diversidade presente em solos pode ser estudada, entendida e explorada de maneira mais ampla, mediante a aplicação de um maior número de primers.

O uso de uma maior quantidade de primers distintos aumenta consideravelmente a representatividade da diversidade taxonômica e funcional, permitindo um maior acesso à diversidade genética e funcional do ambiente.

Por fim, o tamanho das seqüências amplificadas torna possível até mesmo uma tentativa de bioprospecção de novos genes. E essa característica somada a uma maior representatividade do ambiente torna essa técnica uma nova e poderosa alternativa para entender e explorar metagenomas de solo.

Referências

ARAÚJO, E.S.; dos SANTOS, A.M.; AREIAS, R.G.B.M.; de SOUZA, S.R. and FERNANDES, M.S. USO DE RAPD PARA ANÁLISE DE DIVERSIDADE GENÉTICA EM ARROZ. *Agronomia*. v.37, p.33 - 37, 2003.

BERTRAND, H.; POLY, F.; VAN, V.T.; LOMBARD, N.; VOGEL, T.M. and SIMONET, P. High molecular weight DNA recovery from soils prerequisite for biotechnological metagenomic library construction. *Journal of Microbiological Methods* .Vol. 62, p. 1–11, 2005.

BIONDI, G.E; MENGONI, A. and BAZZICALUPO, M. (2006) “Analysis of Microbial Population Genetics” in VARMA, A.; ABBOTT, L.; WERNER, D. and HAMPP, R. (orgs.), *Plant Surface Microbiology*. Berlin: Springer Berlin Heidelberg, 551-565.

CLARK, J.M. Novel non-templated addition reactions catalyzed by prokaryotic DNA polymerases. **Nucleic Acids Research**. Vol. 16, p. 9677-9686, 1998.

CURTIS, T.P.; SLOAN, W.T.; SCANNELL, J.W. Estimating prokaryotic diversity and its limits, **Proceedings of the National Academy of Sciences USA**. Vol. 132, p.10494–10499, 2002.

DANIEL, R. The soil metagenome – a rich resource for the discovery of novel natural products. **Current Opinion in Biotechnology**. Vol. 15, p. 199–204, 2004.

DUBEY, S.K.; TRIPATHI, A.K and UPADHYAY. Exploration of soil bacterial communities for their potential as bioresource. **Bioresource technology**. Vol. 97 p. 2217–2224, 2006.

FERRER, M.; GOLYSHINA, O.; BELOQUI, A. and GOLYSHIN, P.N. Mining enzymes from extreme environments. **Current Opinion in Microbiology**. Vol. 10, p. 207–214, 2007

GARCIA-FRAILE, P.; RIVAS, R.; WILLEMS, A.; PEIX, A.; MARTENS, M.; MARTINEZ-MOLINA, E.; MATEOS, P.F. and VELAZQUEZ, E. *Rhizobium cellulosityticum* sp. nov., isolated from sawdust of *Populus alba*. **International Journal of Systematic and Evolutionary Microbiology**. Vol. 57, p. 844-848, 2007.

GERISHER, U. (2008). **Acinetobacter Molecular Biology**. 1ª edição. Caister Academic Press.

GOMES, N.C.M.; KOSHELEVA, I.A.; ABRAHAM, W.R, SMALLA, K. Effects of the inoculant strain *Pseudomonas putida* KT2442 (pNF142) and of naphthalene contamination on the soil bacterial community. **FEMS Microbiology Ecology**. Vol. 54, p. 21–33, 2005.

GREEN, P. Phrap, SWAT, CrossMatch. Available from the author. University of Washington, 1999.

HARRY, M.; JUSSEAUME, N., GAMBIER, B. & GARNIER-SILLAM, E.. Use of RAPD markers for the study of microbial community similarity from termite mounds and tropical soils. **Soil Biology and Biochemistry**, (33) pgs. 417–427, 2001.

HOTON, FM.; ANDRUP, L.; SWIECICKA, I. and MAHILLON J. "The cereulide genetic determinants of emetic *Bacillus cereus* are plasmid-borne." **Microbiology** 151 (7) 2121-4, 2005.

HUSON, D.H.; AUCH, A.F.;STEPHAN, J.Q. e SCHUSTER, C. MEGAN analysis of metagenomic data. **Genome Research**. Vol. 17. p. 377-386, 2007.

<http://www.fermentas.com/catalog/kits/kitinstaclone.htm>, acessado aos 11.01.2008

<http://www.repeatmasker.org>, acessado aos 24.01.2008

JOHNSON, P.L.F. and SLATKIN, M. Inference of population genetic parameters in metagenomics: A clean look at messy data. **Genome Research**. (112) pgs. 1320-1327, 2006.

KANT, A.D and TRIPLETT, E.W. Microbial communities and their interactions in soil and rhizosphere ecosystems.. **Annual Reviews in Microbiology**. Vol. 56, p. 211–36, 2002.

KIM, Y.J.; CHOI, G.; KIM, S.; YOON, G.; KIM, Y.S.; RYU, Y. Screening and characterization of a novel esterase from a metagenomic library. **Protein Expression and Purification**. Vol. 45, p. 315–323, 2006.

KOTIRANTA, A.; LOUNATMAA, K. and HAAPASALO, M.. "Epidemiology and pathogenesis of *Bacillus cereus* infections". **Microbes Infections**. Vol. 2, p. 189-98, 2000.

LAMMLE, K.; ZIPPER, H.; BREUER, M.; HAUER, B.; BUTA, C.; BRUNNER, H. and RUPP, S. Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. *Journal of Biotechnology* 127 (2007) 575–592.

McHARDY, A.C and RIGOTSOS, I. What's in the mix: phylogenetic classification of metagenome sequence samples. **Current Opinion in Microbiology**. Vol. 10, p. 499–503, 2007.

OGRAM, A.; SAYLER, G.S., BARKAY, T. The extraction and purification of microbial DNA from sediments. **Journal of Microbiological Methods**. Vol. 7, p. 57–66, 1987.

OH, H.J.; CHO, K-W.; JUNG, I-S.; KIM, W-O.; HUR, B-K. and KIM, G-J. Expanding functional spaces of enzymes by utilizing whole genome treasure for library construction. **Journal of Molecular Catalysis B: Enzymatic**. Vol. 26, p. 241–250, 2003.

OJO, O.A. Molecular strategies of microbial adaptation to xenobiotics in natural environment. **Biotechnology and Molecular Biology Review**. Vol. 1, p. 001-013, 2007.

PARK, H.J.; JEON, J.H.; KANG, S.G.; LEE, J-H.; LEE, S-A. and KIM, H-K. Functional expression and refolding of new alkaline esterase, EM2L8 from deep-sea sediment metagenome. **Protein Expression and Purification**. Vol. 113, p. 340–347, 2007.

RAHAL, J. Novel antibiotic combinations against infections with almost completely resistant *Pseudomonas aeruginosa* and *Acinetobacter* species. **Clinical Infectious Diseases**. 43 Suppl 2: S95-9, 2006.

RIESENFELD, C.S; SCHLOSS, P. and HANDELSMAN, J. METAGENOMICS: Genomic **Analysis of Microbial Communities**. **Annu. Rev. Genet**. Vol. 38, p. 525–552, 2004.

ROBE, P; NALIN, R.; CAPELLANO, C.; VOGEL, T.M.; SIMONET, P. Extraction of DNA from soil. **European Journal of Soil Biology** (39) pgs. 183-190, 2003.

RONDON, M.R.; AUGUST, P.R.; BETTERMANN, A.D.; BRADY, S.F.; GROSSMAN, T.H.; LILES, M.R.; LOICONO, K.A.; LYNCH, B.A.; MACNEIL, I.A.; MINOR, C.; TIONG, C.L.; GILMAN, M.; OSBURNE, M.S.; CLARDY, J.; HANDELSMAN, J.; GOODMAN, R.M. Cloning the Soil Metagenome: a Strategy for Accessing the Genetic and Functional Diversity of Uncultured Microorganisms. **Applied and Environmental Microbiology**. (115) p. 2541–2547, 2000.

SANTOS, A.C.F. Diversidade bacteriana do solo de landfarm da refinaria de petróleo Landulpho Alves- Ilhéus, Bahia. 2007. 94 p. (Dissertação de Mestrado). **Universidade Estadual de Santa Cruz**, Ilhéus, BA. 2007.

SCHLOSS, P. and HANDELSMAN, J. Biotechnological prospects from metagenomics. **Current Opinion in Biotechnology**. Vol. 14. p. 303–310, 2003

SHIZUYA, H., Birren, B., KIM, U.-J., MANCINO, V., SLEPAK, T., TACHIRI, Y., and SIMON, M... Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. **International Journal of Systematic and Evolutionary Microbiology**. Vol.89, p. 8794–8797, 1992.

SILVA-MACENA. Análise da biodiversidade metagenômica bacteriana de um solo de Mata Atlântica e isolamento de DNA de alto peso molecular visando aplicações biotecnológicas. - Ilhéus, Bahia. 2007. (Dissertação de Mestrado). **Universidade Estadual de Santa Cruz**, Ilhéus, BA. 2007.

STREIT, Wolfgang R. and SCHMITZ, Ruth A. 2004. Metagenomics – the key to the uncultured microbes. **Current Opinion in Microbiology**. Vol. 7, p. 492–498, 2004.

TRINGE, S.G.; von MERING, C.; KOBAYASHI, A.; SALAMOV, A.A.; CHEN, K.; CHANG, H.W.; PODAR, M.; SHORT, J.M.; MARTHUR, E.J.; DETTER, J.C.; BOORK, P.; HUGENHOLTZ, P. and RUBIN, E.M. Comparative Metagenomics of Microbial Communities. **Science**. Vol. 135, p. 554-557, 2005.

