



Universidade de Brasília

Instituto de Biologia

Departamento de Biologia Celular

Pós-graduação em Biologia Molecular

RIQUEZA DO DOMÍNIO *ARCHAEA* NO SOLO DO BIOMA CERRADO

ELISA CATÃO CALDEIRA PIRES

Dissertação apresentada ao Programa de Pós-Graduação de Biologia Molecular do Departamento de Biologia Celular, Instituto de Biologia, Universidade de Brasília, para obtenção do título de Mestre em Biologia Molecular.

Orientador: Cynthia M. Kyaw

2012



Universidade de Brasília

Instituto de Biologia

Departamento de Biologia Celular

Pós-graduação em Biologia Molecular

RIQUEZA DO DOMÍNIO *ARCHAEA* NO SOLO DO BIOMA CERRADO

ELISA CATÃO CALDEIRA PIRES

Dissertação aprovada pela banca:

Dra. Cynthia M. Kyaw (orientadora) (Universidade de Brasília)

Dra. Cristine Chaves Barreto (Universidade Católica – Brasília)

Dr. Robert Miller (Universidade de Brasília)

Brasília, 2012

Tenho muito orgulho dos meus pais. Não poderia ter melhores exemplos a seguir. A estes, minhas linhagens ancestrais, eu dedico esta dissertação.

Apesar de ser um ramo filogeneticamente mais recente, meu irmão, Francisco, é componente imprescindível da minha árvore e a quem dedico também este trabalho.

AGRADECIMENTOS

Em primeiro lugar quero agradecer ao Woese e aos seus colaboradores pelo seu brilhante trabalho de “descoberta” e proposta de criação de um novo Domínio de vida: *Archaea*.

Agradeço à minha orientadora, Cynthia Kyaw, por sua paciência, humildade, carinho e apoio constantes, cuja orientação sempre se faz presente. Sem dúvida admiro seu conhecimento notável sobre microbiologia e biologia molecular. Ainda quero aprender a identificar bactérias em placa só de olhar e medir OD pelo “olhômetro”.

Agradeço aos professores colaboradores da pesquisa Ricardo Krüger, Mercedes Bustamante e Cristine Barreto pela ajuda na condução dos experimentos.

Posso dizer com segurança que meu trabalho é dos mais prazerosos que existe. Meu dia-a-dia é divertido por causa dos meus companheiros de laboratório: Julianna Peixoto, Helena Magaldi, Cecília Kosmann e Thiago Rodrigues. Por isso a eles agradeço a companhia, a ajuda, as piadas e todas as contribuições intelectuais.

E agradeço ao mestrado por ter me apresentado uma das minhas melhores amigas: a Julianna Peixoto, mulher super inteligente que eu admiro, e que sempre me anima.

Queria agradecer às minhas amigas e professoras particulares, Alinne Castro e Regina Sartori, pela didática, paciência e carinho.

Devo agradecer ao amigo e colega André Bertran por me auxiliar na preparação de células competentes e ao professor Renato Rezende por ceder seu laboratório todas as vezes que necessitei fazer eletroporação.

Agradeço à professora Silviene por ter em seu laboratório um aparelho Nanovue™ (GE) disponibilizado para o uso comum e às suas alunas Tatiana dos Anjos e Mariana Marzullo por me auxiliarem com o uso do aparelho.

Obrigada ao aluno de doutorado Marcus Teixeira e à técnica Adriane Oliveira por realizarem o sequenciamento das minhas placas no Laboratório de Biologia Molecular.

O maior obrigado devo aos meus pais. Busco neles toda a inspiração, perseverança e prazer no trabalho. Impossível observar minha mãe querida, mulher batalhadora, e não querer ser uma grande profissional. E também não consigo observar meu pai, professor da UnB, com grandes projetos de pesquisa, sem querer seguir o mesmo caminho.

Neste sentido agradeço também à minha dinda, tia Helena Catão, antropóloga e pesquisadora também, que me aconselhou ao longo do mestrado.

Um obrigado especial ao meu melhor amigo: meu irmão. Incrivelmente inteligente e maduro que me ouve mais do que qualquer um quando preciso.

Pode parecer ridículo, mas como pesquisa é trabalhar em casa também, quero agradecer à minha maior companheira, minha gata Marie, por me fazer sempre sorrir.

Obrigado meninas! Minhas melhores amigas: Renata, Carol e Rapha. À Carol, bióloga, com quem sempre aprendo um pouco mais sobre evolução. À Rê e à Rapha por mesmo sem entender quase nada do que eu falo sempre me escutam mesmo que em meio a risadas.

Agradeço ao Felipe, que me apoiou no início do meu mestrado e que me lembrou nas vezes que precisei porque gosto de fazer pesquisa. Citando: “*Você gosta muito de trabalhar com seus bichos microscópicos (ou com o DNA desses bichos, que é pior ainda!)*”.

Agradeço à perua-amiga Carmen e ao meu grande amigo Pablo Antunes por confiarem no meu perfil de pesquisadora. Meu obrigado vai também para meu amigo e conselheiro, Flávio Eiró, que me ajuda sempre nas minhas crises.

"If at first the idea is not absurd then there is no hope for it"

Einstein

"—a new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it."

Max Planck

LISTA DE FIGURAS E TABELAS

Figura 1. Árvore filogenética universal, baseada em sequências de rRNA 16S e 18S, revelando os Domínios <i>Archaea</i> , <i>Bacteria</i> e <i>Eucarya</i> (Woese <i>et al.</i> , 1990).....	10
Figura 2. Árvore filogenética de <i>Archaea</i>	15
Figura 3. Árvore filogenética de <i>Archaea</i> baseada em proteínas ribossomais.....	16
Figura 4. Esquema dos tipos de vegetação do bioma Cerrado.	20
Figura 5. Esquema da amostragem de um transecto. Cedido por Alinne Castro.....	23
Figura 6. Esquema do gene de 16S rRNA, suas regiões variáveis e as posições de anelamento dos iniciadores.....	24
Figura 7. Precipitação mensal (em mm) registrada na Reserva Ecológica do IBGE Brasília-DF, em 2010. (http://www.recor.org.br/Estacao/consulta.asp)	32
Figura 8. Perfil eletroforético, em gel de agarose 1%, corado com brometo de etídio, do DNA metagenômico extraído das amostras de solo das fitofisionomias cerrado denso (CDa, CDb) e mata de galeria (MGa, MGb), em duplicata.....	34
Figura 9. Perfil eletroforético, em gel de agarose 1% corado com brometo de etídio, dos produtos amplificados no experimento de PCR empregando os iniciadores 21F e 958R.....	36
Figura 10. Perfil eletroforético, em gel de agarose 1% corado com brometo de etídio, do DNA plasmidial de clones aleatoriamente selecionados a partir das bibliotecas de rDNA das amostras de solo.	38
Figura 11. Histograma da classificação taxonômica das sequências de gene de 16S rRNA de cerrado denso (CDa, CDb) e mata de galeria (MGa, MGb)..	40
Figura 12. Diagrama de Venn representando o número de OTUs (3%) exclusivas em cada réplica e compartilhadas entre estas, para CD e MG.....	46
Figura 13. Diagrama de Venn representando o número de OTUs (3%) exclusivas a cada fitofisionomia e compartilhadas entre estas.....	48
Figura 14. Curvas de rarefação com as porcentagens de dissimilaridade de 3%, 5%, 10%, 20%. Curva de CD e Curva de MG.	49

Figura 15. Curva de rarefação de todas as áreas com o limite de 3% de diferença.	50
Figura 16. Gráfico de Análise das Principais Coordenadas (PCA) construído no programa Unifrac (Lozupone <i>et al.</i> , 2006) para demonstrar agrupamento em quadrantes dos ambientes amostrados.	55
Figura 17. Árvores filogenéticas construídas a partir dos representantes de cada OTU (3%) das amostras CD e MG e CD+MG, após alinhamento com sequências de diferentes isolados de <i>Archaea</i>	62
Tabela 1. Características gerais de <i>Bacteria</i> , <i>Archaea</i> e <i>Eukarya</i>	11
Tabela 2. Iniciadores utilizados	24
Tabela 3. Reagentes e respectivas concentrações finais na reação de PCR	24
Tabela 4. Propriedades físico-químicas das amostras de solo	33
Tabela 5. Concentração de DNA extraído das amostras de solo.....	34
Tabela 6. Número de sequências, OTUs (3%) e singletons por comunidade.	45
Tabela 7 Índices de riqueza (Ace e Chao) e diversidade (simpson e shannon) para as duas fitofisionomias.....	52
Tabela 8. Teste de hipótese Libshuff entre as amostras de cerrado denso e mata de galeria.	54

Resumo.....	1
Abstract.....	2
1. Introdução.....	3
2. Objetivos.....	5
2.1. Objetivo Geral.....	5
2.2. Objetivos Específicos	5
3. Revisão Bibliográfica	6
3.1. Uso de ácidos nucleicos na classificação de organismos	6
3.2. O Domínio Archaea	9
3.3. Microrganismos no solo	18
3.4. Solo de Cerrado	19
4. Materiais e Métodos.....	22
4.1. Coleta do solo.....	22
4.2. Extração de DNA genômico total.....	23
4.3. Iniciadores utilizados e condições de PCR	23
4.4. Purificação e ligação das sequências de rDNA em vetor pGEM-T Easy.....	25
4.5. Preparação de células competentes de <i>Escherichia coli</i> DH5 e transformação	26
4.5.1. Por choque térmico.....	26
4.5.2. Por eletroporação.....	27
4.6. Seleção de clones recombinantes.....	28
4.7. Extração de DNA plasmidial em pequena escala.....	28
4.7.1. <i>PureLinkTM Quick Plasmid Miniprep</i> (Invitrogen).....	28
4.7.2. Protocolo de lise alcalina	28
4.8. Quantificação da concentração do DNA	29

4.9.	Sequenciamento de DNA.....	29
4.10.	Análises de Bioinformática.....	30
5.	Resultados e Discussão.....	32
5.1.	Coleta de solos.....	32
5.2.	Extração de DNA total.....	34
5.3.	Amplificação de sequências de DNA relativas ao gene de 16S rRNA.....	35
5.4.	Transformação de <i>E. coli</i> DH5 com fragmentos do gene de 16S rRNA e seleção de clones recombinantes.....	37
5.5.	Extração de DNA plasmidial em pequena escala.....	37
5.6.	Análises de bioinformática.....	38
5.6.1.	Análise de qualidade das sequências.....	38
5.6.2.	Classificação taxonômica pelo <i>Ribosomal Database Project</i>	39
5.6.3.	Estimativa de riqueza e alfa-diversidade.....	42
5.6.4.	Análises estatísticas de beta-diversidade.....	53
5.6.5.	Árvores filogenéticas.....	56
6.	Considerações finais.....	65
7.	Conclusões e perspectivas.....	67
8.	Referências Bibliográficas.....	68
	ANEXO I. Trabalhos apresentados em congresso.....	82

Resumo

O Domínio *Archaea* vem sendo descrito principalmente através de técnicas moleculares, independentes de cultivo e, atualmente, é dividido em dois filos formalmente aceitos, Crenarchaeota e Euryarchaeota, além de quatro outros propostos – Korarchaeota, Nanoarchaeota, Thaumarchaeota e Aigarchaeota. Uma vez que sequências do gene de rRNA 16S de *Archaea* foram descritas a partir de amostras de diferentes ambientes, as archaeas passaram a ser consideradas organismos ubíquos na natureza. Este trabalho tem como objetivo caracterizar a riqueza de archaeas em amostras de solo de Cerrado - a principal vegetação do Brasil Central, sendo a segunda maior do país. O DNA metagenômico de amostras em duplicata de solos provenientes de duas fitofisionomias distintas, cerrado denso e mata de galeria, foi extraído empregando-se o *PowerSoil DNA Isolation Kit* (MOBio Laboratories Inc.). A construção de bibliotecas genômicas foi feita a partir de experimentos de PCR, com iniciadores *Archaea*-específicos, visando amplificar regiões do gene de rRNA16S a partir do DNA extraído das diferentes amostras. Os resultados obtidos a partir dos fragmentos sequenciados revelaram identidade com o domínio *Archaea* em 96% das sequências, todas pertencentes ao filo Crenarchaeota (possivelmente Thaumarchaeota), principalmente aos grupos I.1b e I.1c. Somente nas amostras de Mata de Galeria foram encontradas sequências do grupo I.1a. A riqueza de *Archaea* foi maior em cerrado denso do que em mata de galeria. Os cálculos não-paramétricos tais como as curvas de rarefação indicam que a amostragem do ambiente foi adequada. As amostras das duas fitofisionomias apresentam diferenças significativas pelas técnicas estatísticas de α -diversidade -Libshuff e pelo gráfico de Análise de Principais Coordenadas (PCA).

Abstract

The domain *Archaea* has been described mostly by culture-independent methods. Currently it is divided into two well accepted phyla, namely the Euryarchaeota and Crenarchaeota, and an additional four proposed phyla: Korarchaeota, Nanoarchaeota, Thaumarchaeota and Aigarchaeota. Since their 16S rDNA sequences have been described from diverse ecosystems not classified as extreme, *Archaea* is now considered ubiquitous and not a domain of only extreme organisms. This work describes the richness of the *Archaea* domain retrieved from soil samples of the Cerrado biome – the main vegetation of Central Brazil and second largest biome in the country. The metagenomic DNA from two replicas of each two distinct phytophysionomies, cerrado denso and mata de galeria, was extracted using the *PowerSoil DNA Isolation Kit* (MOBio Laboratories Inc.). The four DNA libraries were constructed through PCR amplification of the 16S rDNA using *Archaea*-specific primers. Our results reveal that 96% of the sequenced fragments have identity with the domain *Archaea*. All of these sequences are from the phylum Crenarchaeota (possibly Thaumarchaeota), predominantly affiliated to group I.1b and I.1c. Sequences affiliated to the group I.1a were only found on the soil from mata de galeria. The soils from cerrado denso have a greater richness of *Archaea* than those from mata de galeria. The non-parametric estimatives showed that both the environments have been well sampled. There is a significant difference between the soil samples of the two phytophysionomies shown by the statistical test of -Libshuff and by the Principal Coordinates Analysis graphic (PCA).

1. Introdução

O conhecimento sobre a riqueza e distribuição dos organismos em ambientes naturais é essencial para o estudo e entendimento dos mecanismos envolvidos nas relações de comunidades ecológicas, bem como para o desenvolvimento de medidas e políticas de conservação.

Curiosamente, embora os microrganismos compreendam a maior biodiversidade do nosso planeta, pouco é conhecido sobre sua distribuição e seus papéis ecológicos em ambientes naturais. Neste sentido, estudos envolvendo a ecologia microbiana vêm sendo realizados visando descrever as comunidades microbianas para responder às seguintes questões: Quem são? Onde estão? Quantos são? O que fazem? Como estudá-los?

O solo, foco de estudo de microbiologia ambiental, é um ecossistema que possui grande diversidade de microorganismos (Torsvik e Ovreas, 2002). A importância da comunidade bacteriana no solo é reconhecida, mas informações sobre a função das archaeas e sua ubiquidade nos ecossistemas terrestres são ainda recentes (Jurgens *et al.*, 1997; Kemnitz *et al.*, 2007).

O domínio *Archaea*, proposto em 1990 (Woese *et al.*, 1990), é composto por procariotos cujas células diferem em características moleculares, fisiológicas e estruturais, quando comparadas àquelas do domínio *Bacteria*. Na época de sua descoberta, as archaeas foram classificadas erroneamente como ancestrais das bactérias- reino Archaeobacteria. No entanto, foi observado posteriormente que tais microrganismos possuem características peculiares e exclusivas, além de possuírem algumas similaridades com os outros dois domínios. Aliás, foi verificado que *Archaea* compartilha um ancestral comum mais próximo a *Eucarya* que a *Bacteria*.

Inicialmente, *Archaea* era sinônimo de organismos extremófilos, por serem isoladas apenas de ambientes de pH, pressão, salinidade e temperaturas extremas. Entretanto, hoje em dia, estes organismos são encontrados em todo tipo de ambiente e, por isso, são considerados ubíquos.

Esta dissertação de mestrado tem como objetivo analisar a presença de membros do domínio *Archaea* em solos de duas fitofisionomias de Cerrado, o segundo maior bioma do Brasil, por meio da análise comparativa de sequências de DNA amplificadas por reações de polimerização em cadeia, empregando-se iniciadores dirigidos a genes de rRNA 16S. Assim, este trabalho visa trazer informações acerca da distribuição de um grupo de organismos muito pouco estudados no Cerrado.

2. Objetivos

2.1. *Objetivo Geral*

- Analisar a ocorrência de membros do Domínio *Archaea* em solos do bioma Cerrado

2.2. *Objetivos Específicos*

- Caracterizar a riqueza de archaeas no solo de duas regiões fitofisionômicas – Cerrado Denso e Mata de Galeria – da Reserva Biológica do Instituto Brasileiro de Geografia e Estatística (IBGE), reserva esta representante do bioma Cerrado
- Utilizar o gene do rRNA 16S como marcador filogenético para construção de bibliotecas genômicas.
- Correlacionar os grupos taxonômicos encontrados nas duas fitofisionomias

3. Revisão Bibliográfica

3.1. Uso de ácidos nucléicos na classificação de organismos

“We are fortunate to work in a time in which bacterial systematics has been elevated to a scientific multidisciplinary field” (Stackebrandt, 2006). A sistemática microbiana teve início com a descrição da primeira espécie bacteriana em 1872 por Ferdinand Cohn, acompanhado pelas primeiras técnicas de microscopia, coloração e obtenção de culturas puras, desenvolvidas por Robert Koch, Paul Ehrlich e Hans Christian Gram.

A classificação morfológica proposta por Cohn previa o agrupamento das bactérias em espécies distintas de acordo com suas características e já era considerada a existências de variedades dentro da mesma espécie, levando ao reconhecimento de espécies bacterianas como populações (Stackebrandt, 2006).

Em 1986 foi publicada a primeira edição do Manual de sistemática bacteriológica de Bergey em que os organismos foram agrupados segundo suas características morfológicas e fisiológicas. Hoje em dia, com a filogenia baseada em caracteres moleculares, as mesmas características morfológicas e fisiológicas são consideradas polifiléticas (Stackebrandt, 2006).

A sistemática é a classificação dos organismos vivos de acordo com normas de nomenclatura. A sistemática moderna baseia tal classificação em relações filogenéticas. A filogenia tem como objetivo relacionar espécies conhecidas com seus ancestrais mais próximos de forma a descrever sua história evolutiva.

Whittaker, em 1969, propôs a classificação dos seres vivos em cinco Reinos: Monera, Protista, Fungi, Plantae e Animalia. Tal sistema baseava-se em aspectos estruturais (células procarióticas ou eucarióticas), morfológicos e fisiológicos, especialmente os processos empregados no metabolismo energético (fotossíntese, absorção ou ingestão de nutrientes). Entretanto, tal sistema não evidenciava de forma clara as relações filogenéticas entre os seres vivos.

Zuckerlandl e Pauling (1965) postularam que o genoma de um organismo poderia ser considerado como o principal registro de sua história evolutiva. Neste trabalho, os autores utilizaram o sequenciamento de aminoácidos da molécula de hemoglobina de vários primatas, visando inferir as relações evolutivas entre estes. Este princípio foi de grande valia especialmente para a Microbiologia, uma vez que nesta época considerava-se impossível o estudo filogenético de microrganismos. A partir da proposta de Zuckerlandl e Pauling, o sequenciamento de ácidos nucléicos passou também a ser utilizado como ferramenta na sistemática de procariotos (Fox *et al.*, 1977).

Neste sentido, duas interessantes metodologias trouxeram grandes contribuições à sistemática microbiana moderna: a técnica de reassociação DNA:DNA, no início dos anos 70; e a catalogação de oligonucleotídeos de 16S rRNA, no final dos anos 70. Estudos de reassociação de DNA foram os primeiros a agrupar os organismos de acordo com relações filogenéticas. E ainda hoje, esta técnica é utilizada como padrão para definir uma “espécie”, usando o valor de 70% como limiar para diferenciar organismos da mesma espécie (Stackebrandt, 2006).

O RNA presente na subunidade menor do ribossomo, também chamado de SSU rRNA (do inglês *Small Subunit rRNA*), tem sido a molécula modelo para estudos de evolução e de comunidades microbianas, pois apresenta características que a torna um excelente indicador da história evolutiva dos organismos. O 16S rRNA apresenta função estrutural no ribossomo, tanto de definição das posições das proteínas ribossomais, como no contato entre as subunidades menor e maior do ribossomo, ao interagir com o 23S rRNA. Além destas funções, a região 3' do 16S rRNA possui a sequência anti-Shine-Dalgarno que se liga a montante ao códon de iniciação AUG do mRNA e também estabiliza o correto pareamento códon-anticódon entre mRNA e tRNA.

Genes codificadores de rRNA 16S (ou 16S rDNA, gene *rrn*) encontram-se presentes em todos os organismos e possuem homologia funcional e evolutiva, sendo possível inclusive comparar sequências de tais genes de 16S rRNA de procariotos com aqueles de rRNA 18S, de eucariotos. Além disso, este gene possui regiões

conservadas e variáveis, sendo que as porções conservadas e os elementos da estrutura secundária permitem o alinhamento das sequências variáveis. Outra característica que torna o gene de 16S rRNA interessante para análises evolutivas consiste na baixa frequência de eventos de transferência horizontal destes genes, permitindo assim inferências sobre as possíveis relações evolutivas entre organismos (Fox *et al.*, 1977; Olsen *et al.*, 1986).