VROH, B.I., HARVENGT, L., CHANDELIER, A., MERGEAI, G.; and JARDIN, P.D. Improved RAPD amplification of recalcitrant plant DNA by the use of activated charcoal during DNA extraction. **Plant Breeding**. Vol. 115, p. 205–206, 1996.

WHITEHOUSE, C.A and HOTTEL, H.E. Comparison of five commercial DNA extraction kits for the recovery of *Francisella tularensis* DNA from spiked soil samples. **Molecular and Cellular Probes**. Vol. 21, p. 92–96, 2007.

WILSON, I.G. Inhibition and Facilitation of Nucleic Acid Amplification. **Applied and Environmental Microbiology**. Vol. 120 p. 3741–3751, 1997.

YUN, J.; KANG, S.; PARK, S.; YOON, H.; KIM, M.; HEU, S. and RYU, S. Characterization of a Novel Amylolytic Enzyme Encoded by a Gene from a Soil-Derived Metagenomic Library. **Applied and Environmental Microbiology**. Vol. 12, p. 7229–7235, 2004.

YUTIN, N.; SUZUKI, M.T. and BÉJÁ, O. Novel Primers Reveal Wider Diversity among Marine Aerobic Anoxygenic Phototrophs. **Applied and Environmental Microbiology**. Vol. 12, p. 8958–8962, 2005.

5. Capítulo 3- Metagenoma de Mata Atlântica como fonte de substâncias antimicrobiana.

AMORIM, J.H.¹; SILVA-MACENA, T.N.²; COSTA, M.M.C.³; VIDAL, R.O.²; VILELA-DIAS, C.²,
LACERDA-Jr.⁴, G.V.; de OLIVEIRA, J.M.⁴; PIROVANI, C.P.⁵; REZENDE, R.P.⁵; DIAS, J.C.T.⁵;
CASCARDO, J.C.M.⁵

- 1- *Estudante do Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Estadual de Santa Cruz (UESC), Ilhéus, Bahia- Brasil;*
- 2- *Mestre em Genética e Biologia Molecular;*
- 3- *Pesquisador da EMBRAPA/CENARGEN;*
- 4- *Estudante de Graduação no Bacharelado em Ciências Biológicas da UESC;*
- 5- *Docentes do Departamento de Ciências Biológicas da UESC*

Resumo

Muitos microrganismos patogênicos apresentam resistência a antibióticos disponíveis no mercado atualmente. Além disso, os microrganismos solo pode servir como um reservatório de mecanismos de resistência que já emergiram ou têm o potencial de emergir em bactérias de significativa importância clínica, representando um perigo para a saúde pública. A metagenômica é uma importante alternativa para contribuir na solução deste problema porque permite explorar o potencial biotecnológico também de microrganismos não cultiváveis, permitindo o acesso a novos antimicrobianos e enzimas. Nesse trabalho são apresentados os resultados de ensaios de clones de duas bibliotecas metagenômicas, onde um clone apresentou atividade antimicrobiana contra *B. subtilis* e *S. aureus*, mas não contra *E. coli* e *Salmonella*. O estudo da seqüência clonada revelou duas ORFs conhecidas completas. No *contig R*, destacamos uma ORF codificando para uma proteína com sítio ativo de glutamato sintase, como também ORFs menores, sem homologia com seqüências conhecidas. E no *contig F*, destacamos outra codificando para uma proteína que codifica para possível peptídeo sinal, o qual pode estar direcionando a secreção do antimicrobiano. Também foi comprovado que o antimicrobiano tem tamanho maior igual a 5KDa, é secretado e sua expressão pode ser controlada por indução com IPTG.

Palavras-chave: *Antimicrobiano, metagenômica e bioprospecção.*

Introdução

A maioria dos antibióticos utilizados na clínica provém de Actinomicetos presentes no solo (KIESER et al., 2000). Mas o potencial de prover novos agentes antimicrobianos não está concentrado apenas nesta classe de bactérias, pois também nos solos existe uma intensa disputa por espaço e nutrientes que leva os microrganismos a secretarem antimicrobianos contra seus competidores (WRIGHT, 2007). E esses processos envolvem elementos de resistência que quase sempre estão agrupados nos mesmos operons biosintéticos de antibióticos (COSTA et al., 2006). Desta forma, a riqueza em diversidade genética e funcional do solo pode ser acessada através de técnicas de metagenômica para identificar novas substâncias antimicrobianas, o que é extremamente necessário nos dias atuais, quando a medicina moderna enfrenta um de seus mais desafiadores problemas; os *superbugs*, bactérias resistentes a múltiplos antibióticos (WRIGHT, 2007). A presença de antibióticos no ambiente tem promovido à aquisição ou evolução independente de elementos de resistência específicos na ausência da produção de antibióticos inatos, como a resistência à vancomicina em *Streptomyces coelicolor*, *Paenibacillus*, e *Rhodococcus* (HONG et al., 2002; GUARDABASSI & DALSGAARD, 2004). Desta forma, o solo pode servir como um reservatório de mecanismos de resistência que já emergiram ou têm o potencial de emergir em bactérias de significativa importância clínica (COSTA et al., 2006). Isso constitui um perigo para a saúde pública, que pode contar com uma quantidade cada vez menor de opções em antibióticos para tratar infecções. Some-se a isto, a seleção de microrganismos resistentes provocada pelos abandonos de tratamentos, os quais induzem falha clínica por diminuírem a biodisponibilidade dos antibióticos (WALSH, 2003; WRIGHT, 2007).

A busca de novos antimicrobianos através da metagenômica torna-se uma importante alternativa porque permite explorar o potencial biotecnológico também de microrganismos não cultiváveis (DANIEL, 2004). E além de permitir o acesso a novos fármacos, a metagenômica também tem o potencial de promover o acesso a novas enzimas com potenciais aplicações industriais (LAMMLE et al., 2007).

Na maioria das vezes, a preferência é clonar seqüências com o máximo tamanho possível para aumentar a probabilidade de encontrar operons e vias metabólicas funcionais inteiros (BERTRAND et al., 2005). Mas também é possível promover bioprospecção de novos genes clonando-se seqüências menores. E neste trabalho, nós apresentamos uma

busca por atividades antimicrobiana e enzimática em bibliotecas metagenômicas pTZ57R/T de solos de *landfarm* e Mata Atlântica, como também o estudo de um clone de comprovada atividade antimicrobiana cuja seqüência clonada contém pouco mais de 1500 pares de base e foi proveniente de solo de Mata Atlântica.

Metodologia

Screenings enzimáticos. Em placas de Petri grandes contendo agar Mueller-Hinton e IPTG (isopropil- β -D- tiogalactopiranosídeo) 50 mM, inoculou-se os clones, em arranjo de 96 poços de placa de poliestireno, a meios contendo uma única fonte de energia; óleo vegetal a 2% para screening de lipases, caseína a 2% para screening de proteases e óleo mineral a 2% para screening de alcano-desidrogenases. A composição do meio, além da única fonte de energia e ágar a 1,5%, continha KH_2PO_4 a 0,1%, KHPO_4 a 0,1 %, NH_4Cl a 0,1%, MgSO_4 a 0,05% FeSO_4 a 0,001% e CaCl_2 a 0,001%. Este procedimento foi aplicado a 2304 clones de metagenoma de Mata Atântica e 1728 clones de metagenoma de *landfarm*.

Screenings para antimicrobianos. Em placas de Petri grandes contendo ágar Mueller-Hinton e IPTG 50 mM, inoculou-se 200 μL de suspensão de *Staphylococcus aureus* ATCC (10^6 UFC/mL em solução salina a 0,9%) e aplicou-se a técnica “spread plate”. Em seguida, inoculou-se os clones por cima do revelador (*Staphylococcus aureus* ATCC- 25923 no primeiro teste, *Bacillus subtilis* ATCC-168, *Escherichia coli* CCT-1371 e *Salmonella* subsp. *enterica* serovar Newport no segundo teste) conforme arranjo de 96 poços de placa de Elisa. Incubou-se por 16 horas a 37° C e observou-se quanto à formação de halo de inibição de crescimento em volta dos clones. Este procedimento foi aplicado a 2304 clones de metagenoma de Mata Atântica e 1728 clones de metagenoma de *landfarm*. Todos os clones foram obtidos por meio da utilização do vetor pTZ57R/T, segundo instruções do fabricante (InstAclone PCR cloning Kit -Fermentas).

Manipulação do clone Mata II PL08 B3. O inserto do clone MTII PL08 B3 foi amplificado através de PCR da colônia do *screening* para antimicrobianos utilizando-se

primers M13. Em seguida sequenciou-se o produto de reação em Megabace 1000, “forward” (F) e “reverse” (R) e montou-se os consensos no Lasergene. Depois disso, a seqüência foi analisada em buscas por ORFs (Open reading frames), utilizando-se o programa ORF FINDER (<http://www.ncbi.nlm.nih.gov/projects/gorf/>). Para uma das prováveis proteínas, calculou-se o número de aminoácidos, massa molecular, ponto isoelétrico teórico, fórmula e índice de instabilidade através do ProtParam do Expsy Proteomics Tools.

Verificação quanto ao controle de indução por IPTG. O clone MTII PL08 B3 foi testado quanto à capacidade de controle da expressão do antimicrobiano por IPTG. Para isso, o clone foi inoculado em 5 mL de meio Luria-Bertani (LB) com ampicilina (100 mg/mL) e incubado “overnight” a 37°C sob agitação. Em seguida, centrifugou-se a cultura a 5000 rpm por 10 minutos e o precipitado de células desta primeira cultura foi inoculado em 30 mL de meio líquido LB com ampicilina (100 mg/mL) e novamente incubado a 37°C sob agitação. Manteve-se essa incubação até que a cultura atingisse a OD_{600 nm} de 0,7 e induziu-se a expressão com IPTG (concentração final de 0,1 mM), tomando como referência uma cultura sem indução por IPTG. Após a indução, centrifugou-se a cultura a 5000 rpm por 10 minutos e separou-se o precipitado de células do sobrenadante em todos distintos de 50 mL. Parte do sobrenadante foi testado contra *B. subtilis* nos volumes de 1, 2, 4 e 8 µL.

Verificação quanto ao tamanho da molécula do antimicrobiano e seu local de síntese (intra ou extracelular). Os sobrenadantes da etapa anterior foram filtrados em membrana de 0,22 µm, liofilizados, resuspenso em 5 mL tampão fosfato 1x (pH 7,0), centrifugados a 13400 rpm por 10 minutos e os sobrenadantes desta última centrifugação tratados com ácido tricloroacético a 10 % e acetona com betamercapto a 0,007 %. Em seguida, foram filtrados em Centricon de 5 KDa (Millipore) para separar, no sobrenadante, moléculas maiores (Sob. ≥ 5 KDa) e menores (Sob. ≤ 5 KDa) que 5 KDa.

Os precipitados de células foram resuspenso em tampão fosfato 1x (pH 7,0), tratados com lisozima e expostos à sonicação em Ultrasonic Processor GEX 130, com pulsos de 10 segundos, intervalos de 20 segundos, num total de 1 minuto e 10 segundos, com amplitude de 70%. Em seguida, parte das suspensões do extrato total de células (ET). Fotam coletadas e reservadas. Após isso, centrifugou-se o ET a 13400 rpm por 10 minutos;

e os sobrenadantes e os precipitados desta última centrifugação deram origem, respectivamente, à fração solúvel (FS) e fração insolúvel (FI). A fração solúvel foi filtrada em Centricon de 5 KDa (Millipore) para separar moléculas maiores ($FS \geq 5$ KDa) e menores ($FS \leq 5$ KDa) que 5 KDa. Em seguida, um bioensaio em ágar Mueller-Hinton e um inóculo de 200 μ L de solução salina (0,9%) com *B. subtilis* 10^6 cél./mL foi preparado utilizando aproximadamente 1 μ g de proteínas totais do $FS \leq 5$ KDa e do $FS \geq 5$ KDa para difusão em agar. Os Sob. ≤ 5 KDa e Sob. ≥ 5 KDa também foram utilizados no bioensaio, com aproximadamente 70 μ g de proteínas totais para difusão em agar (o controle não induzido foi utilizado como referência para testar sobrenadante e fração solúvel). Em seguida, o ensaio foi incubado a 37°C por 16 horas.

Verificação quanto à expressão diferencial de proteínas por SDS-PAGE. O sobrenadante, ET, FS e FI do clone MTII PL08 B3 foram analisados quanto à expressão diferencial de proteínas quando induzido por IPTG. Para isto, uma eletroforese em gel de poliacrilamida a 18% foi realizada a 140 V por 1 hora e 40 minutos, segundo os procedimentos de SDS-PAGE descritos por AMES, em 1974.

Resultados e Discussão

Screenings enzimáticos. Nenhuma atividade enzimática foi encontrada nos experimentos feitos com os clones de *landfarm* e Mata Atlântica para as análises testadas.

Screenings para antimicrobianos. Foi encontrada atividade antimicrobiana contra *Bacillus subtilis* e *Staphylococcus aureus* (Figuras 1 e 2, respectivamente) de um clone (B3; Mata II PL08) de metagenoma de solo de Mata Atlântica .

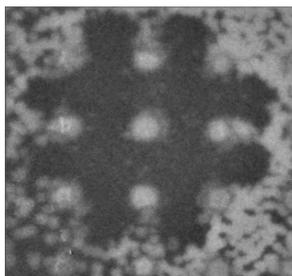


Figura 5.1- O clone B3:MTII PL08, no centro, inibiu de maneira significativa o crescimento do revelador, *Bacillus subtilis* ATCC-168. Notar que o halo de inibição chega a invadir a área de outros clones no agar, sem, no entanto, inibir o crescimento dos mesmos.

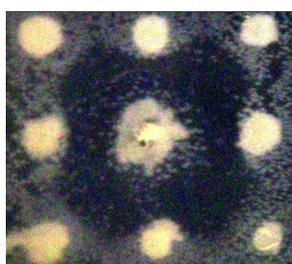


Figura 5.2- O clone B3:MTII PL08, no centro, inibiu de maneira significativa o crescimento do revelador, *Staphylococcus aureus* ATCC-25923. Notar que o halo de inibição de crescimento deste revelador é menor em comparação com o do primeiro revelador, o *Bacillus subtilis* ATCC.

Para que o antimicrobiano tenha se difundido no agar a ponto de provocar a ação inibitória mostrada, é necessário que o mesmo tenha sido secretado em grande quantidade pela hospedeira (*Escherichia coli* TOP 10). Mesmo assim, o antimicrobiano não foi capaz de inibir o crescimento dos outros clones de *E. coli*, visto que chegou a invadir a área do mesmo no agar. Isso pode estar relacionado a diferenças na constituição da parede celular entre as hospedeiras, gram-negativas e os reveladores, gram-positivos; como estes últimos foram afetados e não as hospedeiras gram-negativas, o antimicrobiano em questão pode estar interferindo nos mecanismos de síntese da peptidoglicana, muito mais abundante em parede celular de gram-positivos. Além disso, a membrana externa das gram-negativas pode ser uma barreira bioquímica a essa substância (MEDIGAN & MARTINKO, 2005). Esta idéia é reforçada pelo fato de não ter acontecido inibição contra *E. coli* CCT (Figura 5.3) e o clones não terem crescido na presença de *Salmonella enterica*.

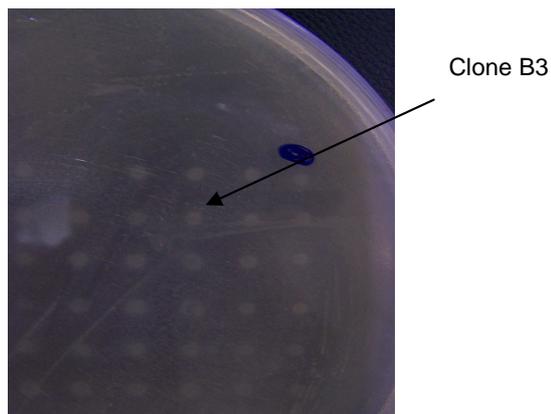


Figura 5.3- Bioensaio de clones contra *E. coli* CCT. Notar que em volta da colônia do clone B3 não há halo de inibição de crescimento do revelador.

Após indução do clone com IPTG, o ensaio do sobrenadante da cultura foi positivo em provocar inibição do crescimento de *S. aureus*. Esse resultado reforça a idéia de que o antimicrobiano é secretado em grande quantidade e o(s) gene(s) responsáveis por sua síntese estão sob a orientação do promotor T7 do vetor.

Análise do clone Mata II PL08 B3. O estudo da seqüência clonada mostrou, inicialmente, um tamanho um pouco maior que 1500 pares de bases (Figura 5.4).

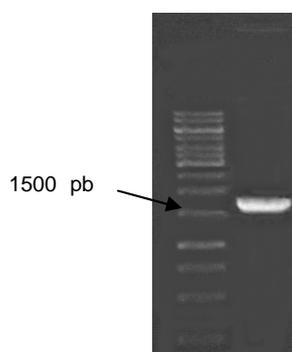


Figura 5.4- Eletroforese do produto de PCR de colônia do clone Mata II PL08 B3 utilizando o primer M13, que anela no múltiplo sítio de clonagem e pode amplificar o inserto. Notar a banda com pouco mais de 1500 pares de base referente ao tamanho da seqüência clonada.

Nos dois contigs montados (Figura 5.5), duas ORFs potenciais foram encontradas; uma com homologia com um peptídeo sinal e outra com homologia com uma glutamato sintetase. Além disso seqüências aparentemente TATA box e Shine-Dalgarno foram destacadas, tornando mais aparente a expressão destas ORFs. Os aparentes TATA box evidenciam a existência de promotores, enquanto as seqüência Shine-Dalgarno evidenciam

sítios de ligação ao rRNA 16S, necessária para acoplar o RNA mensageiro ao complexo ribossomal quando da síntese protéica (SHINE & DALGARNO, 1975; SMALE & KODONAGA, 2003).