Os primeiros estudos desta natureza foram realizados com o rRNA 5S, por ser uma molécula menor e de sequenciamento mais simples. Entretanto, o conteúdo informacional de uma molécula tão diminuta (120 nucleotídeos) não permitia análises de ecossistemas complexos. Neste sentido, o rRNA 16S (com cerca de 1500 nucleotídeos) mostrava-se mais adequado a estudos de taxonomia microbiana, por possuir informação suficiente para comparações com significância estatística (Olsen *et al.*, 1986). Tal abordagem foi utilizada por Woese e Fox (1977), em seu artigo basilar, onde foi proposta a existência de dois reinos (ainda assim chamados na época) distintos entre os procariotos, denominados Eubacteria e Archaeobacteria.

O advento do sequenciamento de 16S rRNA permitiu determinar quão homogêneo é um grupo filogenético, bem como posicionar no sistema hierárquico uma linhagem ainda não classificada (Stackebrandt e Goebel, 1994).

Devido à alta conservação da estrutura primária do 16S rRNA, organismos cuja porcentagem de similaridade da reassociação de DNA:DNA é de 70%, logo são classificados na mesma espécie, possuirão no mínimo 97% de identidade entre suas sequências do gene de 16S rRNA (Stackebrandt e Goebel, 1994).

O gene do RNA ribossomal 16S não é o único marcador filogenético que pode ser utilizado para estudos de diversidade microbiana. Os polimorfismos encontrados no gene de 16S rRNA por vezes não são suficientes para a análise robusta para diferenciar espécies. Outros genes “housekeeping” podem ser utilizados para o estudo da história evolutiva de linhagens específicas (Torsvik e Ovreas, 2006). São eles codificadores de proteínas ribossomais, proteínas de ligação a DNA e todas as aminoacil sintetases. Reconhecidamente para o Domínio *Bacteria*, a subunidade B da RNA polimerase, o gene *rpoB*, provê maior resolução filogenética que o gene de 16S rRNA

(Roux *et al.*, 2011). Outros genes como o que codifica o fator sigma (σ^{70}) da RNA polimerase *rpoD*, o gene *recA*, que codifica a proteína RecA envolvida na recombinação genética e o gene *gyrB* para a subunidade D da DNA girase são outros exemplos de marcadores filogenéticos (Torsvik e Ovreas, 2006). Estes genes, tal como o que codifica para o 16S rRNA, podem ser utilizados como marcadores filogenéticos por possuírem regiões conservadas e variáveis, uma baixa taxa de transferência lateral e, ao contrário do gene do 16S rRNA, possuem apenas uma cópia por genoma (Roux *et al.*, 2011).

Dado o crescente número de genomas completos, o uso de proteínas ribossomais vem sendo realizado como alternativa ao gene do 16S rRNA com o intuito de obter maior resolução em estudos de filogenia. Tais proteínas possuem as mesmas qualidades que o gene do 16S rRNA para estudos evolutivos, e a concatenação das sequências de tais proteínas permite maior resolução na análise das relações filogenéticas entre organismos do domínio *Archaea* (Brochier-Armanet *et al.*, 2008).

No entanto, dado o maior número de estudos com o gene de 16S rRNA (ou escassez de dados com outros marcadores filogenéticos em bancos), não é possível ainda utilizar somente aqueles marcadores genéticos para a descrição de diversidade. Mesmo que tendencioso, o uso do gene de 16S rRNA é ainda indispensável (Roux *et al.*, 2011) e provavelmente continuará sendo uma importante ferramenta no estudo da diversidade microbiana no ambiente (Hong *et al.*, 2009).

Além disso, a ampla utilização do gene de 16S rRNA resultou em bancos de dados enriquecidos para tal gene, como o *Ribosomal Database Project* (RDP), que tornam cada vez mais precisa a caracterização de organismos presentes em amostras de diferentes origens (Oros-Sichler *et al.*, 2006).

3.2. O Domínio *Archaea*

A análise comparativa de sequências de rRNAs 16S e 18S resultou na proposta de diferenciação dos seres vivos em três reinos primitivos: Eubacteria – todas as bactérias conhecidas, Archaeobacteria – inicialmente representado apenas pelas

bactérias metanogênicas, e os Urcariotos – compreendendo todos os organismos eucarióticos (Woese e Fox, 1977).

Estudos posteriores revelaram que os organismos pertencentes ao reino archaeobacteria apresentavam características distintas daquelas observadas tanto em bactérias típicas, como em eucariotos. Por esta razão, foi proposta a divisão sistemática dos organismos em um táxon superior ao Reino, denominado Domínio. Além disso, tal proposição também sugeria que o termo archaeobacteria fosse abandonado, sendo substituído pelo termo *Archaea*, visando denotar a ausência de relação entre tais procariotos e as bactérias típicas. Assim, os seres vivos passariam a ser classificados em três domínios: *Archaea*, *Bacteria* e *Eucarya* (Woese et al., 1990).

A árvore filogenética resultante da análise comparativa de rRNAs 16S e 18S (Figura 1) revelava que o domínio *Archaea* não consistia no ancestral das bactérias e que tais organismos, embora procarióticos, exibiam estreita relação com membros do domínio *Eucarya*.

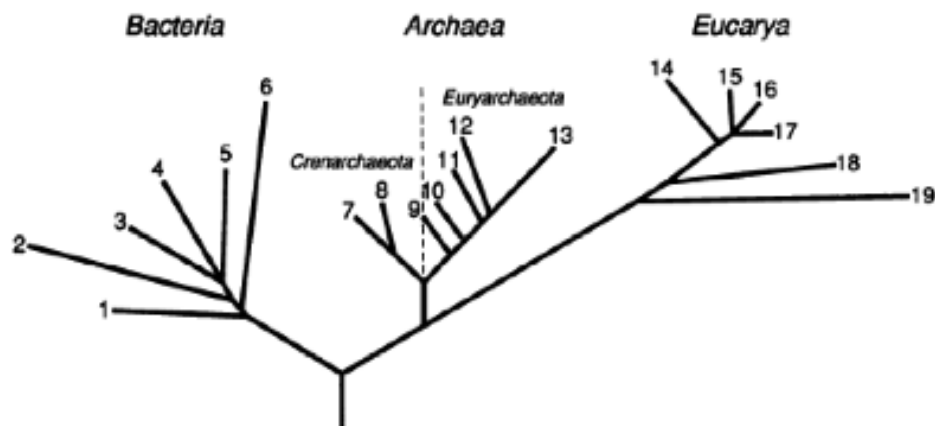


Figura 1. Árvore filogenética universal, baseada em sequências de rRNA 16S e 18S, revelando os Domínios *Archaea*, *Bacteria* e *Eucarya* (Woese et al., 1990).

As archaeas são atualmente consideradas um domínio monofilético, que partilha características com os Domínios *Bacteria* e *Eukarya*, além de possuir características únicas, como pode ser observado na Tabela 1.

Tabela 1. Características gerais de *Bacteria*, *Archaea* e *Eukarya* (adaptado de Madigan *et al.*, 2011)

Características	<i>Bacteria</i>	<i>Archaea</i>	<i>Eukarya</i>
Presença de núcleo	Não	Não	Sim
Parede celular	Peptidoglicano	Não-Peptidoglicano	Não-Peptidoglicano
Lípidios de membrana	Ligação éster	Ligação éter	Ligação éster
Cromossomo	Circular	Circular	Linear
Histonas	Ausentes	Presentes	Presentes
RNA polimerases	Uma (4 subunidades)	Uma (8-12 subunidades)	Três (12-14 subunidades)
Ribossomos (massa)	70S	70S	80S
Iniciador tRNA	Formilmetionina	Metionina	Metionina
Introns	Não	Alguns genes	Sim
Operons	Sim	Sim	Não
5'cap e cauda poli-A no mRNA	Não	Não	Sim
Plasmídeos	Sim	Sim	Raro
Sensibilidade a cloranfenicol, estreptomicina, kanamicina e penicilina	Sim	Não	Não

Uma das características peculiares de archaeas consiste em sua membrana plasmática. Esta é formada por lipídios que, no lugar de ácidos graxos, é composta por cadeias laterais de unidade repetitivas de isopreno, um hidrocarboneto de cinco carbonos, ligadas ao esqueleto principal de glicerol 1-fosfato por meio de ligações do tipo éter (Kates *et al.*, 1993). Algumas archaeas, especialmente as hipertermófilas, apresentam ligações tetraéter entre os lipídeos, possuindo assim membranas do tipo monocamadas, mais resistentes à ruptura (DeRosa *et al.*, 1991).

Quanto à composição das paredes celulares, as archaeas podem apresentar paredes de natureza glicoprotéica, como observado nas halófilas; de pseudomureína, como em várias metanogênicas; ou ainda camadas S como em metanogênicas, dentre vários outros tipos (Kandler e König, 1998). A camada S é um envoltório de natureza protéica observado em quase todas as archaeas, podendo estar presente juntamente

com outros polímeros, incluindo os polissacarídeos pseudomureína e metacondroitina (Albers e Meyer, 2011).

As archaeas são referidas como um mosaico de características eucarióticas e bacterianas, pois seus genes informacionais, ou seja, associados à transcrição, replicação e tradução, são similares àqueles do Domínio *Eucarya*; enquanto os genes operacionais, envolvidos em vias metabólicas, exibem maior semelhança com aqueles presentes no Domínio *Bacteria* (Rivera *et al.*, 1998). No entanto, há exceções como por exemplo, a produção de metano que, até hoje, foi detectada exclusivamente em archaeas.

Apesar do crescente número de estudos relacionados ao isolamento e caracterização de archaeas, grande parte dos genes descritos em tais organismos possui função ainda desconhecida (Graham *et al.*, 2000).

O Domínio *Archaea* é composto por dois filios formalmente aceitos: Euryarchaeota e Crenarchaeota (Woese *et al.*, 1990; Garrity, 2001). No entanto, outros autores propuseram a criação de filios adicionais, “Korarchaeota”, “Nanoarchaeota”, “Thaumarchaeota” e “Aigarchaeota”, ainda não reconhecidos formalmente (Barns *et al.*, 1996; Huber *et al.*, 2002; Brochier-Armanet *et al.*, 2008; Nunoura *et al.*, 2011).

O filo Euryarchaeota consiste em um grupo muito diverso que contém as archaeas metanogênicas, halófilas, alguns organismos redutores de sulfato (como o gênero *Archaeoglobus*) e alguns hipertermófilos (como os gêneros *Thermoplasma* e *Thermococcus*). O filo Crenarchaeota constitui um grupo de organismos inicialmente considerados mais ancestrais, cujo fenótipo apresentava-se mais homogêneo, quando comparado ao filo Euryarchaeota, por conter apenas membros hipertermofílicos (Woese *et al.*, 1990).

Conforme ilustrado nas Figura 2 e 3, a classificação de archaeas em apenas dois filios vem sofrendo modificações, especialmente devido a uma série de trabalhos empregando técnicas de identificação de microrganismos independentes de cultivo. A Figura 2b apresenta uma árvore filogenética construída a partir de sequências do gene de 16S rRNA e pode ser comparada à árvore da Figura 3 que foi construída com base em sequências de aminoácidos de proteínas ribossomais. As relações filogenéticas

entre os organismos representadas em ambas as figuras são similares, no entanto a segunda apresenta maior resolução dos clados.

O sequenciamento de gene de 16S rRNA de amostras oriundas da *Obsidian Pool*, uma fonte termal do Parque Nacional de Yellowstone (EUA), em 1996, revelou a presença de várias archaeas ainda não cultivadas. Dentre essas sequências, duas apresentavam ramificação mais profunda na árvore de *Archaea*, provavelmente anterior à separação dos filos Euryarchaeota e Crenarchaeota. Assim, foi proposto que tais sequências provavelmente formariam um novo filo, denominado “Korarchaeota” (Barns *et al.*, 1996).

Em 2002, foi descrita uma nova archaea, isolada a partir de uma fenda hidrotermal da Islândia, denominada *Nanoarchaeum equitans*. Este organismo apresentava características peculiares, uma vez que encontrava-se em estreita associação com uma outra archaea, a crenarchaeota *Ignicoccus hospitalis*. Além disso, os genes de 16S rRNA de *N. equitans* não eram amplificados em experimentos de PCR que empregavam iniciadores considerados universais para o domínio *Archaea*. Devido a tais características, foi proposto que *N. equitans* representasse um novo filo, denominado “Nanoarchaeota” (Huber *et al.*, 2002). Posteriormente, outros filotipos de “Nanoarchaeota” foram identificados em ambientes hipertermófilos e mesófilos (Casanueva *et al.*, 2008) e a análise da filogenia por meio de proteínas ribossômicas sugeriram que esta espécie é provavelmente uma linhagem de rápida evolução do filo Euryarchaeota, possivelmente relacionada à ordem Thermococcales (Brochier *et al.*, 2005).

Recentemente, outro novo filo foi proposto, denominado “Thaumarchaeota”, o qual seria representado por organismos mesofílicos, anteriormente classificados como Crenarchaeota. Este grupo compreende organismos que possuem genes envolvidos no ciclo do nitrogênio, na etapa da oxidação de amônia. (Brochier-Armanet *et al.*, 2008). Alguns estudos indicam que archaeas são predominantes, em relação a bactérias no processo de nitrificação (Leininger *et al.*, 2006; Adair e Schwartz, 2008), apesar de ter sido mostrado em um solo agriculturável que bactérias dominam a oxidação de amônia (Jia e Conrad, 2009).

Em ambientes de alta temperatura como fontes termais são encontrados organismos putativos termófilos não cultivados pertencentes a um outro grupo irmão das crenarchaeotas. Este grupo foi chamado de *Hot Water Crenarchaeotic Group I* (HWCGI) (Nunoura *et al.*, 2005). A partir de bibliotecas metagenômicas construídas de uma amostra fonte geotermal da subesuperfície de uma mina de ouro foi possível montar o genoma composto do primeiro representante desse grupo, o *Candidatus Caldiarchaeum subterraneum*. A análise deste genoma e a presença de um sistema do tipo ubiquitina, similar ao encontrado em eucarioto, suscitou recentemente a proposta de um novo filo de *Archaea* denominado “Aigarchaeota” (Nunoura *et al.*, 2011). No entanto, há críticas a esta proposta, pois tal genoma codifica características de Thaumarchaeota e poderia ser um ramo antigo deste filo (Brochier-Armanet *et al.*, 2011).

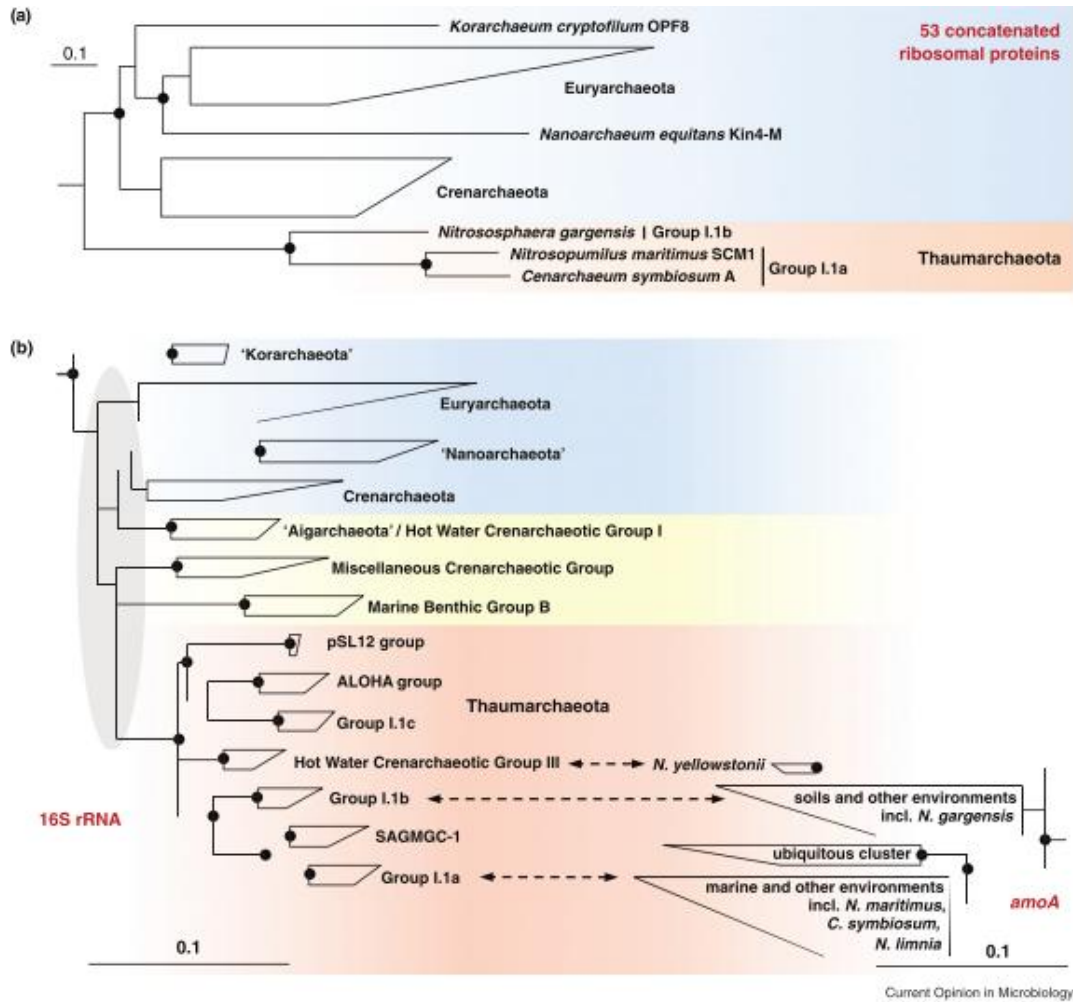


Figura 2. Árvore filogenética de *Archaea* a) construída a partir de 53 proteínas ribossomais concatenadas. b) construída a partir do gene do 16S rRNA (à esquerda) e a partir do gene *amoA* (à direita) (Pester *et al.*, 2011).



Figura 3. Árvore filogenética de *Archaea* baseada em proteínas ribossomais (Brochier-Armanet *et al.*, 2011).

Há vários anos as archaeas passaram a ser consideradas organismos ubíquos, encontrados em uma grande diversidade de ambientes, tais como na cavidade oral e intestino de humanos (Belay *et al.*, 1988; Leep *et al.*, 2004), ecossistemas marinhos costeiros (DeLong, 1992; Furchman *et al.*, 1992), centrais de tratamento de esgoto (Gray *et al.*, 2002), corais (Wegley *et al.*, 2004), solos (Kemnitz *et al.*, 2007), sedimentos (Briée *et al.*, 2007) e lagos de água doce (Llirós *et al.*, 2008), entre outros.

A descrição das comunidades de *Archaea* nos diversos ambientes são baseadas em técnicas independentes de cultivo, especialmente na reação de PCR. Para tanto, são empregados diferentes pares de iniciadores visando amplificar o gene do 16S rRNA, como revisado por Baker (2003) e Gantner (2011).

As sequências de *Archaea* encontradas em ambientes não extremos são divididas em três grupos, sendo o grupo I associado ao filo Crenarchaeota e os grupos II e III pertencentes ao filo Euryarchaeota (DeLong, 1992; DeLong *et al.*, 1998). Apesar de algumas árvores filogenéticas baseadas em SSU rRNA apresentarem as sequências de archaeas mesófilas agrupadas no filo Crenarchaeota (Schleper *et al.*, 2005), Robertson (2005) mostrou que as sequências do grupo I formavam um grupo irmão, dividindo um ancestral comum com crenarchaeotas hipertermófilas, mas não formavam um grupo monofilético com estas. Brochier-Armanet (2008), a partir de árvores filogenéticas de proteínas ribossomais, propôs que o grupo I fosse considerado um grupo relativo às crenarchaeotas mesófilas, denominando-o filo Thaumarchaeota.

O filo Thaumarchaeota proposto é dividido em três grupos: 1.1a, 1.1b e 1.1c, sendo a maioria das sequências de amostras de água pertencente ao grupo 1.1a (DeLong *et al.*, 1998), enquanto que a maior parte das sequências de archaea obtidas de amostras de solo forma o grupo 1.1b (Oschenreiter *et al.*, 2003). As sequências de solo ácido formam outro grupo filogenético, denominado 1.1c, descrito por Jurgens (1997) a partir de amostras de floresta boreal.

Em relação à distribuição e diversidade de archaeas em solos, a literatura permite concluir que estas constituem 5% do total de procariotos em um solo arenoso (Oschenreiter *et al.*, 2003) ou de 12 a 38% em um solo ácido de floresta (Kemnitz *et al.*,

2007), sendo a maioria das sequências encontradas pertencentes ao filo Crenarchaeota (Bintrim *et al.*, 1997).

3.3. Microrganismos no solo

O gene de 16S rRNA passou a ser amplamente utilizado em estudos de ecologia microbiana especialmente após trabalhos indicarem que apenas cerca de 1% dos microrganismos é cultivável pelas técnicas microbiológicas clássicas (Amann *et al.*, 1995; Pace, 1997; Torsvik e Ovreas, 2006). A discrepância observada entre a contagem de microrganismos do ambiente por meio de análises microscópicas e através do cultivo foi designada “grande anomalia da contagem em placa” (Staley e Konopka, 1985). Por meio de microscopia de epifluorescência, foi demonstrado que em 1 grama de solo há mais de 10 bilhões de células procarióticas, sendo possível o cultivo de um número 100 a 1000 inferior de organismos (Torsvik e Ovreas, 2006).

Em um ecossistema, as atividades funcionais e a comunidade microbiana são afetadas por diversos fatores ambientais, como a disponibilidade de carbono e fontes de energia, nutrientes minerais, fatores de crescimento, composição iônica, conteúdo de água disponível, temperatura, pressão, composição de ar, radiação eletromagnética, pH, potencial de oxido-redução, substratos e interações químicas e genéticas entre os microrganismos ali presentes (Nannipieri *et al.*, 2003).

Dada a heterogeneidade do solo, este sistema é considerado o mais diverso em termos de microrganismos. A quantidade de células procarióticas pode variar de 4×10^7 células por grama de solo de floresta, a 2×10^9 células por grama de solo agriculturável, ou de gramíneas (Daniel, 2005).

O solo, de forma geral, é um sistema descontínuo e composto por minerais, componentes orgânicos, água e ar não consolidados. Sua formação resulta de cinco variáveis interdependentes: clima, organismos, material de origem, relevo e tempo (Reatto *et al.*, 1998). Suas transformações ocorrem pela atividade biológica de seres vivos dos três Domínios da vida, sendo os microrganismos responsáveis por 80 a 90% dos processos biológicos que naquele ocorrem (Nannipieri *et al.*, 2003).

Os procariotos são os principais reguladores da ciclagem dos nutrientes e de elementos nos ciclos biogeoquímicos. Apesar dessa ciclagem (ciclos biogeoquímicos do nitrogênio, carbono, enxofre e fósforo) ser mediada majoritariamente por microrganismos, sua taxa depende de componentes da microfauna, como protistas, nematóides e insetos. Após estes revolverem a terra, fungos e bactérias podem atuar na decomposição da matéria orgânica, liberando CO₂ ou CH₄, ou incorporando à matéria orgânica do solo (SOM) (Swift *et al.*, 2008).

Quanto aos papéis desempenhados pelas archaeas no solo muito pouco foi descrito (Van Elsas *et al.*, 2006), sendo a função mais conhecida a de oxidação da amônia, realizada por organismos atualmente classificados como thaumarchaeotas (Spang *et al.*, 2010).

Em relação aos solos brasileiros, a presença de archaeas foi detectada apenas no bioma da Floresta Amazônica (Pazinato *et al.*, 2010), de forma que ainda não existem relatos acerca de outros biomas. Assim, o estudo da diversidade de *Archaea* no solo de Cerrado tem como objetivo inicial avaliar a riqueza desse Domínio em tal ecossistema.

3.4. Solo de Cerrado

O Cerrado consiste na principal vegetação do Brasil Central, estando localizado essencialmente no Planalto Central. Este é o segundo maior bioma do país em área, menor apenas que a Floresta Amazônica. O Cerrado é classificado como uma savana sazonal úmida que cobre cerca 24% da área do Brasil, sendo encontrado em 11 estados brasileiros (Ribeiro e Walter, 1998; Walter *et al.*, 2008).

O clima deste bioma é tropical *Köppen Aw* (tropical chuvoso) com uma precipitação média que varia de 1100 a 1600 mm, apresentando duas estações bem definidas, uma seca - de maio a setembro, e outra chuvosa - de outubro a abril. (Eiten, 1972).

A intemperização do solo pelas chuvas torna o solo um ambiente oligotrófico (Walter *et al.*, 2008), ou seja, com baixo teor de nutrientes, o que pode influenciar quantitativamente e/ou qualitativamente a comunidade microbiana do solo.

Elementos como a textura do solo, a baixa disponibilidade de nutrientes e a boa drenagem, juntamente com a pluviosidade e a frequência de queimadas, influenciam a distribuição florística nos cerrados (Furley e Ratter, 1988), que varia de formações savânicas a florestais e campestres (Figura 4). Esta variação de formações vegetais está relacionada às classes de solo, sendo o latossolo o solo mais comum, constituindo aproximadamente 46% dos solos do bioma Cerrado. Entretanto outras classes de solos também são encontradas, como os solos petroplínticos, cambissolos, podzólicos, plintossolos, gleissolos, areias quartozas, terra roxa estruturada, litólicos e solos orgânicos e aluviais (Reatto *et al.*, 1998).

Os latossolos são solos profundos, bem drenados, distróficos, ou seja, possuem baixa fertilidade e média acidez, com níveis altos de ferro e alumínio, predominantemente argilosos, sendo a maioria deles originário de uma cobertura detritica laterítica datando da era geológica do Terciário (IBGE - <http://www.recor.org.br/cerrado/solos.html>).



Figura 4. Esquema dos tipos de vegetação do bioma Cerrado. Ilustração de José Felipe Ribeiro (Walter *et al.*, 2008).

As características do solo dependem de seus fatores de formação: clima, organismos, material de origem, relevo e tempo. Na Reserva Ecológica do IBGE (RECOR-IBGE), área onde o bioma Cerrado está representado, os solos são diversos, e a classe de latossolos (latossolos vermelho-escuro, vermelho-amarelo e variação una) ocupa cerca de 80% da Reserva.

A flora do cerrado é constituída por muitas espécies endêmicas (Eiten, 1972), sendo a biodiversidade de forma geral considerada a mais rica dentre as savanas do mundo (www.ibama.gov.br).

A vegetação do Cerrado é semidecidual, com um estrato arbóreo descontínuo, com predomínio de árvores baixas (3 a 8 metros de altura), e um estrato arbustivo e outro herbáceo. O Cerrado Denso é uma variação do Cerrado *sensu strictu*, por possuir um estrato arbóreo denso, sendo a cobertura majoritariamente de árvores de copa larga, com 5 a 12 metros de altura. Na RECOR-IBGE esta fitofisionomia está sob um tipo de latossolo vermelho-escuro argiloso. O tipo florestal de mata de Galeria ocorre ao longo de cursos d'água e sua vegetação irá variar de acordo com o grau de umidade e fertilidade do solo (IBGE - <http://www.recor.org.br/cerrado/vegetacao.html>).

Alguns trabalhos descrevendo a comunidade microbiana de Cerrado foram realizados, muitos utilizando como modelo a Reserva Ecológica do IBGE, onde é possível encontrar mais de uma fitofisionomia deste bioma. A comunidade fúngica foi descrita em cerrado *sensu stricto* e na mata de galeria e também em áreas de Cerrado transformadas em pastagem e em cultivo com soja (Castro *et al.*, 2008). A comunidade bacteriana de solo de Cerrado foi também avaliada (Silva, 2004; Quirino *et al.*, 2009; Araújo, 2011). Em outro trabalho, foi estabelecida uma comparação das comunidades bacteriana e fúngica entre um solo de Cerrado nativo e outro de Cerrado sob monocultura de soja (Bresolin, 2010).

Tendo em vista a escassez de dados acerca da distribuição e diversidade de *Archaea* no ecossistema de cerrado, esta dissertação tem como objetivo realizar uma primeira avaliação sobre a riqueza do Domínio *Archaea* no solo de cerrado.

4. Materiais e Métodos

4.1. Coleta do solo

O solo foi coletado, com réplica biológica, na estação de transição chuvosa/seca (no dia 24 de setembro de 2010) de duas fitofisionomias do bioma Cerrado - áreas nativas não perturbadas: Cerrado Denso e Mata de Galeria, na Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR-IBGE) localizada em Brasília – DF. Esta Reserva está localizada 35 km ao sul do centro de Brasília – DF, na BR-251, km 0 (15°55'S, 47°51'W). A RECOR faz parte da Área de Proteção Ambiental (APA) Distrital Gama-Cabeça de Veado.

As duas fitofisionomias amostradas, em replicata, foram: Cerrado denso (CD), Mata de galeria (MG). Cerrado Denso apresenta uma vegetação principalmente arbórea, com cobertura de 50 a 70% e altura média de cinco a oito metros. A Mata de Galeria é uma vegetação florestal perene, presente ao longo de cursos d'água de pequeno porte, com um estrato arbóreo de altura média de 20 a 30 metros e cobertura arbórea de 70 a 95%. (Ribeiro e Walter, 1998).

Em cada fitofisionomia a coleta foi realizada em 5 pontos com distância não-linear de 50 m entre estes. Em cada um dos cinco pontos foram realizadas dez sub amostras, coletadas a 10 cm de profundidade (Figura 5), peneiradas em malha de 2 mm, armazenadas na forma de amostra única, em sacos mantidos em freezer -80°C. Para cada fitofisionomia foram traçados dois transectos de 5 pontos (coleta em duplicata) com distância entre eles aproximadamente de 50 m. O material coletado em cada transecto de cinco pontos foi então reunido, originando uma amostra composta, a qual foi utilizada na etapa de extração de DNA genômico total. Desta forma foram obtidas 4 amostras compostas, sendo duas de cada fitofisionomia, constituindo a réplica biológica. As amostras de solo foram coletadas e gentilmente cedidas pelo grupo de pesquisa coordenado pelo Prof. Dr. Ricardo H. Krüger, da Universidade de Brasília.

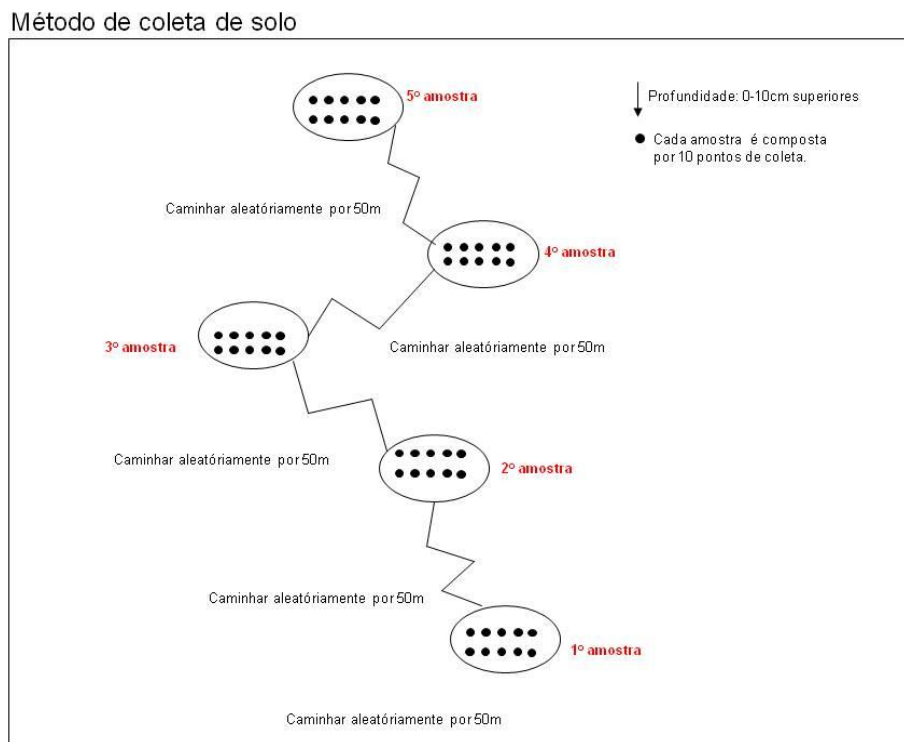


Figura 5. Esquema da amostragem de um transecto. Cedido por Alinne Castro.

4.2. Extração de DNA genômico total

A extração de DNA de solo foi realizada empregando-se o *PowerSoil DNA Isolation Kit* (MO Bio Laboratories Inc.), segundo o protocolo do fabricante. Embora no protocolo original fosse preconizado o uso de 0,25 g de solo em cada procedimento de extração de DNA total, optamos por utilizar 0,5 g, visto que tal quantidade mostrou-se adequada em experimentos prévios realizados pelo grupo do Prof. Dr. Ricardo Krüger. O DNA resultante foi avaliado quanto à concentração e pureza por meio de eletroforese em gel de agarose 1%, corado com brometo de etídio (10 mg/mL).

4.3. Iniciadores utilizados e condições de PCR

Neste trabalho foram utilizados dois pares de iniciadores descritos como universais para *Archaea*: 21F e 958R (DeLong, 1992) e 751F e 1406R (Baker et al., 2003), apresentados na Tabela 2.

Tabela 2. Iniciadores utilizados

Iniciador	Sequência	T _m
21F	5' TTC CGG TTG ATC CYG CCG GA 3'	62,6°C
958R	5' YCC GGC GTT GAM TCC AAT T 3'	56,6°C
751F	5' CCG ACG GTG AGR GRY GAA 3'	54 °C
1406R	5'ACG GGC GGT GWG TRC AA 3'	52 °C

Os iniciadores utilizados neste trabalho amplificam regiões variáveis diferentes do gene do 16S rRNA como apresentado na Figura 6.



Figura 6. Esquema do gene de 16S rRNA, suas regiões variáveis e as posições de anelamento dos iniciadores.

Os experimentos de PCR foram realizados em sistemas de 50 µL, contendo 5 µL de DNA total a 1ng/µL (diluição 1:10) e os demais reagentes, conforme descrito na tabela 3 para ambos os pares de iniciadores.

Tabela 3. Reagentes e respectivas concentrações finais na reação de PCR

Reagente	Concentração final
Tampão de reação (Invitrogen®)	1 X
MgCl ₂ (Invitrogen®)	1,5 mM
Soroalbumina bovina (BSA)	400 ng/µL
Iniciador F	0,5 µM
Iniciador R	0,5 µM
dNTPs	200 µM
TaqDNA polimerase (Invitrogen®)	1,25 U

Os experimentos de PCR foram realizados no termociclador MJ Research PTC-100 (Peltier Thermal Cycles), empregando-se as condições descritas abaixo:

- 1) Desnaturação a 94°C por 1 minuto (2 minutos para o par 751F/1406R)
- 2) Desnaturação a 94°C por 1 minuto
- 3) Anelamento 55°C por 1 minuto (56°C para o par 751F/1406R)
- 4) Extensão 72°C por 1 minuto e 30 segundos (1 minuto para o par 751F/1406R)
- 5) Repetição dos passos 2 a 4 por 29 vezes
- 6) Extensão 5 minutos (10 minutos para o par 751F/1406R)
- 7) Manutenção a 4°C

Em todos os experimentos de PCR realizados foram utilizados dois controles negativos: um tubo onde adicionou-se água em substituição ao DNA, e outro tubo contendo DNA de *Escherichia coli*.

Os produtos resultantes da amplificação foram analisados por meio de eletroforese em gel de agarose 1%, corado com brometo de etídio (10 mg/mL). O tamanho dos produtos amplificados foi estimado com auxílio do marcador de massa molecular *1kb DNA ladder* (Invitrogen®).

Para realizar a transformação por eletroporação, um conjunto de 6 reações de DNA produto de PCR foi purificado com kit Wizard® SV Gel and PCR Clean-Up System (Promega, Wisconsin, USA).

4.4. Purificação e ligação das sequências de rDNA em vetor pGEM-T Easy

Os produtos amplificados por PCR foram inicialmente submetidos à purificação, empregando-se o kit Wizard® SV Gel and PCR Clean-Up System (Promega, Wisconsin, USA). Em seguida, estes foram ligados ao vetor pGEM-T Easy® (Promega, Wisconsin, EUA) na proporção de 3:1 de inserto:vetor, conforme orientações do fabricante. Os sistemas de reação foram incubados por 1 hora à temperatura ambiente e posteriormente armazenados em geladeira, por uma noite.

4.5. Preparação de células competentes de *Escherichia coli* DH5 e transformação

4.5.1. Por choque térmico

Uma colônia isolada de *E. coli* DH5 foi inoculada em 5 mL de meio LB (10 g NaCl, 10 g peptona ou triptose, 5 g extrato de levedura, q.s.p. 1000 mL, pH 7,2) e incubada por uma noite em estufa a 37°C. Após este período, 300 µL desta cultura foram inoculados em um frasco contendo 30 mL de LB, o qual foi incubado a 37°C sob agitação (220 rpm), até a cultura atingir a OD₆₀₀ de 0,2 a 0,3. A cultura foi então centrifugada a 1250 g, por 10 minutos, a 4°C. Após a centrifugação, o sobrenadante foi descartado e o sedimento cuidadosamente ressuspensão em 15 mL de CaCl₂ 100 mM gelado.

As células foram então incubadas em gelo por 20-30 minutos e então centrifugadas a 310 g, por 10 minutos, a 4°C. O sobrenadante foi descartado e o sedimento ressuspensão em 1 mL de CaCl₂ 100 mM gelado. As células assim tratadas foram novamente incubadas em gelo por um período mínimo de 1 hora, quando então podiam ser utilizadas em procedimentos de transformação.

Alíquotas de 100 µL das células tratadas com CaCl₂ foram transferidas para tubos de vidro, mantidos em gelo. A cada sistema de transformação foi adicionado um volume correspondente à metade do sistema de ligação (5 µl) preparado anteriormente, sendo o conjunto incubado em gelo por 30 minutos. Após este período os tubos foram rapidamente transferidos para um banho a 37°C e incubados nesta condição por 5 minutos. Em seguida, acrescentou-se 1 mL de meio LB a cada tubo, sendo estes incubados em estufa a 37°C, durante 1 h. Após este período alíquotas de 25 a 150 µL de cada sistema de transformação foram inoculadas em placas contendo ágar LB (LB adicionado de 1,5% de ágar) suplementado com IPTG (0,5 mM), Xgal (0,00625%), ampicilina (150 µg/mL), sendo a semeadura realizada por espalhamento, com auxílio de pérolas de vidro estéreis.

As placas foram incubadas em estufa a 37°C por 16 a 24 h, até a observação do desenvolvimento de colônias brancas e/ou azuis.

4.5.2. Por eletroporação

Uma colônia de *E. coli* DH5 foi inoculada em 10 mL de meio LB (10 g NaCl, 10 g peptona ou triptose, 5 g extrato de levedura, q.s.p. 1000 mL, pH 7,2) e incubada por uma noite em estufa a 37°C. Após este período, alíquotas de 5 mL desta cultura foram transferidas para frascos contendo 200 mL de meio LB, incubados a 37°C sob agitação (220 RPM), até a cultura atingir a OD₆₀₀ de 0,4. Em seguida, os frascos foram mantidos por 20 minutos em gelo, sendo posteriormente centrifugados a 1180 g, por 20 minutos, a 4° C. Após a centrifugação, o sedimento de células foi cuidadosamente ressuspensão em 50 mL de H₂O Milli-Q gelada, em banho de gelo, adicionando-se em seguida 150 mL de H₂O Milli-Q gelada. As duas suspensões de células foram novamente centrifugadas sob as mesmas condições e o sedimento ressuspensão em 50 mL de glicerol 10% gelado. Em seguida, foram adicionados 50 mL de glicerol 10% gelado, sendo os sistemas novamente centrifugados. Os sedimentos foram ressuspensionados em 5 mL de glicerol 10% gelado, reunidos e acondicionados em um tubo Falcon previamente resfriado. Após nova centrifugação, o sedimento final de células foi ressuspensionado em 1 mL de glicerol 10% gelado. Uma alíquota de 10 µL foi diluída 1:100 e a concentração de células foi avaliada em espectrofotômetro, a qual foi corrigida para $2 - 3 \times 10^{10}$ células/mL. As células competentes foram separadas em alíquotas de 50 µL em tubos tipo eppendorf de 1 mL e armazenadas a -80°C.

Um volume de 2 µL do sistema de ligação a partir de produto de PCR purificado foi adicionado à alíquota de 50 µL. Este sistema foi transferido para uma cubeta com 1 mm de largura. A eletroporação foi realizada de acordo com os parâmetros apresentados por Maranhão (2003): 2,5 kV, 25 µF, 200 . O t obtido variou de 4,5 a 5 ms entre as amostras.

A preparação de células competentes e a posterior transformação por eletroporação foram realizadas no laboratório de Virologia Vegetal do prof. Renato Rezende, com o auxílio de seu aluno de mestrado André Bertran.

Para ambas as técnicas de transformação foi utilizado um plasmídeo fechado PRX1 a 10 ng/ µL e 100 pg/µL, como controle de eficiência da transformação.

4.6. Seleção de clones recombinantes

Números variáveis de colônias brancas foram selecionadas e inoculadas em 5 mL de LB contendo 100 µg/mL de ampicilina. Após a incubação em estufa a 37°C por uma noite, foi retirada uma alíquota da cultura, a qual foi misturada ao mesmo volume de glicerol 70%, e armazenada em freezer -20°C.

4.7. Extração de DNA plasmidial em pequena escala

O DNA plasmidial foi extraído inicialmente de alguns clones com o kit *PureLink™ Quick Plasmid Miniprep* (Invitrogen) e da maioria restante com o método de lise alcalina. O DNA de alguns clones aleatórios foi utilizado para avaliar sua qualidade e pureza por meio de eletroforese em gel de agarose 1%, corado com brometo de etídio. O DNA plasmidial foi quantificado por meio de comparação com os fragmentos de DNA presentes no marcador de massa molecular *High Mass DNA ladder* (Invitrogen®).

4.7.1. *PureLink™ Quick Plasmid Miniprep* (Invitrogen)

A extração de DNA plasmidial dos clones recombinantes foi realizada empregando-se o kit *PureLink™ Quick Plasmid Miniprep* (Invitrogen), de acordo com o protocolo fornecido pelo fabricante.

4.7.2. Protocolo de lise alcalina (adaptado de Dagert e Ehrlich, 1979)

Os clones selecionados foram inoculados em 5 mL de meio LB suplementado com ampicilina (100 µg/mL), e incubados a 37°C por uma noite. A cultura foi transferida para um tubo tipo eppendorf de 1,5 mL e centrifugada a 3500 g por 5 minutos. O sobrenadante foi descartado. O processo foi repetido outras duas vezes com o intuito de obter um maior sedimento de células para extração.