1. Contig F com 793 bases

CCCCCGGTkTGaACGACGGCAGTGAWTCGAGCTgCkGTGTaACTCGGAATGCTCTAGAWTCAGCGCCATGTAAATGCAAACCATGAAG
 CCGAGTTTCGTGGTACCAATATGTTACTCAA**TTTAAAAAAT**CAGTATGCGATTGAGCTATTTCCGGTCTCTGATGAAATTAT**TATTTAAAAA**
 CTTCTTAAGAAAACAGATAGATCGTAATAAATTATCTACTTGCCTGTCT**AGCCGA**AAAT**ATG**CCTGAGCAATATGTTTACTCTATGGTT
 CTTATAAATCAGCCAATCAAGCAAAGCAGGCAGCTGGATCATTTGAATATGGATTTGCCAAAATCAGTCAAACCAGCTGTGCAGCAAATTCAG
 GGATATCAACCTTATGTGAATAACTTAGTCTGATGAGCTGACGCTTMACCAGAGACTTATCAAGTACGATTAATAAATGTGCCTCTA
 CCAAAAGTAGAGGAAATAGTAGTTCCTAAACCGCAGAGCCAGCCCTCATCAAACAGCCAAGTACCACCAGTACAACCATTGtTCGTCGTG
 ATCCATCGGGGAATGTAGTAGATGTGCAGAAGTCTGAAAGCACTGTACCGTCAATGCCACGTGCTCAACCTAATGGTCAACAACGTGCTGA
 AGAACAGAGATAAGTATCCTTTTAACT**TA**TACAAATTTGGGTGACAGGCTAAAAATAGCGCACCCCTGTATGAAAAATACTATcAAaTGTG
 CGTTTTTGGTGC AAAAATACACAATAAcAAATTATCACaTCTGTTCCTAAAATTTGCTAAATCT

5'3' Frame 3
 P R X E R R Q S X E L X C N S R M L S X Q R H C K C K P S S R V S W Y Q I C Y S I S K S V C
 D S A I S G L S S N Y Y S K L L K K T D R S S St I I Y L R L S G E N Met P E Q Y V L L Y G
 S Y K S A N Q A K Q A L D H L N Met D L P K S V K P A V Q Q I Q D Y Q P Y V N N L G S D E L
 T S X Q R L Y Q V R L K N V P L P K V E E I V V P K P Q S Q P S S N T P S T T S T T I V R R
 D P S G N V V D V Q K S E S T V P S Met P R A Q P N G Q Q R A E E Q E I S D P F N Stop Y K
 F G C R L K Stop R T L Y E K I L S N V R F W C K K Y T I T N Y H I C S Stop N L L N

Signal peptide

245 atgcctgagcaatgatgtttactctatggttcttataaatcagcc
 M P E Q Y V L L Y G S Y K S A
 290 aatcaagcaaacgagcactggatcatttgaatatggatttgcca
 N Q A K Q A L D H L N M D L P
 335 aatcagtcaaaccagctgtgcagcaaatcaggattatcaacct
 K S V K P A V Q Q I Q D Y Q P
 380 tatgtgaataacttaggttctgatgagctgacgtctmaccagaga
 Y V N N L G S D E L T S X Q R
 425 ctctatcaagtacgattaaaaaatgtgcctctacaaaagttagag
 L Y Q V R L K N V P L P K V E
 470 gaaatagtagttcctaaaccgagagccagccctcatcaaacacg
 E I V V P K P Q S Q P S S N T
 515 ccaagtaccaccagtacaaccattgttcgctgatccatcgggg
 P S T T S T T I V R R D P S G
 560 aatgtagtagatgtgcagaagtctgaaagcactgtaccgtcaatg
 N V V D V Q K S E S T V P S M
 605 ccacgtgctcaacctaatggtcaacaacgtgctgaagaacaagag
 P R A Q P N G Q Q R A E E Q E
 650 ataagtatccttttaactaa 670
 I S D P F N *

2. Contig R 515 bases

CTCTAATACGACTCACTATAGGGAAAGCTTGCATGACGGCCTCTGCAGTCGACGGGCCCGGATCCGA**TTG**GCGCCATTGTTGGTGGGGa
 CCACAGCCATGCGATAACCCGGCAACCTTCTTCTCAATTTCTTTGGTTAATACTTGTTTTGGCTGTGCTGGCAGTGCAGGATCCAGGTT
 CAGGAAGATCGTACCTACAGCAAAAATTTTCGCTTAAGGTAATATCACTTAAACGCTTTGCTTCTTACGAAAGGATTCCTTTGGCATTGCC
 AAAAGCAAACCACAACCATCCCGGTCTTACCATCTGCGGAATACCACCAGGTGGGTGCATACAACCTTAAACTATGAATGGCGTCTTGA
 CGAGATCATGGCTTGCATCACCCTGCATATGGGGATCAAACCGAAACACAGTTACCTTTAAACTCATCCGGTTGGTATAAACCTTGAGC
 GGGAGCTACATTATAGGCGATGGCATGTG**CAT**AGCGT**ACCCTT**CTTATCACAAGCCCTGGCGGGCTG**TTAAACAATCTAAAT**CCGTCTT
 TGCATGCGCTCAAATTTACTTTTCGAGGAAATTC

3'5' Frame 3
 N F L R K Stop I W T P S Q R R I Stop I V Stop Q P A R A L Stop Stop E G Y A Met H Met P S
 P N N V A P A Q G L Y Q P D E F K G N C G F G L I A H Met Q G D A S H D L V K T A I H S L S
 C Met T H R G I A A D G K T G D G C G L L A Met P K E S F R E E A K R L S D I T L S E I
 F A V G T I F L N L D P A L P A H A K Q V L T K E I E E G C R V I A W R V V P H Q Q W R
 Stop S D P G P V D C R G L H A S F P Y S E S Y Stop S L A Stop S W S Stop L F P V Stop N C
 Y P L T I P H N I R A G S I K C K A W G A Stop Stop V S Stop L T L I A L R S L P A F Q S G

N L S C Q L H Stop Stop I G Q R A G R G G L R I G R S S A S S L T D S L R S V V R L R R A V
 S A H S K A V Y G Y P Q N X G D N A A K E H V S T K A S K K A R E P V K R P R C W P F S

Glutamate synthase

```

915 atgcacatgcccacatgcccataataatgtagctcccgctcaaggttta
    M H M P S P N N V A P A Q G L
870 taccaaccggatgagtttaaaggtaactgtggttccggttgatc
    Y Q P D E F K G N C G F G L I
825 gcccatatgcagggatgcaagccatgatctcgtcaagaccgca
    A H M Q G D A S H D L V K T A
780 attcatagtttaagttgatgaccacggtggtattgcccga
    I H S L S C M T H R G G I A A
735 gatggttaagaccgggatggttggttggctttggcaatgcca
    D G K T G D G C G L L L A M P
690 aaggaatccttccgtgaagaagcaaagcgtttaagtgatattacc
    K E S F R E E A K R L S D I T
645 ttaagcgaatttttgctgtaggtacgatcttccgtaacctggat
    L S E I F A V G T I F L N L D
600 cctgcatgcccagcacacgcaaaacaagtattaaccaagaatt
    P A L P A H A K Q V L T K E I
555 gaagaagaaggttgccgggttatcgcatggcgtggttccccac
    E E E G C R V I A W R V V P H
510 caacaatggcgcTaa 496
    Q Q W R *
  
```

Figura 5.5- Contigs F (1) e R (1) e as ORFs com homologia com peptídeo sinal e glutamato sintetase, respectivamente. Nos contigs, em azul-claro, estão destacados os possíveis TATA box; em rosa, as possíveis seqüências Shine-Dalgarno; em amarelo, os códons de iniciação e em verde os códons de finalização. As seqüência de nucleotídeos das ORFs de peptídeo sinal e glutamato sintetase são acompanhadas das seqüências de aminoácidos das proteínas.

Somados, os contigs têm 1308 bases, com isso existe a possibilidade, ainda que pequena, de existir uma terceira ORF presente nesse contig maior. Os primers necessários para o fechamento desse contig maior já foram desenhado e sua chegada está sendo aguardada.

No consenso F da seqüência clonada, está a ORF com 426 pares de base contendo homologia com peptídeo sinal e que pode estar envolvida com a atividade antimicrobiana, cujo “e-value” do Blast foi de 2^{e-23} (ver Figuras 5.5 e 5.6), a qual apresenta um possível sítio de ligação ao rRNA 16S, homólogo a uma sítio de ligação ao rRNA de uma bacteriocina dependente de peptídeo sinal em *Carnobacterium divergens* (WOROBO et al., 1995).

Os peptídeos sinais estão envolvidas em vias secretórias, direcionando o envio de outras substâncias para fora da célula (SUTCLIFFE & HARRINGTON, 2002; KANDAVILLI et al., 2008). A estratégia celular de utilizar peptídeos sinais para facilitar a movimentação de proteínas através da membrana é um processo conservado em todas as formas de vida (BLOBEL & SABATINI, 1971; KIM et al., 2008). Desta forma, é possível que este peptídeo sinal esteja direcionando a secreção do antimicrobiano pela hospedeira (*E. coli* TOP10). Além disso, trabalhos recentes mostram que muitas vezes, em

procaríotos, as seqüências que codificam para o peptídeo sinal estão agrupadas em operons dos quais a seqüência que codifica para a substância a ser secretada também faz parte. Mas para que o processo de secreção aconteça de maneira completa, é necessário que uma enzima denominada peptidase do peptídeo sinal clive o peptídeo sinal, o qual está ligado à substância a ser secretada, liberando-a para o meio extracelular (NOVAK & DEV, 1988; PENG et al., 2008). Como essa estratégia de secreção é conservada em todos os organismos, é possível que uma peptidase do peptídeo sinal da hospedeira esteja clivando o peptídeo sinal e permitindo a liberação do antimicrobiano para o meio extracelular. Além disso, as outras ORFs encontradas são “no hits” quando comparadas via Blast-x com seqüências do NCBI, o que torna possível estar acontecendo a expressão de alguma substância ainda desconhecida.

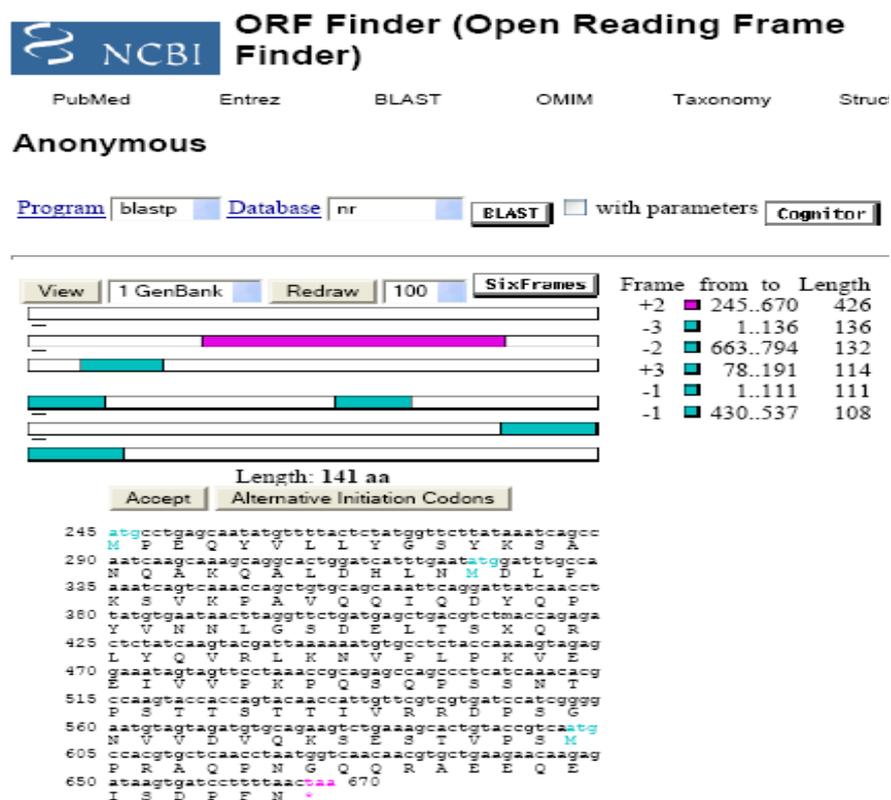


Figura 5.6- Busca por ORFs dentro do consenso F, com 793 pares de bases do clone MTII PL08 B3. A ORF marcada em rosa está completa, possui 426 pares de base, está na “frame” +2 e tem homologia com peptídeo sinal. Sua tradução resultaria numa proteína de 141 aminoácidos (peptídeo sinal). As outras ORFs (menores e marcadas em azul) são “no hits” quando comparadas através de Blast-x com as seqüências do NCBI.

Na figura abaixo, destacamos a ORF de uma possível glutamato sintetase. Em enzimologia, a glutamato sintase é uma enzima que catalisa a seguinte reação:



Ela participa do metabolismo de glutamato e nitrogênio (MOSS, 2006). Mas encontramos, na seqüenciada proteína, um domínio com padrão similar ao de enzimas que participam de síntese de antimicrobianos, não descartando a possibilidade de ter acontecido a clonagem de um operon completamente envolvido com a síntese e secreção deste antimicrobiano.

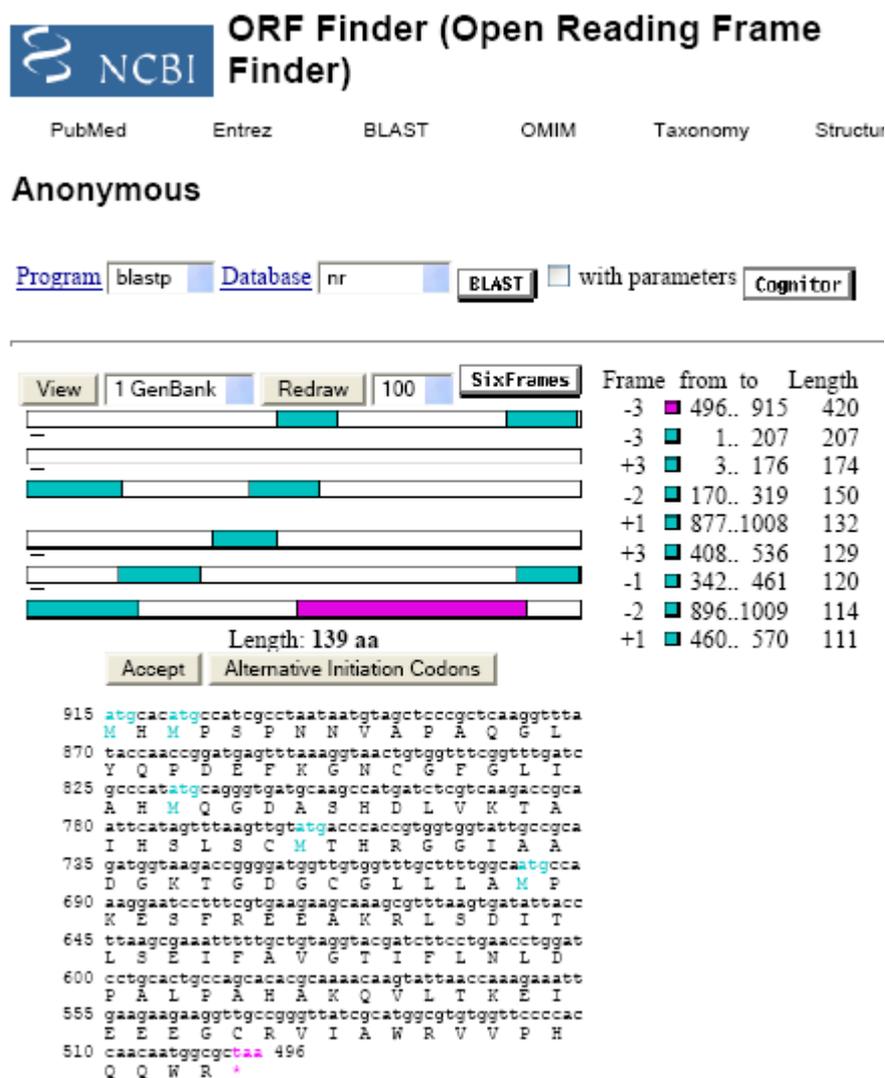


Figura 5.7- Busca por ORFs dentro do consenso R, com 515 pares de base do clone MTII PL08 B3. A ORF marcada em rosa está completa, possui 420 pares de base, está na “frame” -3 e tem homologia com uma proteína com domínio de glutamato sintetase. Sua tradução resultaria numa proteína de 139 aminoácidos.

Verificação quanto ao controle de indução por IPTG, tamanho da molécula do antimicrobiano e expressão diferencial de proteínas. Foi demonstrado que a expressão do antimicrobiano é controlada por indução com IPTG. Além disso, este antimicrobiano é secretado e tem tamanho maior ou igual a 5 KDa, pois no bioensaio contra *B. subtilis*, o sobrenadante da cultura induzida provocou inibição significativa do crescimento do revelador quando aplicado e difundido no agar (Figura 5.8).

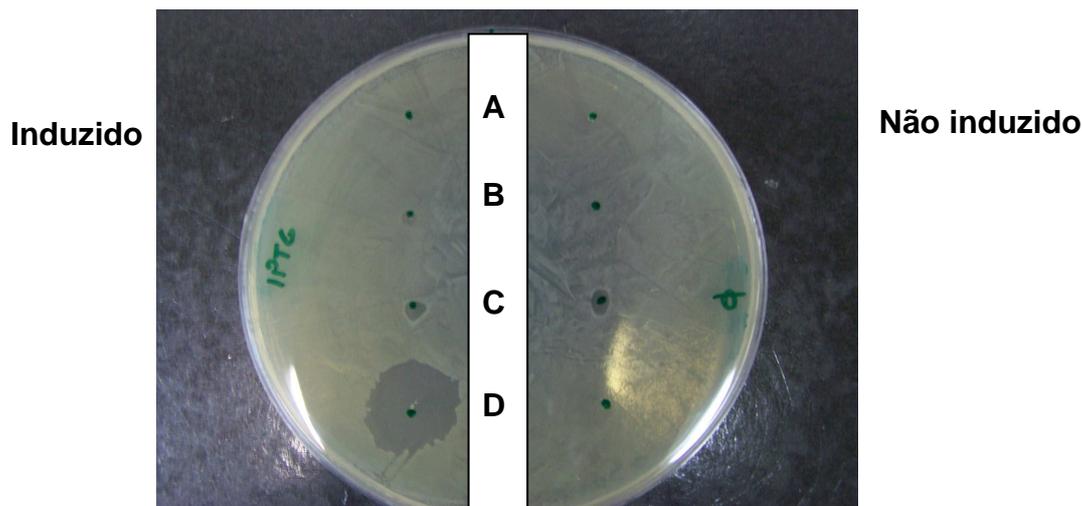


Figura 5.8- Bioensaio contra *B. subtilis* ATCC-168 para verificar indução de expressão do antimicrobiano por IPTG e a faixa de seu tamanho em KDa. **A-** fração solúvel do precipitado de células com moléculas ≤ 5 KDa ($\approx 70 \mu\text{g}$); **B-** fração solúvel do precipitado de células com moléculas ≥ 5 KDa ($\approx 70 \mu\text{g}$); **C-** Sobrenadante da cultura com moléculas ≤ 5 KDa ($\approx 1 \mu\text{g}$); **D-** Sobrenadante da cultura com moléculas ≥ 5 KDa ($\approx 1 \mu\text{g}$). Notar a significativa inibição de crescimento do revelador pela fração do sobrenadante induzido com moléculas ≥ 5 KDa (seta). Notar ainda, pequena inibição da fração do sobrenadante induzido e não induzido com moléculas ≤ 5 KDa. Os pontos verdes indicam o local de aplicação da fração.