O sedimento de células foi ressuscitado com 100 µL da solução I (50 mM de Glicose, 25 mM de Tris-HCl, 10 mM EDTA, pH 8). A solução II (0,2 M NaOH, 1% SDS) foi preparada no momento do uso e 200 µL foram adicionados à ressuspensão anterior. Os tubos foram incubados no gelo por 5 minutos. Por fim, 150 µL de solução III (3 M

Acetato de Potássio, 2 M ácido acético glacial, pH 4,8) foram adicionados e uma nova incubação de 5 minutos no gelo foi realizada. Os tubos foram centrifugados a 12 000 g por 5 minutos a 4° C. O sobrenadante foi transferido para um novo tubo eppendorf de 1,5 mL, ao qual foi adicionado 4 µL de RNase A (20 mg/mL). Os tubos foram incubados a 37°C por 1 a 2 horas. Em seguida, foi adicionado 1 volume de clorofane (fenol, clorofórmio, álcool iso-amílico, na proporção 25:24:1) e os tubos submetidos à agitação em vortex durante 5 segundos. Foi realizada uma centrifugação a 7000 g por 3 minutos à temperatura ambiente. O sobrenadante foi transferido para um novo tubo, ao qual foi adicionado de 2,5 a 3 V de etanol 100% gelado. Os tubos foram mantidos no freezer - 20° C por 24h. Após este período, os tubos foram novamente centrifugados a 12 000 g por 10 minutos a 4° C, sendo os tubos invertidos para o descarte do etanol. Então foram adicionados 400 µL de etanol 70% gelado. E uma nova centrifugação nas mesmas condições anteriores foi realizada. O tubo foi novamente invertido e deixado para secar à temperatura ambiente. O DNA sedimentado foi ressuspendido em 50 µL de H₂O milli-Q.

4.8. Quantificação da concentração do DNA

O DNA foi quantificado no espectrofotômetro Nanovue™ (GE), gentilmente disponibilizado pela professora Silviene de Oliveira, do Laboratório de Genética.

4.9. Sequenciamento de DNA

Os fragmentos de DNA clonados em pGEM-T Easy® (Promega, Wisconsin, EUA) foram sequenciados no Laboratório de Biologia Molecular da Universidade de Brasília em placas de 96 poços, no seqüenciador automático ABI 3130 XL Genetic Analyzer, da Applied Biosystems, com utilização do kit ABI Prism BigDye Terminator v.3.1 Cycle Sequencing kit. Para o sequenciamento foi empregado o mesmo iniciador 21F utilizado para a reação de PCR.

4.10. Análises de Bioinformática

As sequências geradas passaram por uma primeira seleção de qualidade com o algoritmo PHRED (Ewing, 1998), que analisa e verifica a probabilidade de erro base a base. Foram selecionadas apenas sequências que apresentavam mais de 250 nucleotídeos, cujo PHRED apresentava valor superior a 20, isto é: com 99% de probabilidade de estar correta. Tal análise foi realizada *online*, empregando-se a ferramenta *Electropherogram quality analysis*, disponível na página da EMBRAPA (<http://asparagin.cenargen.embrapa.br/phph/>). As sequências foram também submetidas à análise pelo programa Bellerophon (Huber *et al.*, 2004) (<http://comp-bio.anu.edu.au/bellerophon/bellerophon.pl>), para remover artefatos de sequências quiméricas. Tanto as sequências com qualidade inferior a 20 quanto as quiméricas foram descartadas.

As sequências consideradas de qualidade foram então submetidas a análises comparativas nos bancos de dados não redundantes *blastn* do *National Center for Biotechnology Information* (NCBI) (Altschul *et al.*, 1990) e *Ribosomal Database Project* (Michigan State University) - RDP (<http://rdp.cme.msu.edu>) – com o software Classifier (Wang *et al.*, 2007). As sequências que apresentaram identidade igual ou superior a 90% com aquelas depositadas nos bancos de dados foram consideradas e nomeadas.

As sequências de boa qualidade foram alinhadas automaticamente com o algoritmo NAST (DeSantis *et al.*, 2006) disponível no *site* do greengenes (<<http://greengenes.lbl.gov/cgi-bin/nph-index.cgi>> (DeSantis *et al.*, 2006). Com o programa MOTHUR (Schloss *et al.* 2009), um filtro foi aplicado para excluir todas as colunas que consistiam em lacunas (*gaps*), resultantes do alinhamento realizado pelo NAST. O alinhamento foi editado manualmente com auxílio do programa Bioedit (<<http://www.mbio.ncsu.edu/bioedit/bioedit.html>>). As regiões hipervariáveis, com espaços de difícil edição, foram eliminadas, assim como as sequências curtas, com menos de 400 pares de bases.

O método utilizado para construção da árvore filogenética foi o *Neighbour-Joining*, modelo de evolução Jukes-Cantor e teste de *bootstrap* de 1000 repetições, para verificar a robustez da árvore, no programa MEGA (Tamura *et al.*, 2011). O programa

MOTHUR foi utilizado para excluir redundância de sequências pertencentes à mesma OTU em nível de espécie (3%).

A estimativa da riqueza de unidades taxonômicas operacionais (OTUs) foi obtida pelo programa MOTHUR, que foi utilizado para a construção da curva de rarefação – que indica o grau de amostragem da comunidade ao relacionar o número de sequências e o número de OTUs formadas por elas -, empregando-se como limites de cortes os valores de 80% para o nível taxonômico de filo, 90% para classe, 95% para gênero e 97% para espécie.

A riqueza de espécies foi calculada com os estimadores não-paramétricos Ace (*abundance-based coverage estimator*) e Chao1 e a diversidade com os índices de Shannon e Simpson, através do programa MOTHUR, utilizando como base a matriz de distâncias computada para a curva de rarefação.

O teste de hipótese -Libshuff foi utilizado para o cálculo estatístico de significância de diferença entre as comunidades, a partir da matriz de distância entre cada par de amostras pelo Mothur. O valor- foi ajustado pela correção de Bonferroni, de significância 0,05 dividido pelo número de amostras comparadas.

O gráfico de Análise de Coordenadas Principais (PCA) foi construído no programa Unifrac (Lozupone et al., 2006), também para análise estatística de -diversidade. Para tal, foi utilizada uma árvore filogenética construída no MEGA, com o método descrito anteriormente e convertida para formato Newick. A sequência de uma *Acidobacteria* não cultivada foi utilizada para enraizamento da árvore.

5. Resultados e Discussão

5.1. Coleta de solos

Os solos foram coletados em setembro de 2010, no período correspondente à transição seca-chuvosa. Os índices pluviométricos mensais registrados na Estação Meteorológica da Reserva Ecológica do IBGE no ano de 2010 (Figura 7) revelam que a coleta foi realizada em um período de déficit hídrico na maioria dos solos, uma vez que a última precipitação foi verificada no mês de junho.

A comunidade microbiana do solo do Cerrado é alterada conforme as estações de chuva e seca (Silva, 2004), portanto deve-se considerar que este trabalho tem como objetivo descrever a comunidade de *Archaea* no período de seca, e que possivelmente a composição desta comunidade sofre modificações durante a época de chuvas.

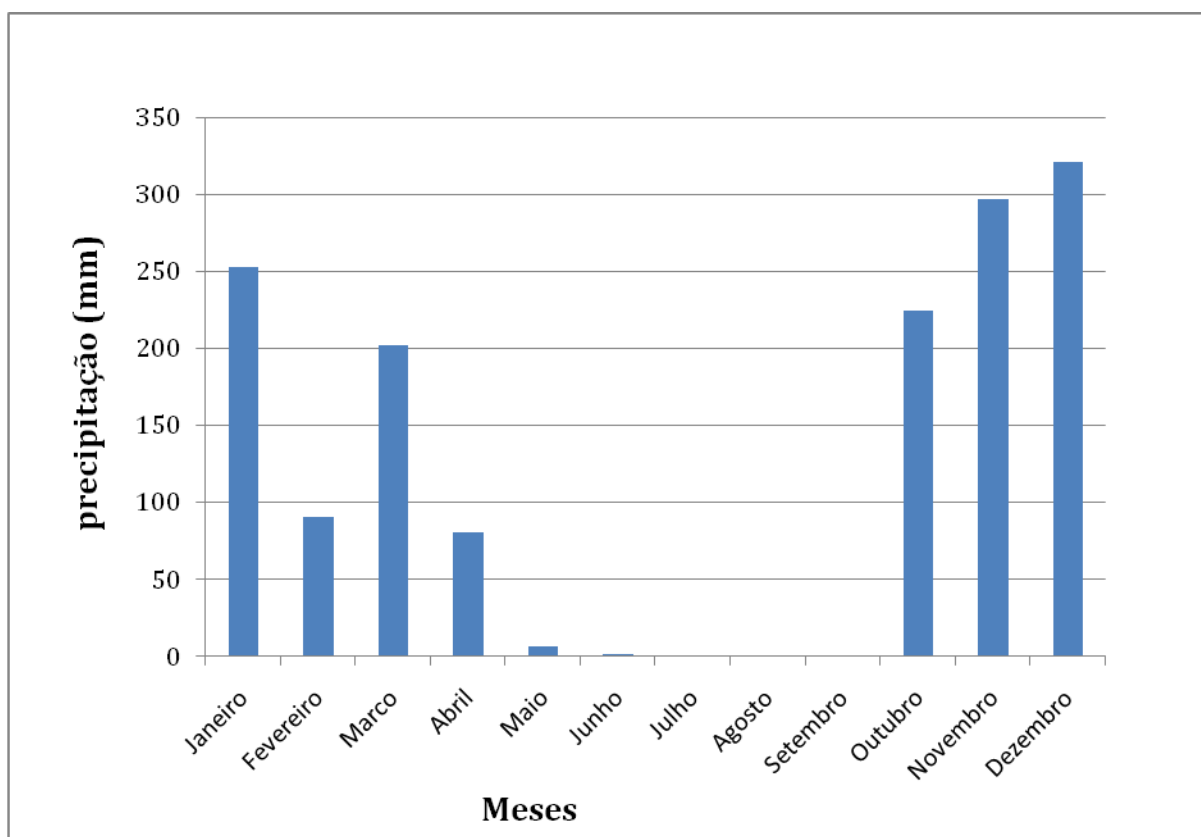


Figura 7. Precipitação mensal (em mm) registrada na Reserva Ecológica do IBGE Brasília-DF, em 2010. (<http://www.recor.org.br/Estacao/consulta.asp>)

Segundo a medição de pH aferida pela empresa SoloQuímica – Análises de Solo Ltda., os solos coletados de cerrado denso (CD) e a réplica A de mata de galeria (MG) apresentam média acidez, com pH 5,0 e a amostra MGb apresentou pH de 4,5 (Tabela 4).

Tal como encontrado por Araújo (2011), em termos de granulometria, o solo de mata de galeria apresenta menor proporção de argila e maior de areia que aquela encontrada para o solo do cerrado. Também as quantidades de fósforo (P), carbono orgânico, matéria orgânica, cálcio (Ca), alumínio (Al) e Nitrogênio (total e nas formas de amônia e nitrato) são maiores para o solo de mata de galeria que para o solo do cerrado denso. Alguns desses parâmetros apresentam valores discrepantes entre as réplicas, mas a média de tais quantificações são similares àquelas descritas por Araújo (2011) para solos das mesmas áreas.

Tabela 4. Propriedades físico-químicas das amostras de solo coletadas na Reserva Ecológica do IBGE realizadas pela SOLOQUIMICA

Parâmetro	CdA	CdB	MgA	MgB
Argila (g/kg)	525	500	450	475
Areia (g/kg)	275	300	425	425
Silte (g/kg)	200	200	125	100
pH em água	5.00	5.00	5.00	4.50
C (g/kg)	45.10	42.80	83.20	14.30
MO (g/kg)	77.60	73.60	143.10	24.60
P (ppm)	0.30	1.50	2.00	2.70
Ca (cmol _c /dm ³)	0.30	0.30	0.60	0.40
Mg (cmol _c /dm ³)	0.20	0.20	0.30	0.30
K (cmol _c /dm ³)	0.02	0.02	0.02	0.02
Na (cmol _c /dm ³)	0.01	0.01	0.01	0.01
Al (cmol _c /dm ³)	0.90	1.00	2.20	2.80
H + Al (cmol _c /dm ³)	8.40	8.40	10.50	14.10
C.T.C (cmol _c /dm ³)	8.93	8.93	11.43	14.83
Nitrogênio total (mg/kg)	3.25	2.38	6.12	3.62
Nitrato - NO ₃ (mg/kg)	1.54	1.03	4.54	1.69
Nitrito - NO ₂ (mg/kg)	0.88	0.75	0.75	0.88
Nitrogênio amoniacal - NH ₃ (mg/kg)	0.63	0.50	0.63	0.75

5.2. Extração de DNA total

O procedimento de extração de DNA total das amostras de solo das duas fitofisionomias mostrou-se bastante satisfatório, uma vez que a análise por meio de eletroforese em gel de agarose 1% corado com brometo de etídio revelou a presença de fragmentos de DNA de alta massa molecular, com pequeno grau de degradação em todas as amostras (Figura 8).

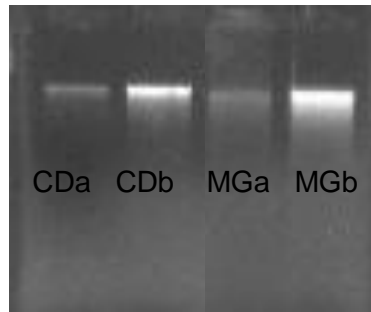


Figura 8. Perfil eletroforético, em gel de agarose 1%, corado com brometo de etídio, do DNA metagenômico extraído das amostras de solo das fitofisionomias cerrado denso (CDa, CDb) e mata de galeria (MGa, MGb), em duplicata.

A concentração do DNA extraído em cada amostra foi aferida no espectrofotômetro Nanovue™ (GE), cujos resultados são apresentados na Tabela 5.

Tabela 5. Concentração de DNA extraído das amostras de solo.

Amostra	Concentração
CDa	7,7 ng/ μ L
CDb	11,4 ng/ μ L
MGa	12,2 ng/ μ L
MGb	20,5 ng/ μ L

O *PowerSoil Kit* (MO Bio) consiste em um procedimento de extração direta, onde as células são lisadas ainda em meio à matriz do solo e a separação de DNA dos restos celulares e da matriz ocorre uma etapa posterior (Ogram *et al.*, 1987). Apesar deste método promover maior recuperação de DNA, este é obtido com menor grau de pureza quando comparado ao procedimento indireto, onde a separação das células é anterior à lise (Gabor *et al.*, 2003).

A extração por lise direta de células aumenta a quantidade de ácidos húmicos e outras substâncias orgânicas contaminantes presentes na matriz do solo que interferem nas reações enzimáticas subseqüentes, como a de amplificação do DNA por PCR ao inibir a atividade da *Taq* polimerase (Roose-Amsaleg *et al.*, 2001). No entanto, a extração com o *PowerSoil Kit* (MOBio) permite uma retirada eficaz dos ácidos húmicos que, quando presentes, se apresentam como uma banda de cor castanha no gel de agarose. A adição de soroalbumina bovina (BSA) na concentração de 400ng/μL diminui a inibição da reação de PCR por substâncias orgânicas provenientes do solo (Kreader, 1996).

5.3. Amplificação de sequências de DNA relativas ao gene de 16S rRNA

Trabalhos previamente realizados em nosso laboratório empregaram o par de iniciadores 751F/1406R (Tabela 2), descrito como universal para os genes de 16S rRNA de *Archaea* por Baker *et al.* (2003). No entanto, quando os produtos de PCR foram submetidos ao sequenciamento, verificou-se que praticamente todas as sequências geradas pertenciam ao domínio *Bacteria*. Desta forma concluiu-se que tais iniciadores não eram adequados para a detecção de membros do Domínio *Archaea*.

Em 1992, DeLong detectou a presença de membros de *Archaea* em ambientes marinhos costeiros, por meio de experimentos de PCR empregando um novo par de iniciadores, 21F e 958R, que amplifica as regiões variáveis V1 a V5 (Figura 6) do gene de 16S rRNA. Este par de iniciadores foi utilizado com sucesso em outro trabalho desenvolvido por nosso grupo, cujo objetivo visava a detecção de *Archaea* a partir de amostras de sedimentos de água doce do Parque Nacional Sempre-Vivas. Os resultados obtidos revelaram que este par de iniciadores permitiu a amplificação de sequências de rDNA tanto de Crenarchaeota e Euryarchaeota, demonstrando seu potencial para a amplificação de sequências do domínio *Archaea* de outros ambientes. Além disso, o par de iniciadores descrito por DeLong (1992) vem sendo utilizado na descrição de comunidades de *Archaea* em diferentes ecossistemas (Hansel *et al.*, 2008; Pazinato *et al.*, 2010), inclusive de rúmen bovino (Skillman *et al.*, 2006). Assim,

tal par foi escolhido para os ensaios de PCR das amostras provenientes de solo de Cerrado.

O DNA extraído das amostras de solos – duplicatas de cerrado denso e mata de galeria – foi utilizado para a amplificação do fragmento do gene de 16S rRNA por PCR, com obtenção de fragmentos de DNA de tamanho esperado (cerca de 940pb) em todas as amostras, como ilustrado na Figura 9. A eficiência da reação de amplificação foi similar entre as amostras. O resultado obtido com o controle negativo contendo DNA de *E. coli*, indica que o par de iniciadores utilizados foi bastante específico, uma vez que não foi observada a presença de produtos de amplificação no gel.

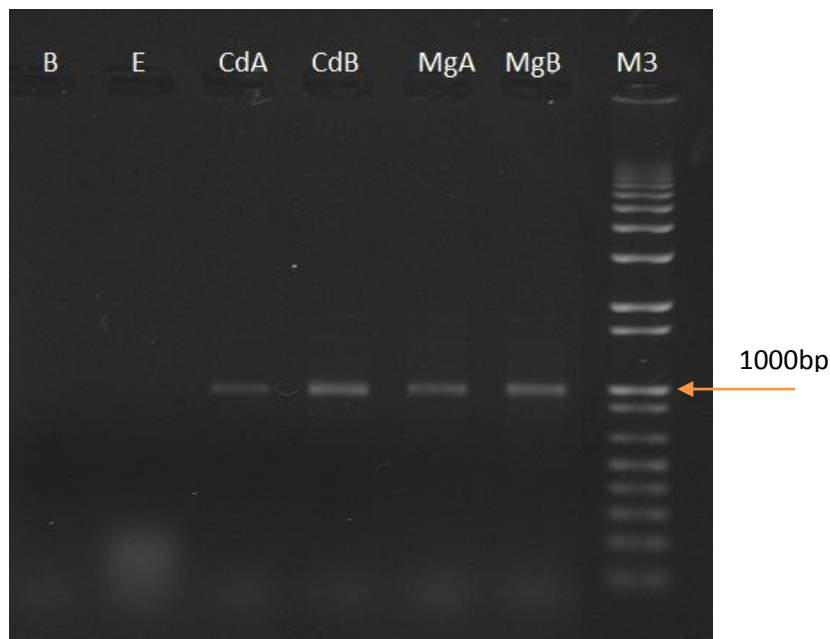


Figura 9. Perfil eletroforético, em gel de agarose 1% corado com brometo de etídio, dos produtos amplificados no experimento de PCR empregando os iniciadores 21F e 958R. B – “branco”, controle negativo, sem DNA; E – DNA de *Escherichia coli*; CD – cerrado denso, réplicas A e B; MG – mata de galeria, réplicas A e B, M3 – Marcador de massa molecular *1kb DNA ladder* (Invitrogen).

5.4. Transformação de *E. coli* DH5 com fragmentos do gene de 16S rRNA e seleção de clones recombinantes

Os fragmentos de DNA amplificados foram ligados ao vetor pGEM-T Easy® (Promega, Wisconsin, EUA), conforme orientações do fabricante. Os plasmídeos foram inseridos em células de *E. coli* da linhagem DH5 por meio da técnica de transformação por choque térmico e por eletroporação, alternativamente.

Alíquotas de volume variável (25 a 200µL) do sistema de transformação (1 mL) foram semeadas em placas de ágar LB suplementado com Xgal, IPTG e ampicilina, que foram incubadas por 16h a 37°C. As colônias brancas foram selecionadas e estocadas em LB com glicerol 35%, a -20°C. Foram obtidas 485 colônias da amostra CDa, 485 de MGa, 300 de CDb e 344 de MGb.

5.5. Extração de DNA plasmidial em pequena escala

O DNA plasmidial dos clones selecionados de cada amostra de solo foi extraído empregando-se o *PureLink™ Quick Plasmid Miniprep* (Invitrogen), ou pelo método de lise alcalina.

A concentração e pureza dos DNAs obtidos foram analisadas por meio de eletroforese de gel de agarose corado com brometo de etídio, como ilustrado na Figura 10. O marcador de massa molecular *High Mass DNA ladder* foi utilizado como parâmetro para estimar a concentração do DNA plasmidial de clones aleatórios e também para confirmar os valores de concentração obtidos por espectrofotometria.

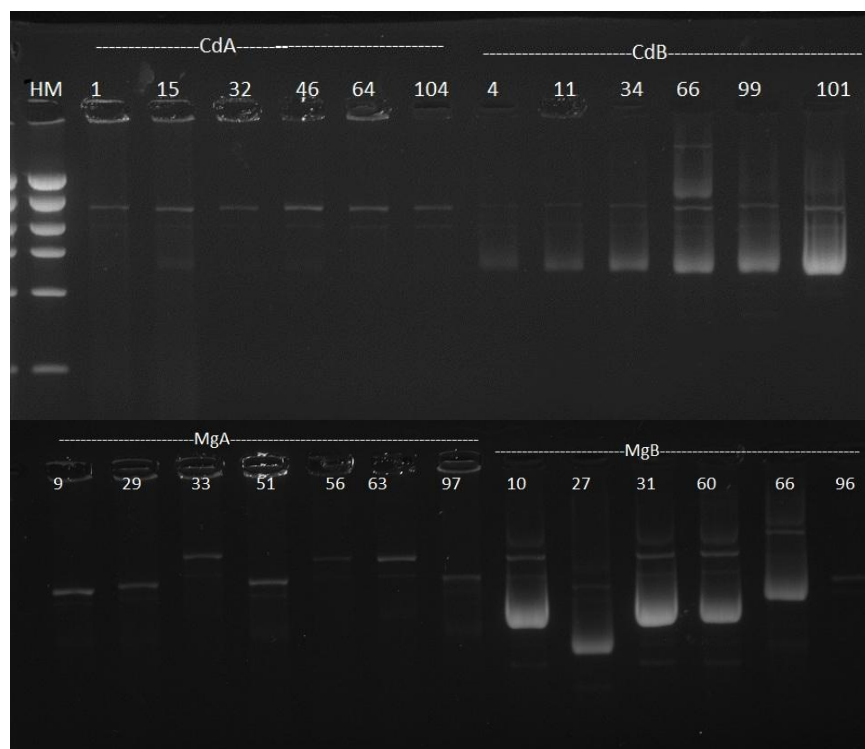


Figura 10. Perfil eletroforético, em gel de agarose 1% corado com brometo de etídio, do DNA plasmidial de clones aleatoriamente selecionados a partir das bibliotecas de rDNA das amostras de solo. HM – *High Mass DNA ladder* (Invitrogen), CdA e CdB – cerrado denso A e B, MgA e MgB – mata de galeria A e B. Cada poço apresenta a nomenclatura das colônias selecionadas.

5.6. Análises de bioinformática

5.6.1. Análise de qualidade das sequências

Um total de 552 clones, oriundos das quatro bibliotecas de DNA, foram submetidos ao sequenciamento automático, empregando o iniciador 21F, específico para *Archaea*. Destes, apenas 275 sequências apresentaram Phred (Ewing *et al.* 1998) de qualidade superior a 20 em mais de 250 nucleotídeos e, por esta razão, as demais sequências foram descartadas. Outras 13 sequências foram também eliminadas por serem curtas demais, uma vez que continham menos de 400 pares de bases. Assim, as análises de bioinformática foram realizadas empregando-se as 262 sequências - 51 de CdA, 65 de CdB, 75 de MgA e 71 de MgB - que apresentavam qualidade adequada para a avaliação da riqueza de *Archaea* no solo do Cerrado.

A análise das sequências com o programa *Bellerophon* (Huber *et al.*, 2004) revelou a ocorrência de 53 sequências quiméricas. No entanto, a título de confirmação de tal evento, cada uma destas 53 sequências foi alinhada com sequências depositadas no banco de dados não redundante do NCBI (*National Center for Biotechnology Information*), empregando-se o programa *blastn* (Altschul *et al.*, 1990). Nesta análise, todos os alinhamentos obtidos relacionavam-se a sequências de *Archaea* não cultiváveis. Desta forma, nenhuma das sequências foi considerada um híbrido de outros fragmentos de DNA e, portanto foram mantidas nas análises subsequentes.

5.6.2. Classificação taxonômica pelo *Ribosomal Database Project*

O banco de dados não redundante do NCBI é normalmente utilizado para obter informações sobre sequências de DNA. Entretanto, dado que a maior parte das espécies microbianas ainda não é cultivada, muitas vezes o alinhamento de sequências desconhecidas com aquelas depositadas no banco de dados não fornece informações detalhadas sobre o posicionamento taxonômico das amostras sob análise.

No entanto, no caso de sequências relativas aos genes que codificam os rRNAs, é preferível a utilização do banco de dados *Ribosomal Database Project (RDP)*, uma vez que este é específico para sequências dos genes que codificam os rRNAs 16S dos Domínios *Bacteria* e *Archaea* e o 18S de fungos. Uma das ferramentas disponíveis no RDP consiste no *Classifier* que, como o nome indica, busca classificar as sequências sob análise em posições taxonômicas hierárquicas, comparando-a às sequências depositadas. Desta forma, este banco apresenta as possíveis classificações em cada nível taxonômico para cada sequência submetida pelo usuário.

As 262 sequências com boa qualidade e tamanho superior a 400 pb foram submetidas à análise utilizando o algoritmo *Classifier* com limite de confiança (*confidence threshold*) de 90%. Dentre as sequências analisadas, 250 apresentaram níveis de confiança entre 90-100% com o domínio *Archaea* e apenas 12 não puderam ser classificadas de acordo com o limite estabelecido, sendo então denominadas “Domínio não identificado” (Figura 11). Das 250 sequências classificadas como

Archaea, 147 (47 CDa, 62 CDb, 71 MGa, 70 MGb) foram classificadas como pertencentes ao filo Crenarchaeota e as 103 restantes apresentaram nível de confiança inferior a 90% para classificação de filo e, portanto, foram nomeadas como “*Archaea* sem classificação”. Todas as sequências pertencentes ao filo Crenarchaeota são da classe Thermoprotei, a única classe descrita para este filo.

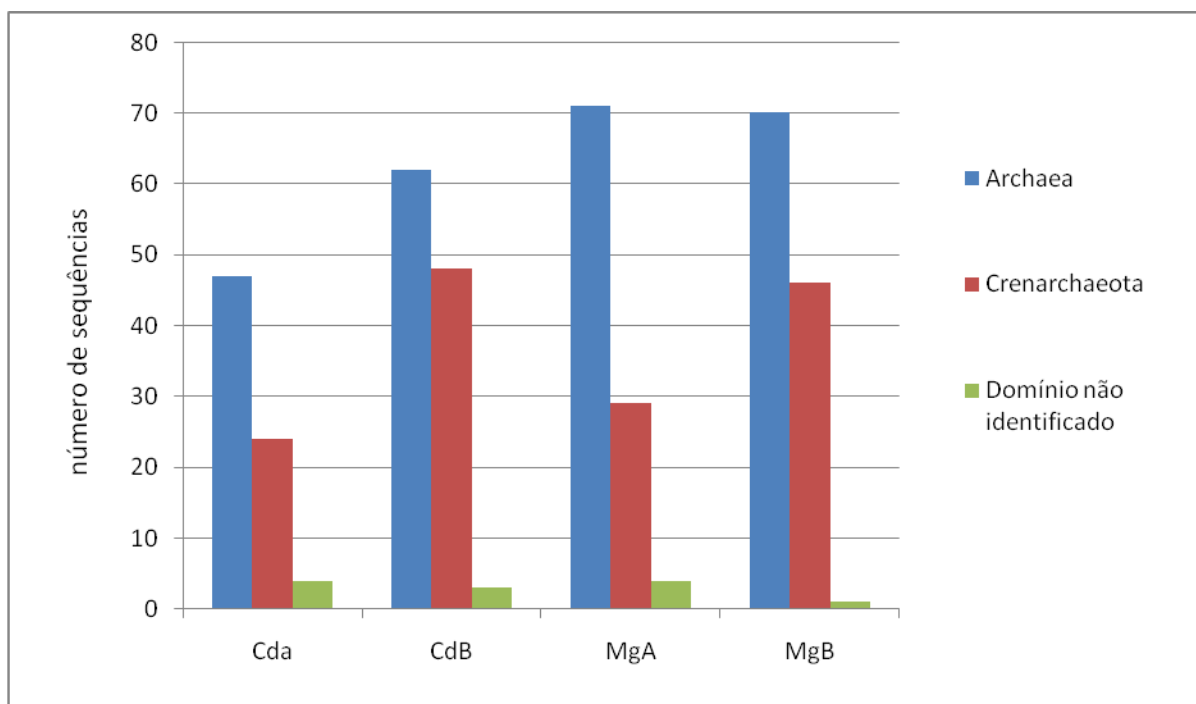


Figura 11. Histograma da classificação taxonômica das sequências de gene de 16S rRNA de cerrado denso (CDa, CDb) e mata de galeria (MGa, MGb). As sequências foram alinhadas com aquelas depositadas no banco de dados RDP, por meio do programa *Classifier* (Wang *et al.*, 2007).

As 12 sequências que não puderam ser classificadas não foram utilizadas nas análises estatísticas e na construção de árvores filogenéticas, por não ser possível afirmar que são sequências de *Archaea*. Desta forma, foram utilizadas: 47 sequências de CDa, 62 de CDb, 71 de MGa e 70 de MGb.

A classificação das sequências do Cerrado no filo Crenarchaeota está de acordo com a literatura. Os estudos de *Archaea* relatam que este filo é o predominante em solo (Bintrim *et al.*, 1997; Buckley *et al.*, 1998; Oschenreiter *et al.*, 2003). Em trabalhos

de biogeografia, Auguet (2009) e Bates (2010) mostraram ainda que o grupo I.1b do filo Crenarchaeota é o mais abundante em ambientes terrestres.

Embora a presença predominante de membros de Crenarchaeota em nossas amostras confirme os dados da literatura, tal resultado deve também ser avaliado em relação às limitações impostas pela metodologia obtida para a obtenção das sequências de DNA. Neste sentido, deve-se considerar os vieses introduzidos pela metodologia de PCR para o estudo de diversidade microbiana. Dentre estes, pode-se citar: variações na eficiência dos experimentos de PCR devido à presença de contaminantes na preparação de DNA total; a especificidade do iniciador em relação ao gene 16S rRNA alvo, resultando na amplificação diferencial de produtos; o número de cópias de genes de 16S rRNA por célula e suas variações intraespecíficas; formação de produtos artefatos quiméricos; e DNA contaminante não proveniente da amostra (Wintzingerode *et al.*, 1997).

Além disso, como a técnica de PCR é baseada em uma reação enzimática competitiva, DNAs mais abundantes na amostra terão maior probabilidade de serem amplificados. Portanto, organismos raros no ambiente podem não ter seu gene de 16S rRNA amplificado, mesmo que suas células tenham sido lisadas (Forney *et al.*, 2004). Além disso, também a técnica de clonagem impõe uma tendência de assimilação de clones de fragmentos presentes em maior proporção no sistema de ligação. Portanto, pode-se supor que nesse solo as euryarchaeotas sejam menos abundantes e por isso não foram detectadas nesta amostragem. Para garantir o acesso inclusive a organismos mais raros é importante seqüenciar um maior número de clones.

Entretanto, Bates (2011) relata que sequências de Euryarchaeota foram raramente encontradas nos solos amostrados em seu trabalho de biogeografia de *Archaea* em ambientes terrestres. Quando encontradas, compunham apenas 1,5% das sequências de *Archaea*. Uma baixa porcentagem de Euryarchaeota, cerca de 0,05% do total de archaeas, foi também detectada em amostras de solo de Cerrado, sendo mais abundantes em mata de galeria que cerrado denso (Araújo, 2011).

Mesmo possuindo sequências longas e com bons valores de phred, não foi possível obter classificação taxonômica inferior à classe, provavelmente devido a uma

falta de assinaturas genômicas para níveis taxonômicos mais específicos, e pelo fato das sequências de rDNA não possuírem sinal filogenético suficiente para resolver os nós mais internos talvez devido ao seu tamanho (Brochier-Armanet *et al.*, 2008). Vale ainda ressaltar que há baixo grau de concordância taxonômica entre as diversas sequências depositadas em bancos de dados, provavelmente em virtude da escassez de representantes cultivados (Auguet *et al.*, 2010). Um aspecto relevante, citado por Kan (2011) refere-se à ausência de sequências renomeadas como Thaumarchaeota nos bancos de dados do RDP e do greengenes, observando que todas as sequências desse filo permanecem classificadas como Crenarchaeota.

No momento, os bancos de dados dispõem de 122 genomas completos (Acesso em 23 de Janeiro de 2012. <<http://www.ebi.ac.uk/genomes/archaea.html>>), sendo 78 pertencentes ao filo Euryarchaeota (especialmente archaeas halófilas e metanogênicas); 39 ao filo Crenarchaeota; 1 ao filo Korarchaeota; 1 ainda depositado como Nanoarchaeota e 3 de Thaumarchaeota (*N. Maritimus* e *C. symbiosum*), sendo que *Candidatus Caldiarchaeum subterraneum*, originalmente proposto como pertencente a um novo filo - Aigarchaeota - é considerado membro do filo Thaumarchaeota.

Vale ainda ressaltar que outras abordagens, além da utilização de sequências de genes de rRNA, vêm sendo utilizadas no sentido de otimizar e refinar os achados filogenéticos. Uma estratégia que tem se mostrado bastante consistente, com resultados que permitem observar a distribuição dos organismos nas árvores filogenéticas em nós mais internos, consiste na utilização de conjuntos de proteínas ribossomais concatenadas (Brochier-Armanet *et al.*, 2008; Spang *et al.*, 2010; Brochier-Armanet *et al.*, 2011; Nunoura *et al.*, 2011).

5.6.3 Estimativa de riqueza e alfa-diversidade

Medir a diversidade de um sistema biológico implica em calcular a riqueza e a abundância relativa dos organismos em tal ambiente. A riqueza é a medida de diferentes espécies em determinada comunidade e a abundância é a homogeneidade da distribuição (*evenness*) dos indivíduos nessas espécies. Ao comparar dois

ambientes de mesma riqueza, aquele que possuir maior uniformidade de organismos divididos entre os táxons, logo maior abundância relativa, possui maior diversidade.

Quando procariotos são utilizados como modelo de estudos filogenéticos ou de diversidade, não se emprega o conceito de espécie, mas sim de unidade taxonômica operacional (OTUs), uma vez que, de acordo com Coenye (2005), para procariotos “espécies são populações e não entidades genômicamente coerentes”. Neste contexto, um grupo de sequências de DNA que apresentem 97% de identidade são consideradas como pertencentes a uma mesma OTU para nível de espécie.

Nesta mesma linha, o posicionamento hierárquico das OTUs nas diferentes categorias taxonômicas emprega valores de similaridade amplamente utilizados, propostos por Schloss e Handelsman, em 2004. De acordo com estes autores, sequências com 80% de similaridade (20% de diferença ou dissimilaridade) são consideradas como de um mesmo filo; com 90% de similaridade (10% de dissimilaridade) como de uma mesma classe/família; 95% de similaridade (5% de dissimilaridade) como de um mesmo gênero e 97% de similaridade (3% de dissimilaridade) como de uma mesma espécie.

O valor de 97% de identidade entre duas sequências de gene de 16S rRNA foi definido como padrão para a classificação de dois organismos em uma mesma espécie por Stackebrand e Goebel (1994). Entretanto, outros grupos sugerem que o valor de 99% possa ser mais acurado, uma vez que muitas espécies reconhecidas apresentam apenas 1% de dissimilaridade entre suas sequências do gene de 16S rRNA (Fox *et al.*, 1992; Yassin *et al.*, 2003).

Para obter o agrupamento das sequências segundo os níveis presuntivos de taxonomia primeiramente é necessário realizar o alinhamento múltiplo das sequências. Este foi realizado com o algoritmo NAST (*Near Alignment Space Termination*) (DeSantis *et al.*, 2006), programado especificamente para sequências de genes que codificam rRNA, considerando a estrutura secundária deste. As regiões contendo somente espaços gerados pelo alinhamento segundo a estrutura secundária foram filtradas com o programa Mothur (Schloss *et al.* 2009). As regiões hipervariáveis foram excluídas das sequências alinhadas, dado que mudanças múltiplas tendem a se

acumular nas regiões hiperváriaveis, podendo assim ocultar a genealogia da molécula ao simular identidades falsas e alterar a estimacão do número de eventos evolutivos (Stackebrandt e Goebel, 1994).

Esta etapa da bioinformática é limitante e crucial para as análises subsequentes, visto que a edição manual das sequências pode alterar os resultados estatísticos posteriores. Entretanto, White (2010) concluiu que o alinhamento realizado por meio das ferramentas NAST, MUSCLE e CLUSTALW não provoca diferenças significativas na estimativa de riqueza de OTUs.

Após o alinhamento, realizado pelo programa Mothur, o arquivo foi convertido para o formato phylip e, a partir deste, foi calculada a matriz de distâncias (método de *average neighbor*). Esta foi utilizada para o cálculo de agrupamento das sequências em táxons, segundo os valores presuntivos descritos anteriormente.

As sequências foram alinhadas de três formas: 1) as réplicas de cada fisionomia separadamente; 2) as réplicas de cada fisionomia em conjunto (CDa + CDb e MGa+ MGb) e 3) as sequências referentes às quatro amostras, em conjunto.

Isoladamente, foram obtidos os seguintes valores de OTUs para cada biblioteca, para o valor de 97% de similaridade: em CDa, as 42 sequências formaram 19 OTUs; CDb com 33 OTUs a partir de 63 sequências; MGa originou 19 OTUs a partir de suas 70 sequências e MGb 21 OTUs com suas 70 sequências (Tabela 6). A partir destes dados é possível concluir que mata de galeria é uma área com menor riqueza de espécies de *Archaea* do que cerrado denso, por possuir menor número de OTUs em relação ao número de sequências, ao contrário do que foi encontrado para *Bacteria* (Araújo, 2011).

Curiosamente, embora o número de sequências da amostra CDa fosse bastante inferior ao de CDb, a grande diferença de OTUS detectadas nestas duas amostras obtidas de cerrado denso, foi marcante. As réplicas de mata de galeria parecem mais homogêneas se considerarmos que ambas possuem quantidade similar de sequências e de OTUs. Na tabela 6, os singletons indicam as OTUs formadas por apenas uma sequência.

Tabela 6. Número de sequências, OTUs (3%) e singletons por comunidade.

	Sequências	OTUs	Singletons
CDa	42	19	10
CDb	62	33	21
MGa	70	19	9
MGb	70	21	8

Para confirmar quantas OTUs eram compartilhadas entre as réplicas foi realizado o alinhamento múltiplo de CDa com CDb e MGa com MGb.

O alinhamento dos CDs resultou em 37 OTUs totais, a partir de 104 sequências. Três sequências tiveram que ser excluídas de forma a manter o alinhamento múltiplo com tamanho maior que 350pb, sendo este valor considerado significativo para a inferência dos dados estatísticos. O diagrama de Venn gerado mostra que 11 OTUs são compartilhadas entre CDa e CDb (Figura 12). Apesar deste número parecer não ser significativo, dado que CDa apresenta 12 OTUs exclusivas e CDb 14, estas exclusivas são formadas majoritariamente por singletons. As 11 OTUs compartilhadas representam 74 sequências. Desta forma, os resultados obtidos sugerem que CDa e CDb são réplicas provavelmente sujeitas a alterações decorrentes de fatores ambientais, que resultam nas diferenças observadas nas amostras.

O mesmo resultado foi observado nas amostras de mata de galeria. Novamente três sequências tiveram que ser excluídas para manter o tamanho do alinhamento acima de 350pb. Das 140 sequências do alinhamento múltiplo de MGa e MGb, a riqueza total de mata de galeria foi de 19 OTUs, sendo 10 compartilhadas, 5 exclusivas de MGa e 4 de MGb (Figura 12). Novamente, estas OTUs exclusivas são, em sua maioria, compostas por singletons, de forma que as 10 OTUs compartilhadas representam 120 sequências. A similaridade entre as réplicas de mata de galeria é mais expressiva quando comparada ao número de OTUs compartilhadas entre os Cds. Provavelmente isto ocorre devido à menor riqueza de espécies de *Archaea* na fitofisionomia de mata de galeria.

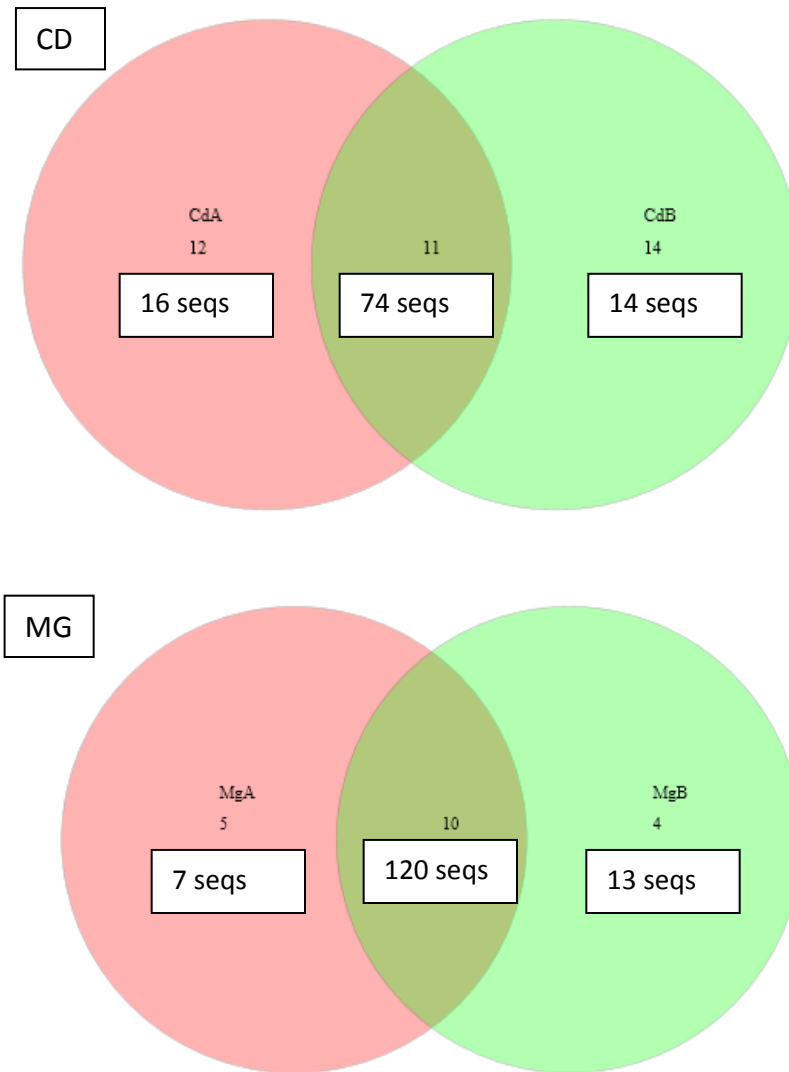


Figura 12. Diagrama de Venn representando o número de OTUs (3%) exclusivas em cada réplica e compartilhadas entre estas, para CD e MG. Os quadrados apresentam o número de seqüências que compõem cada grupo de OTUs.