O SDS-PAGE mostra uma expressão diferencial em nível da fração insolúvel do conteúdo celular da cultura do clone B3 induzida (Figura 5.9). Como o antimicrobiano é secretado e induzido por IPTG, podemos sugerir uma maquinaria molecular de secreção em nível de membranas, uma vez que a fração insolúvel apresenta expressão diferencial.

Destacando novamente a ORF com homologia a um peptídeo sinal (similar ao peptídeo sinal de *Acinetobacter*), uma relação mais direta é estabelecida entre esta ORF e o controle de secreção do antimicrobiano, pois existe uma situação de condição para atividade extracelular e indução por IPTG. Um operon inteiro pode ter sido clonado, com a seqüência codificadora do antimicrobiano e as seqüências codificadoras de proteínas de membrana envolvidas em sua secreção, ou seja, uma via secretória inteira, talvez com lipoproteínas transportadoras do tipo não-ABC, as quais estão covalentemente ligadas à superfície celular e atuam na internalização e/ou secreção de açúcares, aminoácidos, metais, minerais e peptídeos antiibióticos. E o peptídeo sinal atua direcionado a secreção dessas substâncias. Essa idéia é reforçada pelo fato de haver uma expressão diferencial em nível de fração insolúvel (membranas) das células induzidas por IPTG. A ORF com homologia com seqüências de peptídeo sinal também pode vir a ser o próprio antimicrobiano, uma vez que proteínas com atividade de lipase cujas seqüências codificadoras continham homologia com peptídeo sinal já foram encontradas (LEE et al., 2006; KHANNAVILLI et al., 2008).

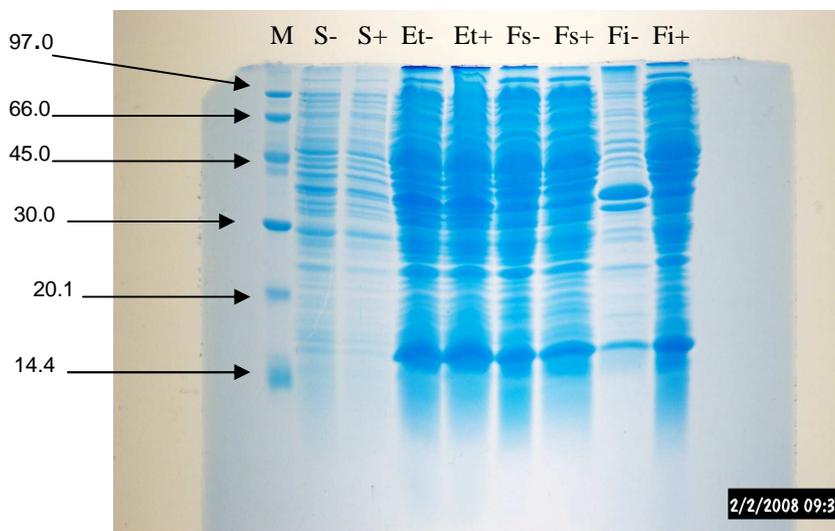


Figura 5.9- SDS-PAGE para avaliação de expressão diferencial de proteínas com indução por IPTG. M: marcador LMW (peso molecular em KDa); S-: sobrenadante não induzido; S+: sobrenadante induzido; Et-: extrato total de células não induzido; Et+: extrato total de células induzido; Fs-: fração solúvel de células não induzida; Fs+: fração solúvel de células não induzida; Fi-: fração insolúvel de células não induzida; Fi+: fração solúvel de células induzida. Notar a expressão diferencial na fração insolúvel de células induzida.

O fato de apenas 1 µg de proteínas totais do sobrenadante em suas frações maior ou igual e menor ou igual a 5 KDa ter sido testado no bioensaio contra *B. subtilis* e ter inibido significativamente o crescimento deste revelador indica que existe um forte mecanismo de secreção, que pode enviar todo o antimicrobiano sintetizado para fora da célula. Isso porque uma quantidade setenta vezes maior de proteínas totais da fração solúvel das células não foi capaz de causar inibição de crescimento do revelador (Figura 5.8), mostrando ausência ou concentração muito baixa deste antimicrobiano no meio intracelular. Desta forma, o antimicrobiano é secretado, sua expressão pode ser controlada por indução com IPTG, seu tamanho é igual ou superior a 5 KDa e sua atividade antibiótica é condicionada a Gram-positivos. Como tem o tamanho mencionado e sua atividade é condicionada a Gram-positivos, pode-se tratar de uma hidrolase de parede celular bacteriana, pois é uma característica desta classe de antimicrobianos não ser efetivo contra a maioria dos Gram-negativos (PARISIEN et al., 2007). Essa provável proteína está sendo enviada para o laboratório de bacteriologia da UNICAMP onde será testada para sua atividade contra MRSA. As ORFs também estão sendo clonadas e as proteínas recombinantes serão expressas em vetores pET para análise posterior.

Referências

- AMES, GF-L. Resolution of bacterial proteins by polyacrylamide gel electrophoresis on slabs. Membrane, soluble and periplasmatic fractions. **Journal of Biological Chemistry**. Vol. 249, p. 634-644, 1974.
- BERTRAND, H.; POLY, F.; VAN, V.T.; LOMBARD, N.; VOGEL, T.M. and SIMONET, P. High molecular weight DNA recovery from soils prerequisite for biotechnological metagenomic library construction. **Journal of Microbiological Methods**. Vol. 62, p.1– 11, 2005.

BLOBEL, G. & SABATINI, D. D. Ribosome– membrane interaction in eukaryotic cells. **Biomembranes**. Vol. 2, p. 193–195, 1971

DANIEL, R. The soil metagenome – a rich resource for the discovery of novel natural products. **Current Opinion in Biotechnology**. Vol. 15, p. 199–204, 2004.

HONG, H. J.; PAGET, M. S.; BUTTNER, M. Characterization of an inducible vancomycin resistance system in *Streptomyces coelicolor* reveals a novel gene (*vanK*) required for drug resistance. **Journal of Molecular Microbiology**. Vol. 44, p. 1199, 2002.

KANDAVILLI, S.; HOMER, K.A.; YUSTE, J.; BASAFANNA, S.; MITCHELL, T and BROWN, J. Maturation of *Streptococcus pneumoniae* lipoproteins by a type II signal peptidase is required for ABC transporter function and full virulence. **Molecular Microbiology**. Vol.67, p. 541–557, 2008.

KIESER, T.; BIBB, M. J.; BUTTNER, M. J.; CHATER, K. F. and HOPWOOD, D. A, (orgs) *Practical Streptomyces Genetics* (John Innes Foundation, Norwich, UK, ed. 1, 2000).

KIM, A.C.; OLIVER, D.C. and PAETZEL, M. Crystal Structure of a Bacterial Signal Peptide Peptidase. **Journal of Molecular Microbiology**. Vol. 376, p. 352–366, 2008.

GUARDABASSI, L.; PETERSON, A.; OLSEN, J.E.; and DALSGAARD, A. Antibiotic Resistance in *Acinetobacter* spp. Isolated from Sewers Receiving Waste Effluent from a Hospital and a Pharmaceutical Plant. A. **Applied Environmental Microbiology**. Vol. 70, p. 984, 2004.

LAMMLE, K.; ZIPPER, H.; BREUER, M.; HAUER, B.; BUTA, C.; BRUNNER, H. and RUPP, S. Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. **Journal of Biotechnology**. Vol.127, p. 575–592, 2007.

LEE, M-H; LEE, C-H; OH, T-K; SONG, J.K and YONG, J-H. Isolation and Characterization of a Novel Lipase from a Metagenomic Library of Tidal Flat Sediments:

Evidence for a New Family of Bacterial Lipases. **Applied Environmental Microbiology**. Vol. 72, p. 7406–7409, 2006.

MADIGAN M; MARTINKO J. (2005). *Brock Biology of Microorganisms*, 11th ed., Prentice Hall.

MOSS, G.P. Recommendations of the Nomenclature Committee. **International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes by the Reactions they Catalyse**. 14.03.2006.

NOVAK, P. & DEV, I. K. . Degradation of a signal peptide by protease IV and oligopeptidase **American Journal of Bacteriology**. Vol.170, p. 5067–5075, 1988.

PARISIEN, A.; ALLAIN, B.; ZHANG, J.; MANDEVILLI, R. and LAN, C.Q. Novel alternatives to antibiotics: bacteriophages, bacterial cell wall hydrolases, and antimicrobial peptides. **Journal of Applied Microbiology**. Vol. 104, p.1364-5072, 2007.

PENG, C.; LUI, L.; SHAO, X.; CHI, C. and WANG, C. Identification of a novel class of conotoxins defined as V-conotoxins with a unique cysteine pattern and signal peptide sequence. **Peptides**. doi:10.1016/j.peptides.2008.01.007 (artigo aceito)

SUTCLIFFE, I.C., and HARRINGTON, D.J. Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes. **Microbiology**. Vol. 148, p. 2065–2077, 2002.

WALSH, C. Where will new antibiotics come from? **Nature Reviews/ Microbiology**. Vol. 1, p. 65-70, 2003.

WOROBO, R.W.; BELKUM, MJV.; SAILER, M.; ROY, K.; VEREDAS, J. and STILES, M. A Signal Peptide Secretion-Dependent Bacteriocin from *Carnobacterium divergens*. **Journal of Bacteriology**. Vol. 177, p. 3143–3149, 1995.

WRIGHT, G.D. The antibiotic resistome: the nexus of chemical and genetic diversity. *Nature Reviews/Microbiology*. Vol. 5, p. 175-186, 2007.