A réplica biológica tem por objetivo confirmar a veracidade dos dados encontrados, mas como se pode observar, ao trabalhar com amostras compostas de solo, cada microhabitat sofre variações de acordo com fatores locais, como proximidade a raízes por exemplo. De forma que são verificadas seqüências que estão presentes em apenas uma das réplicas. De toda maneira, a maior parte das seqüências está presente na interseção das duas subamostras e, portanto, CDa-CDb e MGa-MGb serão consideradas réplicas, trabalhando-se assim com as comunidades CD (CDa+CDb) e MG (MGa+MGb).

Das 250 sequências totais caracterizadas como pertencentes ao domínio *Archaea*, nove foram descartadas na etapa de realização do alinhamento múltiplo de todas as bibliotecas, de forma a conservar o alinhamento com no mínimo 350 pares de bases. Desta forma, restaram então 42 sequências alinhadas de CDa, 60 de CDb, 71 de MGa e 68 de MGb.

Os mesmos cálculos foram utilizados para comparar as comunidades de cerrado denso e mata de galeria, considerando as réplicas como uma amostra única. Das 241 sequências do analisadas foram obtidas 59 OTUs, (Figura 13), sendo possível observar no diagrama de Venn que a comunidade de CD forma 28 OTUs exclusivas e MG 19 exclusivas, corroborando os gráficos anteriores que apresentam CD com maior riqueza. No entanto, CD e MG compartilham 12 OTUs. Estas OTUs compartilhadas são representadas por 121 sequências, metade do total de sequências alinhadas.

Este resultado sugere que há uma sobreposição de organismos nas duas fitofisionomias. Provavelmente tais organismos são generalistas, sofrendo pouca influência das características do solo de cada fitofisionomia, como os maiores valores encontrados de matéria orgânica, alumínio, fósforo e nitrogênio (Tabela 4) nos solos de mata de galeria.

A formação das comunidades de *Archaea* específicas em cerrado denso e mata de galeria, representadas pelas OTUs exclusivas de organismos especialistas, podem ser afetadas por um ou mais desses parâmetros cujos valores são diferentes para CD e MG. Da mesma forma, o conteúdo hídrico, também pode ser um fator responsável por esta diferença.

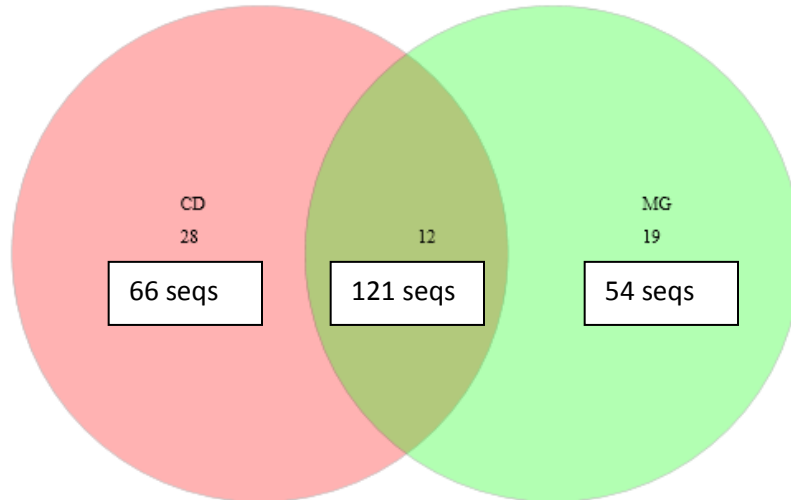
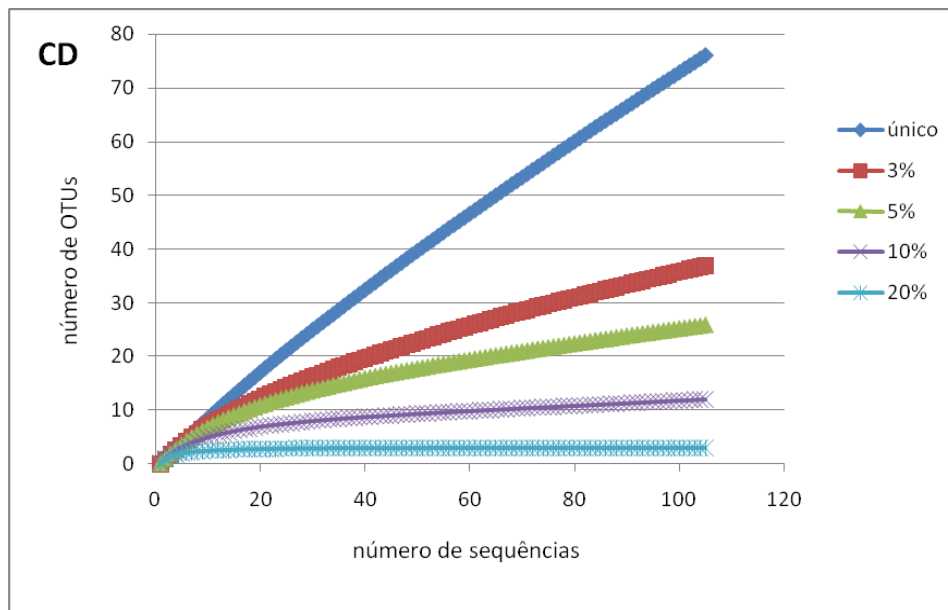


Figura 13. Diagrama de Venn representando o número de OTUs (3%) exclusivas a cada fitofisionomia e compartilhadas entre estas. CD: cerrado denso (CDa+CDb), MG (mata de galeria (MGa+MGb)). Os quadrados apresentam o número de sequências que compõem cada grupo de OTUs.

A Figura 14 apresenta a correlação entre número de OTUs e número absoluto de sequências foi estabelecida nas curvas de rarefação com o objetivo de se observar a cobertura da amostragem, para os níveis de espécie (3%), gênero (5%), classe (10%) e filo (20%).



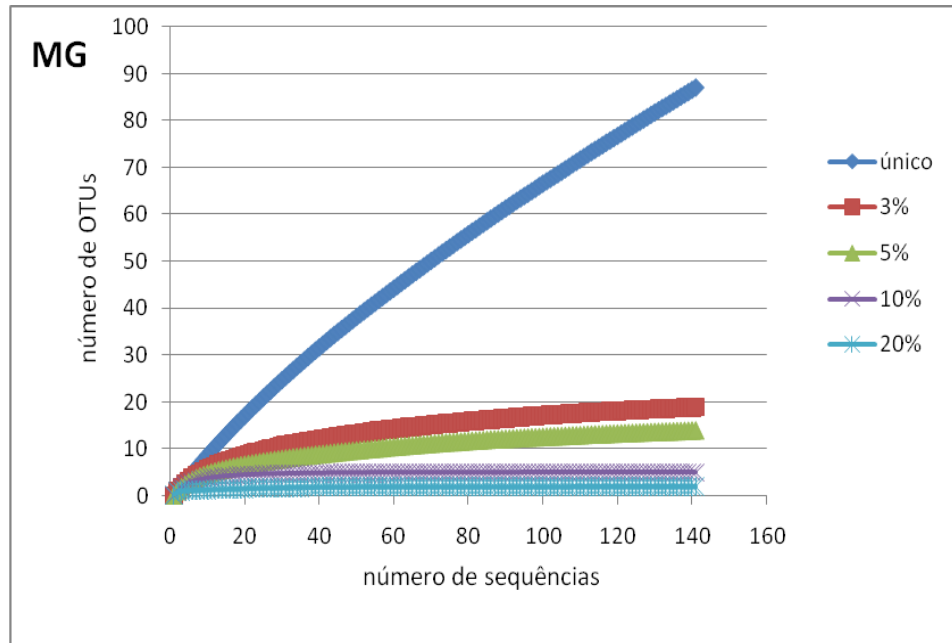


Figura 14. Curvas de rarefação com as porcentagens de dissimilaridade de 3%, 5%, 10%, 20%. Curva de CD e Curva de MG.

Enquanto o gráfico relativo à amostra MG (Figura 14) indica que a amostragem de 141 sequências parece ter sido suficiente para descrever a comunidade de *Archaea*, visto que para todos os níveis taxonômicos, a curva está próxima de atingir o platô, a análise da cobertura de cerrado denso (CD) indica haver a necessidade de aumentar a amostragem de clones, uma vez que para os níveis de espécie, gênero e classe a curva ainda não atingiu um platô.

A curva construída a partir de todas as bibliotecas para nível de espécie (3%) (Figura 15) corrobora as curvas obtidas com as amostras de cada fisionomia separadamente. CDa e CDb e MGa e MGb apresentam inclinações similares e novamente é possível observar que a amostra de mata de galeria está mais próxima de descrever toda a comunidade do que a amostra de cerrado denso.

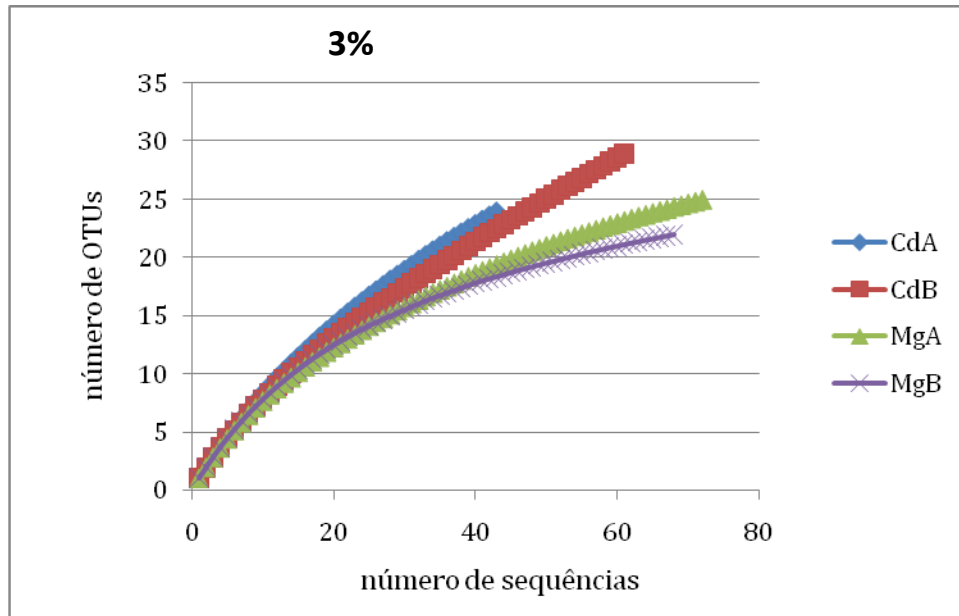


Figura 15. Curva de rarefação de todas as áreas com o limite de 3% de diferença.

Alguns estudos revelam que raramente bibliotecas de *Archaea* são tão diversas quanto as de *Bacteria*, sugerindo que as primeiras apresentam comunidades menos ricas, com maior proporção de organismos abundantes que de raros. Via de regra, a comparação de comunidades de *Bacteria* e *Archaea* de um mesmo ambiente revela que a quantidade de sequências necessárias para atingir a saturação da curva é menor em *Archaea* (Aller e Kemp, 2008; Robertson *et al.*, 2009).

Especificamente para o ambiente terrestre, é comum encontrar menor porcentagem de *Archaea* que *Bacteria*, compondo até 5% da microbiota do solo (Roesch *et al.*, 2007; Araújo, 2011; Bates *et al.*, 2011). Apesar de Kemnitz (2007) ter descrito um aumento proporcional de 12 a 38% de sequências do gene de 16S rRNA de *Archaea* com o aumento da profundidade do solo amostrado.

Archaeas apresentam maior diversidade em habitats de fendas hidrotermais e ecossistemas de água doce. No entanto, os organismos encontrados em solo são mais filogeneticamente agrupados, sendo a diversidade de *Archaea* resultante de um alto número de filotipos altamente relacionados, e não devido à presença de diferentes linhagens (Auguet *et al.*, 2010).

Hong (2009) defende ainda que mesmo que uma amostragem seja maior, o viés introduzido pela escolha dos iniciadores na PCR implica na descrição de apenas uma fração da real comunidade.

Trabalhos de ecologia microbiana também utilizam outros métodos para o cálculo de riqueza de comunidades de microorganismos, empregando a terminologia adotada em estudos de macroecologia, a saber: α -diversidade - aquela que descreve as espécies encontradas em determinado local; β -diversidade – que descreve as diferenças na composição de espécies entre dois sítios; e por fim, γ -diversidade – que descreve as diferenças de todos os sítios ao longo de uma paisagem. (Forney *et al.*, 2004)

A diversidade de um local é determinada pela riqueza de espécies, que consiste no número e uniformidade de espécies no local, ou seja, a abundância relativa de cada espécie (Forney *et al.*, 2004). Entretanto, em trabalhos baseados em genes de 16S rRNA somente a riqueza de filotipos da comunidade local é analisada, ou seja, fornecem resultados quantitativos apenas. Neste caso, não é possível inferir a diversidade, pelo fato da PCR ser uma técnica enzimática competitiva, na qual as fitas de DNA não são amplificadas de forma homogênea, isto é, há uma tendência de amplificação de moléculas de DNA presentes em maior concentração. Assim, não é possível afirmar a abundância de determinados grupos frente a outros.

Por outro lado, a diversidade pode ser analisada se mantida a ressalva de que está sendo realizada uma comparação do número de sequências em relação ao total amostrado naquele local. Por exemplo, no caso da descrição da abundância relativa principalmente de archaeas e bactérias, onde verifica-se que sequências de archaeas em solo representam somente 2% da população microbiana (Bates *et al.*, 2010).

A título de comparação entre as comunidades, os índices de diversidade de Shannon e de Simpson foram utilizados. O primeiro é utilizado para verificar a comunidade mais diversa. O segundo atribui valores de 0 a 1 para cada comunidade, sendo que valores mais próximos de 0 indicam maior diversidade.

A riqueza foi estimada também pelos índices não paramétricos de Ace e Chao1. Ambos estimam o número de OTUs esperadas. Estes valores podem ser comparados ao número de OTUs observadas (sobs) (Tabela 6).

Tabela 7 Índices de riqueza (Ace e Chao) e diversidade (simpson e shannon) para as duas fitofisionomias

		CD	MG
	Número de sequências	104	140
3%	Cobertura	76,92%	96,43%
	Sobs	37	19
	Ace	190,94	23,25
	Chao1	106,00	21,50
	Simpson	0,07	0,16
	Shannon	3,03	2,25
5%	Cobertura	86,54	97,14%
	Sobs	26	14
	Ace	71,61	17,04
	Chao1	56,33	16,00
	Simpson	0,08	0,20
	Shannon	2,72	1,89
10%	Cobertura	95,19%	100%
	Sobs	12	5
	Ace	22,00	5,00
	Chao1	29,28	5,00
	Simpson	0,21	0,27
	Shannon	1,87	1,41
20%	Cobertura	100%	100%
	Sobs	3	2
	Ace	3,00	2,00
	Chao1	3,00	2,00
	Simpson	0,49	0,92
	Shannon	0,84	0,18

Tal como observado para os valores de OTUs (Tabela 6 e Figura 12) e nas curvas de rarefação (Figura 14 e 15), os índices de riqueza e de diversidade, mostram que cerrado denso é não só mais rico como apresenta maior diversidade de organismos que mata de galeria.

Todos os resultados são congruentes no sentido de que a riqueza de espécies no cerrado denso e na mata de galeria foi praticamente coberta. Entretanto, métodos não

paramétricos são influenciados pelo tamanho da amostragem. Assim, caso esta seja muito pequena, a estimativa de riqueza também o será (Hughes *et al.*, 2001; Curtis *et al.*, 2006). Portanto, deve ser levada em conta a possibilidade da riqueza de *Archaea* em solo de Cerrado ter sido subestimada.

Os índices de Ace e Chao1 são calculados de forma diferente, portanto é usual que apresentem números diferentes de OTUs esperadas dentro de uma mesma amostra. Empregando-se valores de *cutoffs* mais baixos, como de 10 e 20%, os resultados tornam-se mais consistentes.

O índice de Shannon apresenta valores absolutos de diversidade e para todos os níveis taxonômicos, a amostra de cerrado denso apresentou maiores valores. Tais dados foram corroborados pelo índice de Simpson, no qual a amostra de cerrado denso apresentou valores mais próximos de 0, sugerindo ser um ambiente mais diverso que a mata de galeria. Mesmo apresentando menor diversidade, a amostra de mata de galeria também apresentou alta diversidade até o nível de classe (10% de dissimilaridade). Desta forma, considerando que mata de galeria apresenta menor riqueza, mas níveis de diversidade mais elevados, sua composição deve ser caracterizada por uma maior abundância relativa de organismos.

5.6.4. Análises estatísticas de beta-diversidade

A análise comparativa de diversidade entre as amostras é denominada *beta*-diversidade e pode ser estimada por meio de testes de hipótese alternativos aos baseados em OTUs, tais como *beta*-Libshuff (Schloss *et al.*, 2004) e Unifrac (Lozupone *et al.*, 2006).

O teste *beta*-Libshuff compara as comunidades par a par e verifica se uma é subamostra da outra e permite concluir qual a significância de cada comunidade ser única do local.

O valor de *p* corrigido pela correção de Bonferroni indica a significância entre as amostras. Se o resultado do valor *p* for inferior ao *p* corrigido, as comunidades apresentam diferença significativa.

Tabela 8. Teste de hipótese Libshuff entre as amostras de cerrado denso e mata de galeria.

Comparação	dCXYScore	value	Significância se =0,05/12=0,0041
CDa-CDb	0.0002015	0.6815	Não
CDb-CDa	0.0001000	0.7717	Não
CDa-MGa	0.0024346	<0.0001	Significativo
MGa-CDa	0.0017213	0.0107	Não
CDa-MGb	0.0028176	<0.0001	Significativo
MGb-CDa	0.0038394	0.0001	Significativo
CDb-MGa	0.0037450	<0.0001	Significativo
MGa-CDb	0.0042325	<0.0001	Significativo
CDb-MGb	0.0042753	<0.0001	Significativo
MGb-CDb	0.0065050	<0.0001	Significativo
MGa-MGb	0.0001553	0.4339	Não
MGb-MGa	0.0002927	0.1293	Não

As réplicas não apresentam diferença significativa entre si, como esperado. A baixa sobreposição de sequências entre as bibliotecas das áreas de cerrado denso e mata de galeria permitiu concluir se tratarem de amostras diferentes, com comunidades únicas. A comparação do par reverso permitiu identificar qual comunidade é subamostra da outra, como no caso de MGa-CDa, em que CDa é subamostra de MGa, provavelmente devido à pequena amostragem de CDa (Tabela 8).

O conceito de Baas-Becking de que “tudo está em todo lugar, o ambiente seleciona” (“*everything is everywhere; the environment selects*”), mostrado inicialmente em seleção de placa (Wit e Bouvier, 2006), é aceito hoje em dia na microbiologia ambiental, sendo confirmado por estudos de sequenciamento de genes marcadores filogenéticos. Para a comunidade bacteriana do solo de cerrado foi observado agrupamento de sequências conforme a fitofisionomia analisada, ficando significativamente separados os grupos de mata de galeria e campo sujo, do cerrado

denso e cerrado *sensu strictu* (Araújo, 2011). Auguet (2010) mostrou para *Archaea* haver também um agrupamento de determinados grupos conforme o habitat.

O programa Unifrac (Lozupone *et al.*, 2006) foi utilizado para avaliar a diferença da estrutura entre as comunidades das duas fitofisionomias. Uma árvore filogenética no formato Newick é utilizada como *input* para esse programa, no qual é possível observar em quais linhagens estão concentradas as diferenças entre as comunidades. O gráfico de Análise das Principais Coordenadas é a forma de visualização de tais diferenças a partir das distâncias entre as sequências.

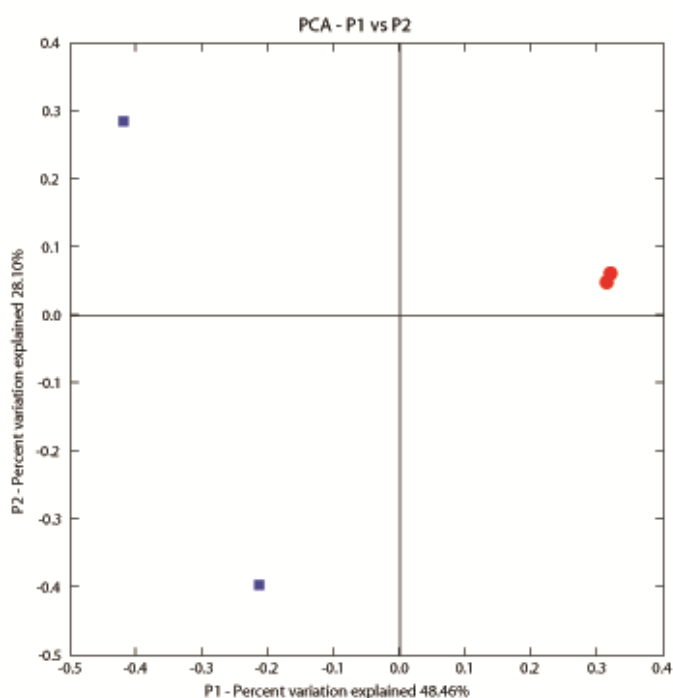


Figura 16. Gráfico de Análise das Principais Coordenadas (PCA) construído no programa Unifrac (Lozupone *et al.*, 2006) para demonstrar agrupamento em quadrantes dos ambientes amostrados. Quadrados azuis indicam as réplicas de cerrado denso e círculos vermelhos indicam as réplicas de mata de galeria.

No gráfico de Análise de Principais Coordenadas (PCA) foi possível observar diferença entre mata de galeria e cerrado denso para sequências de *Archaea*, separadas em quadrantes diferentes (Figura 16). Todavia, CDa e CDb, réplicas biológicas da fitofisionomia de cerrado denso não foram agrupadas em um mesmo quadrante, provavelmente devido à maior semelhança entre as sequências de MGa e

MGb, talvez relacionado à menor riqueza de espécies deste grupo. Este resultado está coerente com aqueles obtidos nos Diagramas de Venn (Figura 12), em que se observa maior similaridade entre as réplicas de MG do que entre as de CD.