6. Conclusões gerais

- A metodologia convencional desenvolvida (Capítulo I) permite um alto rendimento na extração de DNA de solo. Além disso, a qualidade do DNA extraído também é adequada aos processos de clonagem.
- A extração direta de DNA de solo mostrada no Capítulo II permite uma exposição homogênea das células à pressão de lise, sendo indicada como importante alternativa na busca da redução do viés genético e funcional em bibliotecas metagenômicas.
- O número de primers distintos utilizados na amplificação e cobertura do metagenoma é proporcional a diversidade taxonômica e metabólica encontrada na biblioteca.
- O tamanho das seqüências clonadas neste trabalho já exhibe potencial biotecnológico, podendo ser evoluído na medida em que esta técnica seja aprimorada.
- A técnica e o conhecimento gerados neste trabalho devem ser considerados em metagenômica a fim de se evitar subestimativas de caráter genético e funcional em trabalhos futuros.
- Foi identificado um clone metagenômico proveniente de uma biblioteca de solo de Mata Atlântica com potencial antimicrobiano.

7. REFERÊNCIAS BIBLIOGRÁFICAS

ALTSCHUL, S.F; GISH, W.; MILLER, W.; MYERS, E.W and LIPMAN DJ. Basic local alignment search tool. **Journal of Molecular Biology**. Vol. 215, p. 403–10, 1990.

AMES, GF-L. Resolution of bacterial proteins by polyacrylamide gel electrophoresis on slabs. Membrane, soluble and periplasmatic fractions. **Journal of Biological Chemistry**. Vol. 249, p. 634-644, 1974.

ARAÚJO, E.S.; dos SANTOS, A.M.; AREIAS, R.G.B.M.; de SOUZA, S.R.and FERNANDES, M.S. USO DE RAPD PARA ANÁLISE DE DIVERSIDADE GENÉTICA EM ARROZ. **Agronomia**. Vol.37, p. 33 - 37, 2003.

BERTRAND, H.; POLY, F.; VAN, V.T.; LOMBARD, N.; VOGEL, T.M. and SIMONET, P. High molecular weight DNA recovery from soils prerequisite for biotechnological metagenomic library construction. **Journal of Microbiological Methods** . Vol.62, p. 1–11, 2005.

BIONDI, G.E; MENGONI, A. and BAZZICALUPO, M. (2006) “Analysis of Microbial Population Genetics” in VARMA, A.; ABBOTT, L.; WERNER, D. and HAMPP, R. (orgs.), **Plant Surface Microbiology**. Berlin: Springer Berlin Heidelberg, 551-565.

BLOBEL, G. & SABATINI, D. D. Ribosome– membrane interaction in eukaryotic cells. **Biomembranes**. Vol. 2, p. 193–195, 1971.

CLARK, J.M. Novel non-templated addition reactions catalyzed by prokaryotic DNA polymerases, **Nucleic Acids Research**. Vol. 16, p. 9677-9686, 1998.

CURTIS, T.P.; SLOAN, W.T.; SCANNELL, J.W. Estimating prokaryotic diversity and its limits, **Proceedings of National Academy of Science. USA**. Vol.132, p. 10494–10499, 2002.

DANIEL, R. The soil metagenome – a rich resource for the discovery of novel natural products. **Current Opinion in Biotechnology**. Vol. 15, p. 199–204, 2004.

DUBEY, S.K.; TRIPATHI, A.K. and UPADHYAY, S.N . Exploration of soil bacterial communities for their potential as bioresource. **Bioresource Technology**. Vol. 55, p. 2217–2224, 2006.

FERRER, M.; GOLYSHINA, O.; BELOQUI, A. and GOLYSHIN, P.N. Mining enzymes from extreme environments. **Current Opinion in Microbiology**. Vol. 10, p. 207–214, 2007.

GARCIA-FRAILE, P.; RIVAS, R.; WILLEMS, A.; PEIX, A.; MARTENS, M.; MARTINEZ-MOLINA, E.; MATEOS, P.F. and VELAZQUEZ, E. *Rhizobium cellulosilyticum* sp. nov., isolated from sawdust of *Populus alba*. **International Journal of Systematic and Evolutionary Microbiology**. Vol.57, p. 844-848, 2005.

GERISHER, U. (2008). **Acinetobacter Molecular Biology**. 1ª edição. Caister Academic Press

GOMES, N.C.M.; KOSHELEVA, I.A.; ABRAHAM, W.R, SMALLA, K. Effects of the inoculant strain *Pseudomonas putida* KT2442 (pNF142) and of naphthalene contamination on the soil bacterial community. **FEMS Microbiology Ecology**. Vol. 1, p. 21–33, 2005.

GUARDABASSI, L.; PETERSON, A.; OLSEN, J.E.; and DALSGAARD, A. Antibiotic Resistance in *Acinetobacter* spp. Isolated from Sewers Receiving Waste Effluent from a Hospital and a Pharmaceutical Plant. A. **Applied and Environmental Microbiology**. Vol. 70, p. 984, 2004.

HARRY, M.; JUSSEAUME, N., GAMBIER, B. & GARNIER-SILLAM, E.. Use of RAPD markers for the study of microbial community similarity from termite mounds and tropical soils. **Soil Biology and Biochemistry**. Vol. 33, p. 417–427, 2001.

HONG, H. J.; PAGET, M. S.; BUTTNER, M. Characterization of an inducible vancomycin resistance system in *Streptomyces coelicolor* reveals a novel gene (*vanK*) required for drug resistance. **Journal of Molecular Microbiology**. Vol. 44, p. 1199 ,2002.

HOTON, FM.; ANDRUP, L.; SWIECICKA, I. and MAHILLON J. "The cereulide genetic determinants of emetic *Bacillus cereus* are plasmid-borne." **Microbiology**. Vol.7, p. 2121-4, 2005.

<http://www.fermentas.com/catalog/kits/kitinstaclone.htm>, acessado aos 11.01.2008.

HUBBARD, B. K. and WALSH, C. T. Exploiting resistance prokaryotic genes. **Angewandte Chemie International Edition in English**. Vol. 42, p. 12-21, 2003.

HUBER, T.; FAULKNER, G. and HUGENHOLTZ, P. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. **Bioinformatics**. Vol. 20, p. 2317–9, 2004.

HUSON, D.H.; AUCH, A.F.;STEPHAN, J.Q. e SCHUSTER, C. MEGAN analysis of metagenomic data. **Genome Research**. Vol. 17, p. 377-386, 2007.

JOHNSON, P.L.F. and SLATKIN, M. Inference of population genetic parameters in metagenomics: A clean look at messy data. **Genome Research**. Vol. 133, p. 1320-1327, 2006.

KANDAVILLI, S.; HOMER, K.A.; YUSTE, J.; BASAFANNA, S.; MITCHELL, T and BROWN, J. Maturation of *Streptococcus pneumoniae* lipoproteins by a type II signal peptidase is required for ABC transporter function and full virulence. **Molecular Microbiology**. Vol. 67, p. 541–557, 2008.

KANT, A.D and TRIPLETT, E.W. MICROBIAL COMMUNITIES AND THEIR INTERACTIONS IN SOIL AND RHIZOSPHERE ECOSYSTEMS. **Annual Review of Microbiology**. Vol. 56, p. 211–36, 2002.

KIESER, T.; BIBB, M. J.; BUTTNER, M. J.; CHATER, K. F.and HOPWOOD, D. A, **Practical Streptomyces Genetics** (John Innes Foundation, Norwich, UK, ed. 1, 2000).

KIM, A.C.; OLIVER, D.C. and PAETZEL, M. Crystal Structure of a Bacterial Signal Peptide Peptidase. **Journal of Molecular Biology**. Vol. 376, p. 352–366, 2008.

KIM, Y.J.; CHOI, G.; KIM, S.; YOON, G.; KIM, Y.S.; RYU, Y. Screening and characterization of a novel esterase from a metagenomic library. **Protein Expression and Purification**. Vol. 15, p. 315–323, 2006.

KOTIRANTA, A.; LOUNATMAA, K.and HAAPASALO, M.. "Epidemiology and pathogenesis of Bacillus cereus infections". **Microbes Infections**. Vol. 2, p. 189-98, 2000.

LAMMLE, K.; ZIPPER, H.; BREUER, M.; HAUER, B.; BUTA, C.; BRUNNER, H. and RUPP, S. Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. **Journal of Biotechnology**. Vol. 127, p. 575–592, 2007.

LEE, M-H; LEE, C-H; OH, T-K; SONG, J.K and YONG, J-H. Isolation and Characterization of a Novel Lipase from a Metagenomic Library of Tidal Flat Sediments: Evidence for a New Family of Bacterial Lipases. **Applied and Environmental Microbiology**. Vol. 72, p. 7406–7409, 2006.

MACIEL, B.M ; DIAS, J.T ; SANTOS, A. C. F. ; ARGOLO FILHO, R. C. ; FONTANA, R. ; LOGUERCIO, L. L. ; REZENDE, R.P. Microbial surfactant activities from a petrochemical landfarm in a humid tropical region of Brazil. **Canadian Journal of Microbiology**. Vol. 53, p. 937-943, 2007.

MADIGAN M; MARTINKO J. (2005). **Brock Biology of Microorganisms**, 11th ed., Prentice Hall.

MAIDAK, B.L.; OLSEN, G.J.; LARSEN, N.; OVERBEEK, R.; McCAUGHEY, M.J. and WOESE, C.R. The Ribosomal Database Project (RDP). **Nucleic Acids Research**. Vol. 24, p. 82–5, 1996

McHARDY, A.C and RGOUTSOS, I. What's in the mix: phylogenetic classification of metagenome sequence samples. **Current Opinion in Microbiology**. Vol. 10, p.499–503, 2007.