Como se observa nas análises de beta-diversidade (Tabela 8 e Figura 16), as comunidades das duas fitofisionomias são significativamente diferentes, independentemente da alta porcentagem de compartilhamento de OTUs entre as comunidades de CD e MG apresentado na Figura 13.

No programa Unifrac, foi utilizado o *Lineage Specific Analysis*, com uma distância de 0,1, a fim de identificar as OTUs que contribuíram para a diferença entre as fitofisionomias. As sequências que promoveram essa diferença maior entre as réplicas de cerrado denso comparativamente às de mata de galeria foram alinhadas no programa *blastn* do banco de dados do NCBI.

Apesar do alinhamento das sequências das quatro áreas com o banco de dados não redundante do NCBI ter mostrado heterogeneidade entre os *hits*, as sequências dos grupos que promoveram a diferença entre as amostras nos quadrantes do PCA foram todas da fitofisionomia de cerrado denso, de ambas amostras da réplica, e apresentaram todas os mesmos melhores *hits*: uma crenarchaeota proveniente de intestino de cupim (AF293587), uma archaea não classificada de solo de um vulcão (DQ791525) e uma crenarchaeota descrita de um solo sob cultivo de arroz (AB380040). Estes foram diferentes de todos os *hits* obtidos com o restante das sequências, indicando que tal grupo de sequências de CD difere do restante da mesma fitofisionomia e daquelas de mata de galeria.

5.6.5. Árvores filogenéticas

A construção de árvores filogenéticas é uma metodologia baseada, essencialmente, no alinhamento múltiplo das sequências sob análise com aquelas disponíveis em banco de dados.

Um representante de cada OTU em nível de espécie (3%) foi escolhido para a construção da árvore filogenética de cada fitofisionomia (amostras CD e MG).

Sequências de isolados de *Archaea* dos filos Euryarchaeota e Crenarchaeota, assim como dos filos propostos Thaumarchaeota, Korarchaeota e Aigarchaeota foram alinhados às sequências representantes de OTUs para construção da árvore, no sentido de permitir a avaliação da qualidade do alinhamento e da construção da árvore, como também para obter uma melhor estimativa das possíveis classificações dos clones obtidos no solo de cerrado (Figura 17).

Para a construção da árvore foram escolhidos genomas de organismos atualmente considerados do filo Thaumarchaeota, ou sequências parciais do gene de rRNA 16S de outros organismos não cultivados, amplamente descritos em outros trabalhos. Isolados conhecidos de cada filo, com genoma sequenciado, foram também incluídos na árvore. Do filo Euryarchaeota: *Haloferax volcanii*, *Methanocella paludicola*, *Methanocalculus halotolerans*. Do filo Crenarchaeota, as archaeas hipertermófilas: *Sulfolobus solfataricus*, *Ignicoccus hospitalis*, *Thermoproteus neutrophilus* e *Pyrobaculum islandicum*. Do filo Korarchaeota o único representante com genoma sequenciado, *Candidatus Korarchaeum cryptofilum*.

O método *Neighbor-Joining* foi o escolhido para a construção da árvore por ser o mais simples e o mais utilizado na construção de árvores a partir de sequências do gene 16S rRNA. Este método apresenta uma aproximação à mínima evolução ao aglomerar as OTUs a partir da matriz de distância gerada, produzindo uma árvore com os menores comprimentos (Saitou e Nei, 1987).

Embora o método de máxima verossimilhança apresente a melhor representação evolutiva das sequências analisadas, requer maior esforço computacional e baseia-se em modelos evolutivos complexos, desnecessários ao estudo de riqueza de espécies por meio de sequência de gene de 16S rRNA. No entanto, ambos os métodos foram testados para construção da árvore e as sequências de cerrado obtidas neste trabalho formaram os mesmos clados.

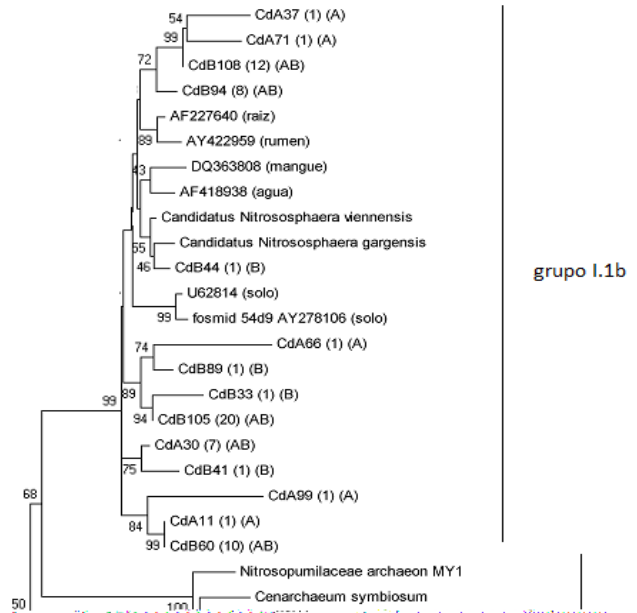
Conforme mencionado anteriormente, os isolados do grupo I de Crenarchaeota são classificados por alguns autores como pertencentes ao filo Thaumarchaeota. Esta nomenclatura será utilizada também para fins de classificação das sequências nesta dissertação.

Neste grupo I, as sequências de *Cenarchaeum symbiosum* (Preston *et al.*, 1996); Candidatus *Nitrosopumilus maritimus* (Konneke *et al.*, 2005); *Giganthauma karukareense*, *Giganthauma insulaporcus* (Muller *et al.*, 2010); e da recente espécie proposta Candidatus *Nitrosoarchaeum koreensis* (*Nitrosopumilaceae* archaeon MY1, figura 17) (Jung *et al.*, 2011) são pertencentes ao grupo I.1a. Todas são isoladas de ambiente aquático, exceto a última, obtida de solo agriculturável.

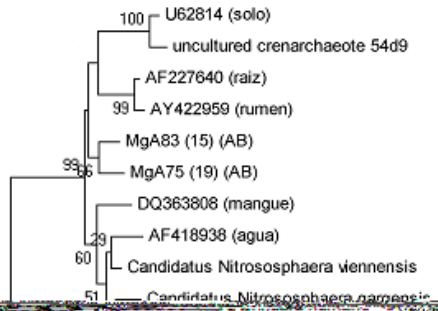
Outros isolados, representantes do grupo I.1b de Thaumarchaeota foram também utilizados. São eles *Nitrososphaera gargensis* (Hatzenpichler *et al.*, 2008) e *Nitrososphaera viennensis* (Tournai *et al.*, 2010), além de um isolado que se apresenta em um ramo irmão ao grupo I.1b, Candidatus *Nitrosocaldus yellowstonii* (de la Torre *et al.*, 2008).

A maioria destes isolados apresenta o gene *amoA* e para alguns já foi demonstrada a capacidade de oxidação de amônia. Somente para as espécies de *Giganthauma* não foi encontrado tal potencial de nitrificação.

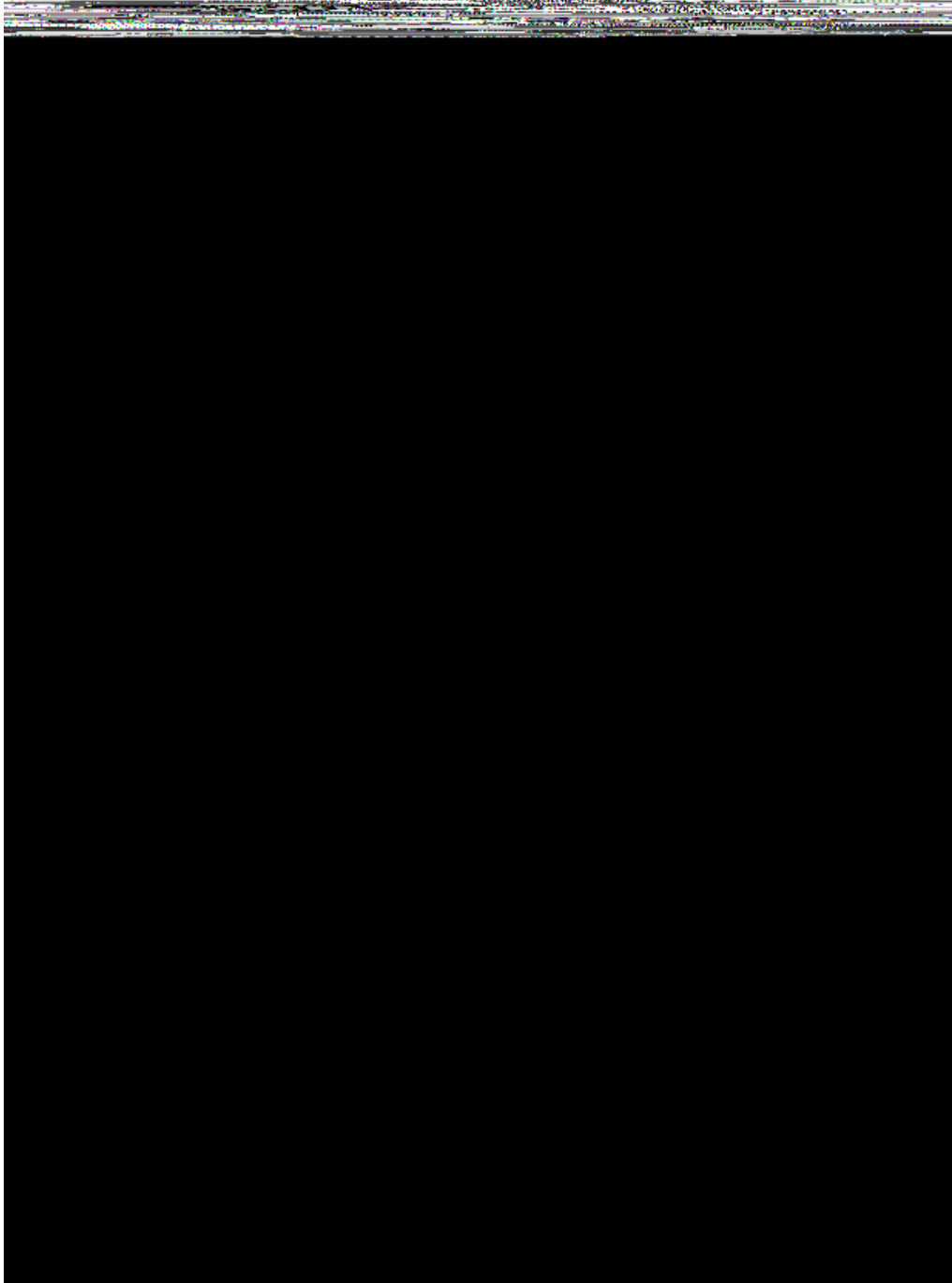
CD



MG



grupo I.1b



CD+MG

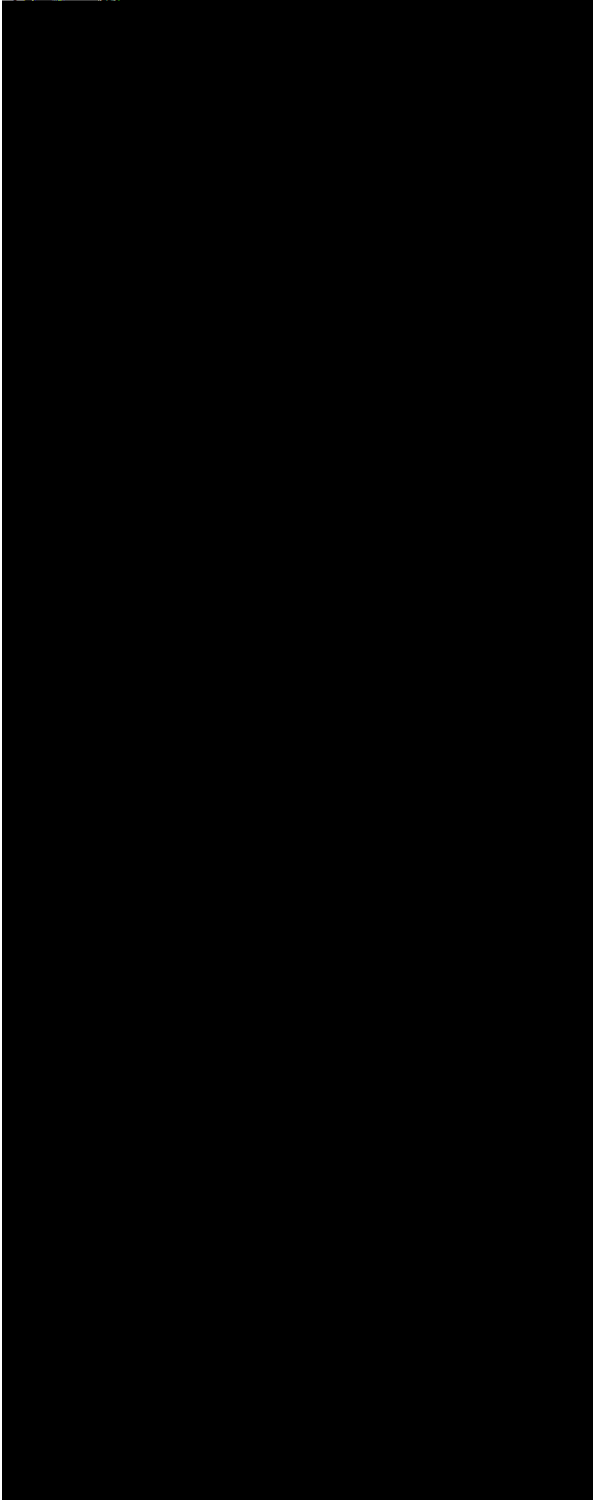


Figura 17. Árvores filogenéticas construídas a partir dos representantes de cada OTU (3%) das amostras CD e MG e CD+MG, após alinhamento com sequências de diferentes isolados de *Archaea*. A árvore foi construída pelo programa MEGA5 (Tamura *et al.*, 2011) pelo método de *Neighbor Joining* (Saitou e Nei, 1987), com modelo de Jukes-Cantor (Jukes e Cantor, 1969) e teste de 1000 *bootstrap* (Felsenstein, 1985). Valores de *bootstrap* menores que 50 foram excluídos da árvore. A escala indica a distância da similaridade dos ramos. Uma sequência de *Acidobacteria* não cultivada do grupo gp1 foi utilizada como grupo externo. E: *Euryarchaeota*; C: *Crenarchaeota*; K: *Korarchaeota*; A: *Aigarchaeota*; I.1a, I.1b e I.1c: *Thaumarchaeota*. Os valores entre parentesis em cada clone indicam o número de sequências agrupadas na OTU daquele representante, e a indicação se OTU é representada por sequências exclusivamente da réplica A ou B ou se de ambas.

A análise das árvores filogenéticas revela que todas as sequências recuperadas das amostras de cerrado denso e mata de galeria encontram-se posicionadas no grupo anteriormente classificado como grupo I de crenarchaeotas mesófilas, atualmente considerado por vários autores como o novo filo Thaumarchaeota (Brochier-Armanet *et al.*, 2011; Nunoura *et al.*, 2011).

As relações filogenéticas observadas na Figura 17 da árvore filogenética de solos de cerrado denso e mata de galeria podem ser comparadas aquelas apresentadas por Pester (2011) na Figura 2.

As 113 sequências utilizadas para a construção da árvore da fitofisionomia de cerrado denso, originaram 36 OTUs, sendo 14 posicionadas no grupo 1.1b e 22 em 1.1c. No entanto, apesar de parecer desta forma que a maioria das sequências estão afiliadas ao grupo I.1c de Thaumarchaeota, aquelas 14 OTUs do grupo I.1b são formadas pela maioria das sequências: 66. As 47 sequências restantes formam o grupo I.1c.

No caso da amostra de mata galeria, foram obtidas 140 sequências, que geraram 19 OTUs assim distribuídas: 4 OTUs (6 sequências) em I.1a, 3 OTUs (39 sequências) em I.1b e 12 OTUs (95 sequências) em I.1c. Nesta fitofisionomia há uma clara diferença, com maior predominância de sequências no grupo I.1c. Este resultado está consistente com a literatura, que relata sequências deste grupo isoladas de formações florestais, como é considerada a mata de galeria (Jurgens *et al.*, 1997).

Apesar do número de OTUs esperadas (Tabela 7) indicar que mata de galeria apresenta menor riqueza e diversidade que cerrado denso para todos os níveis presuntivos mesmo de classe e filo, na árvore (Figura 17) pode-se observar formação das sequências de mata de galeria nos três grupos de Thaumarchaeota, contra apenas dois grupos em cerrado denso.

O primeiro clado de CD, constituído por sequências tanto da réplica A quanto da B, está afiliado com sequências isoladas de rizosfera (AF227640) e de rúmen (AY422959). Os isolados de *Candidatus Nitrososphaera gargensis* e *Nitrososphaera viennensis* formam um clado com uma sequência de mangue (DQ363808) e de sedimento de água doce (AF418938) e com a OTU CDb44.

As sequências de solo utilizadas para construção das árvores foram descritas por trabalhos de referência. Jurgens (1997) - X96688 e Kemnitz (2007) - AM291987 são artigos importantes na descrição de crenarchaeotas mesófilas de solos ácidos. Oschenreiter (2003) - AY278095 e Bintrim (1997) - U62814 são os artigos mais citados sobre a predominância do filo Crenarchaeota, principalmente do grupo I.1b, em solos.

Na árvore de CD, observa-se a forte afiliação de sequências do cerrado denso com as sequências de solo descritas por Jurgens (1997) e Oschenreiter (2003) no grupo I.1c. Outro clado está fortemente relacionado à sequência descrita por Kemnitz (2007). Todos estes corroboram a qualidade da construção da árvore por terem descrito ambientes com alta porcentagem de sequências do grupo I.1c de Thaumarchaeota.

Na árvore de MG é possível observar claramente um menor número de OTUs, com grande número de sequências. A OTU mais representativa, com 47 sequências se encontra no grupo I.1c fortemente afiliada às sequências de solo descritas por Oschenreiter (2003) e Jurgens (1997) e também a uma sequência amostrada de um lixiviado de mina de urânio (AJ535136).

Há a formação de um clado irmão de uma pequena porcentagem das sequências de mata galeria com o grupo I.1a, representado este por sequências de ambiente aquático.

Com o intuito de verificar a formação de clados entre as sequências de cerrado denso e mata de galeria foi construída também uma árvore filogenética com todas as sequências representantes para nível de 3% das comunidades de cerrado denso e mata de galeria. Verifica-se novamente a formação de clados unicamente de sequências do Cerrado e muitos destes constituídos por sequências de CD e MG.

No grupo I.1b, a maioria das sequências de mata de galeria estão próximas filogeneticamente das sequências descritas a partir de rizosfera (AF227640) e de rúmen (AY422959) e de solo descrito por Bintrim (1997) (U62814). Mas é possível observar uma OTU representada por 5 sequências que forma um ramo irmão à sequência do isolado de *Candidatus Nitrososphaera gargensis* descrito por Hatzenpichler (2008).

É possível observar em ambas as árvores a formação de clados exclusivamente compostos por sequências de Cerrado. Não é possível afirmar que sejam sequências novas sem afiliação ainda descrita, pois seria necessário utilizar todas as sequências descritas para construção da árvore, no entanto é um indicativo de sugere a existência de organismos ainda não descritos no solo do Cerrado.

Os solos de cerrado denso e de mata de galeria são levemente ácidos, apresentando pH 5,0. Portanto seria esperado encontrar um maior porcentagem de sequências afiliadas ao grupo I.1c, como encontrado por Kemnitz (2007). Enquanto nas amostras de cerrado denso, pode-se verificar uma distribuição homogênea das sequências nos grupos I.1b (de solos) e I.1c (de solos ácidos), o mesmo não ocorreu em relação às amostras de mata de galeria. Neste caso, houve clara predominância de sequências afiliadas ao grupo I.1c, que foi definido como um grupo predominante de solos ácidos.

A composição deste grupo é diretamente afetada pelo pH do solo, sendo que sua abundância aumenta com a diminuição do pH (Lehtovirta *et al.*, 2009) e pode chegar a 85% do total de sequências de *Archaea* em solo ácido de floresta (Kemnitz *et al.*, 2007).

Devido ao maior teor de água em solos de mata de galeria, não é surpreendente encontrar sequências do grupo I.1a apenas nessas amostras. Apesar de,

recentemente, ter sido descrito um novo candidato ao filo *Thaumarchaeota* isolado de solo de agricultura e pertencente ao grupo I.1a (Jung, 2011).

A quantificação da umidade relativa de solos das mesmas áreas (cerrado denso e mata de galeria), coletados para outro trabalho, revelaram que amostras de mata de galeria contêm 37% de conteúdo de água, contra 18% presente no solo de cerrado denso (dados cedidos gentilmente pela doutoranda Maria Regina da Silva Sartori do Laboratório de Ecologia da Universidade de Brasília).

Outros parâmetros físico-químicos, como aqueles apresentados na Tabela 4, podem também influenciar a diferenciação das comunidades de *Archaea* em mata de galeria e cerrado denso, como por exemplo as quantidades e matéria orgânica, fósforo e mesmo a concentração de nitrogênio.

Como as sequências de cerrado denso e mata de galeria encontradas neste trabalho formaram grupo monofilético, com *bootstrap* maior do que 70, com organismos oxidantes de amônia do filo Thaumarchaeota, é possível que algumas destas sejam de organismos capazes de metabolizar amônia.

Curiosamente, o posicionamento das sequências de isolados dos outros filios difere entre as duas árvores. Tal fato poderia ser explicado pelos baixos valores de *bootstrap* para os ramos que formam tais filios. Atualmente, considera-se que inferências mais robustas sobre o posicionamento filogenético de determinado grupo de sequências requerem a utilização de sequências concatenadas de SSU e LSU e/ou proteínas ribossomais (Spang *et al.*, 2010; Nunoura *et al.*, 2011).