MENCONI, G. and MARANGONI, J. A Compression-Based Approach for Coding Sequences Identification. I. Application to Prokaryotic Genomes. **Journal of Computational Biology**. Vol. 8, p. 1477-1488, 2006.

MOSS, G.P. Recommendations of the Nomenclature Committee. **International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes by the Reactions they Catalyse**. 14.03.2006.

NOVAK, P. & DEV, I. K. (). Degradation of a signal peptide by protease IV and oligopeptidase **American Journal of Bacteriology**. Vol. 170, p. 5067–5075, 1988.

OGRAM, A.; SAYLER, G.S., BARKAY, T. The extraction and purification of microbial DNA from sediments, **Journal of Microbiological Methods**. Vol.7, p. 57–66, 1987.

OH, H.J.; CHO, K-W.; JUNG, I-S.; KIM, W-O.; HUR, B-K. and KIM, G-J. Expanding functional spaces of enzymes by utilizing whole genome treasure for library construction. **Journal of Molecular Catalysis B: Enzymatic**. Vol. 26, p. 241–250, 2003.

OJO, O.A. Molecular strategies of microbial adaptation to xenobiotics in natural environment. **Biotechnology and Molecular Biology Review**. Vol. 1, p. 1-13, 2007.

PACE, N.R.; STAHL, D.A.; LANE, D.J and OLSEN G.J: Analyzing natural microbial populations by rRNA sequences. **ASM News**. Vol. 46, p. 51:4-12, 1985.

PAPPAS Jr, G.J; BENABDELLAH, K.; ZINGALES, B.and GONZALEZ, A. Expressed sequence tags from the plant trypanosomatid *Phytomonas serpens*. **Molecular Biochemical Parasitology**. Vol. 142, p. 149–57, 2005.

PARISIEN, A.; ALLAIN, B.; ZHANG, J.; MANDEVILLI, R. and LAN, C.Q. Novel alternatives to antibiotics: bacteriophages, bacterial cell wall hydrolases, and antimicrobial peptides. **Journal of Applied Microbiology**. Vol. 104, p.1364-5072, 2007.

PARK, H.J.; JEON, J.H.; KANG, S.G.; LEE, J-H.; LEE, S-A.and KIM, H-K. Functional expression and refolding of new alkaline esterase, EM2L8 from deep-sea sediment metagenome. **Protein Expression and Purification**. Vol. 155, p. 340–347, 2007.

PENG, C.; LUI, L.; SHAO, X.; CHI, C. and WANG, C. Identification of a novel class of conotoxins defined as V-conotoxins with a unique cysteine pattern and signal peptide sequence. **Peptides**. doi:10.1016/j.peptides.2008.01.007 (artigo aceito)

RAHAL, J. Novel antibiotic combinations against infections with almost completely resistant *Pseudomonas aeruginosa* and *Acinetobacter* species. 2006. **Clinical and Infectious Diseases**. 43 Suppl 2: S95-9.

RIESENFELD, C.S; SCHLOSS, P.and HANDELSMAN, J. METAGENOMICS: Genomic Analysis of Microbial Communities. **Annual Review of Genetics**. Vol. 38, p. 525–52, 2004.

ROBE, P; NALIN, R.; CAPELLANO, C.; VOGEL, T.M.; SIMONET, P. Extraction of DNA from soil. **European Journal of Soil Biology**. Vol. 39, p.183-190, 2003.

RONDON, M.R.; AUGUST, P.R.; BETTERMANN, A.D.; BRADY, S.F.; GROSSMAN, T.H.; LILES, M.R.; LOICONO, K.A.; LYNCH, B.A.; MACNEIL, I.A.; MINOR, C.; TIONG, C.L.; GILMAN, M.; OSBURNE, M.S.; CLARDY, J.; HANDELSMAN, J.; GOODMAN, R.M. Cloning the Soil Metagenome: a Strategy for Accessing the Genetic and Functional Diversity of Uncultured Microorganisms. **Applied and Environmental Microbiology**. Vol. 86, p. 2541–2547, 2000.

SANTOS, A.C.F. Diversidade bacteriana do solo de landfarm da refinaria de petróleo Landulpho Alves- Ilhéus, Bahia. 2007. 94 p. (Dissertação de Mestrado). **Universidade Estadual de Santa Cruz**, Ilhéus, BA. 2007.

SCHLOSS, P. D., and HANDELMAN, J. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. **Applied and Environmental Microbiology**. Vol. 71, p. 1501–1506, 2005.

SHIZUYA, H., Birren, B., KIM, U.-J., MANCINO, V., SLEPAK, T., TACHIRI, Y., and SIMON, M... Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. **Proceedings of the National Academy of Science. USA**. Vol. 89, p. 8794–8797, 1992.

SILVA-MACENA. Análise da biodiversidade metagenômica bacteriana de um solo de Mata Atlântica e isolamento de DNA de alto peso molecular visando aplicações biotecnológicas. - Ilhéus, Bahia. 2007. (Dissertação de Mestrado). **Universidade Estadual de Santa Cruz**, Ilhéus, BA. 2007.

STREIT, Wolfgang R. and SCHMITZ, Ruth A. 2004. Metagenomics – the key to the uncultured microbes. **Current Opinion in Microbiology**. Vol. 7, p. 492–498, 2004.

SUTCLIFFE, I.C., and HARRINGTON, D.J. Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes. **Microbiology**. Vol. 148, p. 2065–2077, 2002.

TRINGE, S.G.; von MERING, C.; KOBAYASHI, A.; SALAMOV, A.A.; CHEN, K.; CHANG, H.W.; PODAR, M.; SHORT, J.M.; MARTHUR, E.J.; DETTER, J.C.; BOORK, P.; HUGENHOLTZ, P. and RUBIN, E.M. Comparative Metagenomics of Microbial Communities. **Science**. Vol. 77, p. 554-557, 2005.

VROH, B.I., HARVENGT, L., CHANDELIER, A., MERGEAI, G.; and JARDIN, P.D.,. Improved RAPD amplification of recalcitrant plant DNA by the use of activated charcoal during DNA extraction. **Plant Breeding**. Vol. 115, p. 205–206, 1996.

WALSH, C. Where will new antibiotics come from? **Nature Reviews/ Microbiology**. Vol. 1, p. 65-70, 2003.

WHITEHOUSE, C.A and HOTTEL, H.E. Comparison of five commercial DNA extraction kits for the recovery of *Francisella tularensis* DNA from spiked soil samples. **Molecular and Celular Probes**. Vol. 21, p. 92–96, 2007.

WILSON, I.G. Inhibition and Facilitation of Nucleic Acid Amplification. **Applied and Environmental Microbiology**. Vol. 23, p. 3741–3751, 1997.

WRIGHT, G.D. The antibiotic resistome: the nexus of chemical and genetic diversity. **Nature Reviews/Microbiology**. Vol. 5, p. 175-186, 2007.

YUN, J.; KANG, S.; PARK, S.; YOON, H.; KIM, M.; HEU, S. and RYU, S. Characterization of a Novel Amylolytic Enzyme Encoded by a Gene from a Soil-Derived Metagenomic Library. **Applied and Enviromental Microbiology**. Vol. 15, p. 7229–7235, 2004.

YUTIN, N.; SUZUKI, M.T. and BÉJÁ, O. Novel Primers Reveal Wider Diversity among Marine Aerobic Anoxygenic Phototrophs. **Applied and Environmental Microbiology**. Vol. 12, p. 8958–8962, 2005.

8. Apêndice

Tabela com as sequências dos primers utilizados nos experimentos referentes ao capítulo II.

<i>Primer</i>	Sequência	<i>Primer</i>	Sequência
OPB03	CATCCCCCTG	OPJ16	CTGCTTAGGG
OPB06	TGCTCTGCCC	OPK08	GAACACTGGG
OPB07	GGTGACGCAG	OPK12	TGGCCCTCAC
OPB13	TTCCCCGCT	OPK13	GGTTGTACCC
OPB14	TCCGCTCTGG	OPK19	CACAGGCGGA
OPC01	TTCGAGCCAG	OPK20	GTGTCGCGAG
OPC03	GGGGGTCTTT	OPM03	GGGGGATGAG
OPC15	GACGGATCAG	OPM04	GGCGGTTGTC
OPC17	TTCCCCCAG	OPM05	GGGAACGTGT
OPC19	GTTGCCAGCC	OPM06	CTGGGCAACT
OPD07	TTGGCACGGG	OPM13	GGTGGTCAAG
OPD11	AGCGCCATTG	OPN09	TGCCGGCTTG
OPD13	GGGGTGACGA	OPN14	TCGTGCGGGT
OPD16	AGGGCGTAAG	OPN17	CATTGGGGAG
OPD17	TTTCCACGG	OPN18	GGTGAGGTCA
OPD19	CTCCGGACTT	OPN19	GTCCGTACTG
OPI04	CCGCCTAGTC	OPQ06	GAGCGCCTTG
OPI06	AAGGCGGCAG	OPQ08	CTCCAGCGGA
OPI07	CAGCGACAAG	OPQ10	TGTGCCCGAA
OPI09	TGGAGAGCAG	OPQ11	TCTCCGCAAC
OPI15	TCATCCGAGG	OPQ14	GGACGCTTCA
OPI17	GGTGGTGATG	OPQ15	GGGTAACGTG
OPI18	TGCCCAGCCT	OPR08	CCCGTTGCCT
OPI19	AATGCGGGAG	OPS07	TCCGATGCTG
OPJ04	CCGAACACGG	OPS10	ACCGTTCCAG
OPJ06	TCGTTCCGCA	OPT17	CCAACGTCGT
OPJ14	CACCCGGATG	OPT20	GACCAATGCC