Por outro lado, é importante ressaltar que na afiliação dos clados das sequências de Cerrado com sequências descritas para os grupos I.1a, I.1b e I.1c, os valores de *bootstrap* foram maiores que 70, o que pode ser considerado significativo para fins de classificação taxonômica.

6. Considerações finais

Este trabalho apresenta a primeira descrição da riqueza de *Archaea* no solo do bioma Cerrado. No entanto, uma vez que este bioma é caracterizado por estações bem definidas com diferentes níveis pluviométricos, provavelmente a composição da

comunidade de *Archaea* na estação de transição seca-chuva – quando as amostras foram coletadas – não é a mesma presente nas outras estações (seca, chuvosa, transição chuva-seca).

Devido ao grande número de trabalhos descrevendo a predominância de membros do filo Crenarchaeota em solos, é provável que este seja também o mais encontrado em outras épocas do ano, principalmente o grupo I.1b, o mais ubíquo em ambientes terrestres. Considerando-se também que o conteúdo de água influencia no pH do solo, é possível que em outras estações climáticas haja uma maior proporção de organismos do grupo I.1c. Por outro lado, organismos do filo Euryarchaeota, tais como archaeas metanogênicas podem estar em menor número em épocas secas, tornando-se mais representativas quando o solo apresenta maior conteúdo de água (Angel *et al.*, 2011).

Os resultados obtidos neste trabalho não permitem inferências acerca do metabolismo das archaeas no solo do cerrado. É possível, no entanto, especular o potencial nitrificador dos organismos encontrados, dado que suas sequências agruparam-se, nas árvores filogenéticas, com organismos do filo Thaumarchaeota, reconhecidamente participantes do ciclo do nitrogênio.

A distribuição terrestre de *Archaea*, em uma escala global, apresenta grande variabilidade quanto à estrutura da comunidade e à sua abundância relativa, o que enfatiza a necessidade de amostrar o maior número de solos antes de apresentar-se conclusões gerais sobre a ecologia e riqueza de *Archaea* (Bates *et al.*, 2010).

Este é o primeiro trabalho de avaliação da composição da comunidade de *Archaea* no solo de Cerrado. Entretanto, um maior entendimento sobre a real diversidade e endemidade dos filotipos encontrados requer a utilização de outros marcadores filogenéticos em concomitância com o gene de 16S rRNA, ou mesmo, mais de um conjunto de iniciadores para este gene.

Dada a baixa resolução do gene de 16S rRNA, estudos futuros sobre comunidades de *Archaea* poderão ser realizados empregando sequências de proteínas ribossomais, a fim de permitir uma possível elucidação da classificação dos organismos presentes no solo do cerrado em nível taxonômico mais específico.

7. Conclusões e perspectivas

- Pode-se concluir que crenarchaeotas mesófilas, possivelmente também denominadas Thaumarchaeota, são dominantes nos solos de cerrado denso e mata de galeria do Cerrado da Reserva Ecológica do IBGE.

- As fitofisionomias de cerrado denso e mata de galeria apresentam composições distintas de *Archaea*, com a predominância de membros do grupo I.1c em Mg e de I.1b em Cd.

- A presença de organismos do grupo I.1a, de ambientes aquáticos, só foi detectada em mata de galeria.

- Apesar da significativa diferença verificada com os testes estatísticos de diversidade, Cd e Mg compartilham 121 sequências divididas em 12 OTUs (3%).

- A curva de rarefação de mata de galeria começa a atingir o platô para nível de espécie com 70 sequências. No entanto, será necessário o sequenciamento de um maior número de clones a fim de obtermos uma maior cobertura em ambas as fitofisionomias, especialmente na de cerrado denso.

- Este trabalho será ampliado para as outras fitofisionomias encontradas no bioma Cerrado.

- Nosso grupo já iniciou trabalhos com os iniciadores para o gene *amoA*, envolvido no processo de oxidação de amônia, dado que este é um metabolismo comumente encontrado em organismos do filo proposto *Thaumarchaeota*.

- Ensaios de PCR em tempo real poderão ser utilizados para averiguar a abundância do filo *Thaumarchaeota* no solo, a qual será correlacionada com a expressão dos genes de oxidação de amônia de forma a estimar a importância de *Archaea* de solo de cerrado no ciclo do nitrogênio.

- O cultivo de *Archaea* do solo de Cerrado poderá ser realizado com o intuito de descrever novos organismos desse domínio e, se possível, mostrar a capacidade de nitrificação ou outros metabolismos para esses.

8. Referências Bibliográficas

- ADAIR, K.L.; SCHWARTZ, E. (2008) Evidence that ammonia-oxidizing archaea are more abundant than ammonia-oxidizing bacteria in semiarid soils of northern Arizona, USA. **Microbiology and Ecology**, 56: 420-426.
- ALBERS, S-V e MEYER, B.H. (2011) The archaeal cell envelope. **Nature Reviews Microbiology**, 9: 414-426.
- ALLER, J.Y.; KEMP, P.F. (2008) Are Archaea inherently less diverse than Bacteria in the same environments? **FEMS Microbiology Ecology**, 65: 74-87.
- ALTSCHUL, S.F.; GISH, W.; MILLER, W.; MYERS, E.W.; LIPMAN, D.J. (1990) Basic local alignment search tool. **Journal of Molecular Biology**, 215: 403-410.
- AMANN, R.; LUDWIG, W.; SCHLEIFER, K. (1995). Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. **Microbiological Reviews**, 59(1): 143-169.
- ANGEL, R.; MATTHIES, D.; CONRAD, R. (2011) Activation of Methanogenesis in Arid Biological Soil Crusts Despite the Presence of Oxygen. **PLoS One**, 6(5): 1-8.
- ARAÚJO, J.F. 2011. **Diversidade bacteriana do solo em diferentes fitofisionomias do bioma Cerrado e perspectivas biotecnológicas**. 154f. Tese (Doutorado em Ciências Genômicas e Biotecnologia) – Universidade Católica de Brasília, Brasília.
- AUGUET, J.C.; BARBERAN, A.; CASAMAYOR, E.O. (2009) Global ecological patterns in uncultured archaea. **The ISME Journal**, 4: 182-190.
- BAKER, G.C.; SMITH, J.J.; COWAN, D.A. (2003) Review and re-analysis of domain-specific 16S primers. **Journal of Microbiological Methods**, 55: 541-555.
- BARNS, S.M.; DELWICHE, C.F.; PALMER, J.D.; PACE, N.R. (1996) Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. **Proceedings of the National Academy of Sciences of USA**, 93: 9188-9193
- BATES, S.T.; BERG-LYONS, D.; CAPORASO, J.G.; WALTERS, W.A.; KNIGHT, R.; FIERER, N. (2010) Examining the global distribution of dominant archaeal populations in soil. **The ISME Journal**, 5: 908-917.

- BELAY, N.; JOHNSON, R.; RAJAGOPAL, B.S.; MACARIO, E.C.; DANIELS, L. (1988) Methanogenic Bacteria from Human Dental Plaque. **Applied and Environmental Microbiology**, 54(2): 600-603.
- BINTRIM, S.B.; DONOHUE, T.J.; HANDELSMAN J.; ROBERTS, G.P. (1997) Molecular phylogeny of archaea from soil. **Proceedings of the National Academy of Sciences of USA**, 94: 277-282.
- BRESOLIN, J.D.; BUSTAMANTE, M.M.C.; KRUGER, R.H.; SILVA, M.R.S.S.; PEREZ, K.S. (2010) Structure and composition of bacterial and fungal community in soil under soybean monoculture in the brazilian cerrado. **Brazilian Journal of Microbiology**, 41: 391-403.
- BRIÉE, C.; MOREIRA, D.; LÓPEZ-GARCÍA, P. (2007) Archaeal and bacterial community composition of sediment and plankton from a suboxic freshwater pond. **Research in Microbiology**, 158: 213-227.
- BROCHIER, C.; GRIBALDO, S.; ZIVANOVIC, Y.; CONFALONIERI, F.; FORTERRE, P. (2005) Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? **Genome Biology**, 6: R42.
- BROCHIER-ARMANET, C.; BOUSSAU, B.; GRIBALDO, S.; FORTERRE, P. (2008) Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. **Nature Reviews Microbiology**, 6: 245-252.
- BUCKLEY, D.H.; GRABER, J.R.; SCHMIDT, T.M. (1998) Phylogenetic analysis of nonthermophilic members of the kingdom Crenarchaeota and their diversity and abundance in soils. **Applied and Environmental Microbiology**, 64: 4333-4339.
- CASANUEVA, A. ; GALADA, N.; BAKER, G.C.; GRANT, W.D.; HEAPHY, S.; JONES, B.; YANHE, M.; VENTOSA, A.; BLAMEY, J.; COWAN, D.A. (2008) Nanoarchaeal 16S rRNA gene sequences are widely dispersed in hyperthermophilic and mesophilic halophilic environments. **Extremophiles**, 12: 651-656.
- CASTRO, A.P. de; QUIRINO, B.F.; PAPPAS, G.Jr.; KUROKAWA, A.S.; NETO, E.L.; KRÜGER, R.H. (2008) Diversity of soil fungal communities of Cerrado and its closest surrounding agriculture fields. **Archives of Microbiology**, 190: 129-139.

- COENYE, T.; GEVERS, D.; VAN DE PEER, Y.; VANDAMME, P.; SWINGS, J. (2005) Towards a prokaryotic genomic taxonomy. **FEMS Microbiology Reviews**, 29: 147-167.
- CURTIS, T.P.; HEAD, I.M.; LUNN, M.; WOODCOCK, S.; SCHLOSS, P.D.; SLOAN, W.T. (2006) What is the extent of prokaryotic diversity? **Philosophical Transactions of The Royal Society**, 361: 2023-2037.
- DANIEL, R. The metagenomics of soil. (2005) **Nature Reviews Microbiology**, 3: 470-478.
- DAGERT, M. e EHRLICH, S.D. (1979) Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. **Gene**, 6: 23-28.
- DE LA TORRE, J.R.; WALKER, C.B.; INGALLS, A.E.; KONNEKE, M.; STAHL, D.A. (2008) Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. **Environmental Microbiology**, 10 (1): 810-818.
- DELONG, E. (1992) Archaea in coastal marine environments. **Proceedings of the National Academy of Sciences of USA**, 89: 5685-5689.
- DELONG, E.F. (1998) Everything in moderation: Archaea as “non-extremophiles”. **Current Opinion in Genetics & Development**, 8: 649-654.
- DEROSA, M.; TRINCONE, A.; NICOLAUS, B.; GAMBACORTA, A. In: Life Under Extreme Conditions (Ed. diPrisco, G.) p. 61-87. (Springer, Berlin Heidelberg, 1991). Apud ALBERS, S.V.; SZABÓ, Z.; DRIESSEN, J.M. (2006) Protein secretion in the Archaea: multiple paths towards a unique cell surface. **Nature Reviews Microbiology**, 4: 537-47.
- DESANTIS, T.Z.; HUGENHOLTZ, P.; KELLER, K.; BRODIE, E.L.; LARSEN, N.; PICENO, Y.M.; PHAN, R.; ANDERSEN, G.L. (2006). NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. **Nucleic Acids Research**, 34: 394-399.
- DESANTIS, T.Z.; HUGENHOLTZ, P.; LARSEN, N.; ROJAS, M.; BRODIE, E.L.; KELLER, K.; HUBER, T.; DALEVI, D.; HU, P.; ANDERSEN, G.L. (2006) Greengenes, a Chimera-checked 16S rRNA gene database and workbench

- compatible with ARB. **Applied and Environmental Microbiology**, 72(7): 5069-5072.
- EITEN, G. The Cerrado Vegetation of Brazil.(1972) **The Botanical Review**, 38: 201-341.
- EWING, B.; GREEN, P. (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. **Genome Research**, 8:186-194.
- FELSENSTEIN J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. **Evolution**, 39:783-791.
- FORNEY, L.J.; ZHOU, X.; BROWN, C.J. (2004) Molecular microbial ecology: land of the one-eyed king. **Current Opinion in Microbiology**, 7:210-220.
- FOX, G.E.; PECHMAN, K.R.; WOESE, C.R. (1977) Comparative cataloging of 16S ribosomal ribonucleic acid: molecular approach to prokaryotic systematics. **International Journal of Systematic Bacteriology**, 27(1): 44-57.
- FOX, G.E.; WISOTZKEY, J.D.; JURTSCHUK, P. (1992) How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. **International Journal of Systematic Bacteriology**, 42(1): 166-170.
- FUHRMAN, J.A.; MCCALLUM, K.; DAVIS, A.A. (1992) Novel major archaebacterial group from marine plankton. **Nature**, 356: 148–149.
- FURLEY, P.A. e RATTER, J. A. (1988) Soil resources and plant communities of the central Brazilian cerrado and their development. **Journal of Biogeography**, 15: 97-108. Apud SILVA, M.R.S.S. **Produção de serrapilheira, biomassa e diversidade de comunidades bacterianas do solo em áreas de Cerrado sob diferentes usos e manejos**. 2004. Dissertação de mestrado, Departamento de Ecologia, Instituto de Ciências Biológicas, Universidade de Brasília.
- GABOR, E.M.; VRIES, E.J. de, JANSSEN, D.B. (2003) Efficient recovery of environmental DNA for expression cloning by indirect extraction methods. **FEMS Microbiology Ecology**, 44: 153-163.

- GANTNER, S.; ANDERSSON, A.F.; ALONSO-SÁEZ, L.; BERTILSSON, S. (2010) Novel primers for 16S rRNA-based archaeal community analyses in environmental samples. **Journal of Microbiological Methods**, 84(1): 12-18.
- GARRITY, G. M. Bergey's manual of systematic bacteriology, Volume One: The Archaea and the Deeply Branching and Phototrophic Bacteria. (Garrity, G. M.; Boone, D.R.; Castenholz, R.W.; eds.) New York: Springer-Verlag, 2^a edição, 2001.
- GRAHAM, D.E.; OVERBREEK, R.; OLSEN, G.J.; WOESE, C.R. (2000) An archaeal genomic signature. **Proceedings of the National Academy of Sciences of USA**, 97: 3304-3308.
- GRAY, N.D.; MISKIN, I.P.; OKSANA, K.; CURTIS, T.P.; HEAD, I.M. (2002) Occurrence and activity of Archaea in aerated activated sludge wastewater treatment plants. **Environmental Microbiology**, 4(3): 158-168.
- HANDELSMAN, J; RONDON, M.R.; BRADY, S.F.; CLARDY, J.; GOODMAN, R.M. (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. **Chemistry & Biology**, 5(10): 245-249.
- HANSEL, C.M.; FENDORF, S.; JARDINE, P.M.; FRANCIS, C.A. (2008) Changes in bacterial and archaeal community structure and functional diversity along a geochemically variable soil profile. **Applied and Environmental Microbiology**, 74(5): 1620-1633.
- HATZENPICHLER, R.; LEBEDEVA, E.V.; SPIECK, E.; STOECKER, K.I RICHTER, A.; DAIMS, H.; WAGNER, M. (2008) A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. **Proceedings of the National Academy of Sciences of USA**, 105(6): 2134-2139.
- HONG, S.; BUNGE, J.; LESLIN, C.; JEON, S.; EPSTEIN, S.S. (2009) Polymerase chain reaction primers miss half of rRNA microbial diversity. **The ISME Journal**, 3: 1365-1373.
- HUBER, H.; HOHN, M.J.; RACHEL, R.; FUCHS, T.; WIMMER, V.C.; STETTER, K.O. (2002) A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. **Nature**, 417: 63-68.

- HUBER, T.; FAULKNER, G.; HUGENHOLTZ, P. (2004) Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. **Bioinformatics**, 20: 2317-2319.
- HUGHES, J.B.; HELLMANN, J.J.; RICKETTS, T.H.; BOHANNAN, B.J.M. (2001) Counting the Uncountable: Statistical Approaches to Estimating Microbial Diversity. **Applied and Environmental Microbiology**, 67(10): 4399-4406.
- JIA, Z. e CONRAD, R. (2009) Bacteria rather than Archaea dominate microbial ammonia oxidation in an agricultural soil. **Environmental Microbiology**, 11(7): 1658-1671.
- JUKES T.H. and CANTOR C.R. (1969). Evolution of protein molecules. In: **Mammalian Protein Metabolism** (Munro HN, editor), Vol. III, pp. 21-132, Academic Press, New York.
- JUNG, M-Y.; PARK, S-J.; MIN, D.; KIM, J-S.; RIPJSTRA, W.I.C.; DAMSTÉ, J.S.S.; KIM, G-J.; MADSEN, E.L.; RHEE, S-K. (2011) Enrichment and characterization of an autotrophic ammonia-oxidizing archaeon of mesophilic crenarchaeal group I.1a from an agricultural soil. **Applied and Environmental Microbiology**, 77 (24): 8635-8647.
- JURGENS, G.; LINDSTROM, K.; SAANO, A. (1997) Novel group within the kingdom Crenarchaeota from boreal forest soil. **Applied and Environmental Microbiology**, 63: 803-805.
- KAN, J.; CLINGENPEEL, S.; MACUR, R.E.; INSKEEP, W.P.; LOVALVO, D.; VARLEY, J.; GORBY, Y.; MCDERMOTT, T.R.; NEALSON, K. (2011) *Archaea* in Yellowstone Lake. **The ISME Journal**, 1-12.
- KANDLER, O.; KÖNIG, H. Cell wall polymers in Archaea (Archaeobacteria). (1998) **Cellular and Molecular Life Sciences**, 54: 305-308.
- KATES, M. (1993) Membrane lipids of archaea. In: **The Biochemistry of Archaea (Archaeobacteria)**. (Kates, M.; Kushner, D.J.; Matheson, A.T.; eds). Amsterdam: Elsevier, 1993. Capítulo 9, p. 261-295.
- KEMNITZ, D.; KOLB, S.; CONRAD, R. (2007) High abundance of Crenarchaeota in a temperate acidic Forest soil. **FEMS Microbiology Ecology**, 60: 442-448.

- KÖNNEKE, M.; BERNHARD, A.E.; TORRE, J.R.; WALKER, C.B.; WATERBURY, J.B. STAHL, D.A. (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. **Nature Letters**, 437: 543-546.
- KOSKI LB e GOLDING GB (2001) The closest BLAST hit is often not the nearest neighbor. **Journal of Molecular Evolution**, 52: 540–542.
- KREADER, C. (1996) Relief of Amplification Inhibition in pcr with bovine serum albumin or T4 gene 32 protein. **Applied and Environmental Microbiology**, 62(3): 1102-1106.
- LEHTOVIRTA, L.E.; PROSSER, J.I.; NICOL, G.W. (2009) Soil pH regulates the abundance and diversity of Group 1.1c Crenarchaeota. **FEMS Microbiology Ecology**, 70: 367-376.
- LEININGER, S.; URICH, T.; SCHLOTTER, M.; SCHWARK, L.; QI, J.; NICOL, G.W.; PROSSER, J.L.; SCHUSTER, S.C.; SCHLEPER, C. (2006) Archaea predominate among ammonia-oxidizing prokaryotes in soils. **Nature Letters**, 442: 806-809.
- LEPP, P.W., BRINIG, M.M., OUVENEY, C.C., PALM, K., ARMITAGE, G.C., RELMAN, D.A.(2004) Methanogenic archaea and human periodontal disease. **Proceedings of the National Academy of Sciences of USA**, 101: 6176–6181.
- LLIRÓS, M.; CASAMAYOR, E.O.; BORREGO, C. (2008) High archaeal richness in the water column of a freshwater sulfurous karstic lake along an interannual study. **FEMS Microbiology Ecology**, 66: 331-342.
- LOZUPONE, H. e KNIGHT.(2006) Unifrac – an online tool for comparing microbial community diversity in a phylogenetic context. **BMC Bioinformatics**, 7: 371.
- MADIGAN, M.T.; MARTINKO, J.M.; STAHL, D.A.; CLARK, D.P. Capítulo 16: Microbial Evolution and Systematics In: **Brock Biology of Microorganisms**. Benjamin Cummings (Pearson) , 13a edição, p. 447-476, 2011.
- MARANHÃO, A. Capítulo VIII: Transformação Bacteriana.In: **Técnicas Básicas em Biologia Molecular** (eds. Azevedo, M.O.; Felipe, M.S.S.; Brígido, M.M.; Maranhão, A.Q.; De-Souza, M.T.) Brasília: Editora UnB, p. 129-141, 2003.
- MULLER, F.; BRISSAC, T.; LE BRIS,N.; FELBECK, H.; GROS, O. (2010) First description of giant Archaea (Thaumarchaeota) associated with putative bacterial

- ectosymbionts in a sulfidic marine habitat. **Environmental Microbiology**, 12(8): 2371-2383.
- NANNIPIERI, P.; ASCHER, J.; CECCHERINI, M.T.; LANDI, L.; PIETRAMELLARIA, G.; RENELLA, G. (2003) Microbial diversity and soil functions. **European Journal of Soil Science**, 54: 655-670.
- NUNOURA, T.; TAKAKI, Y.; KAKUTA, J.; NISHI, S.; SUGAHARA, J.; KAZAMA, H.; CHEE, G.-J.; HATTORI, M.; KANAI, A.; ATOMI, H.; TAKAI, K.; TAKAMI, H. (2011) Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. **Nucleic Acids Research**, 39 (8): 3204-3223.
- NUNOURA, T.; HIRAYAMA, H.; TAKAMI, H.; OIDA, H.; NISHI, S.; SHIMAMURA, S.; SUZUKI, Y.; INAGAKI, F.; TAKAI, K.; NEALSON, K.H.; HORIKOSHI, K. (2005) Genetic and functional properties of uncultivated thermophilic crenarchaeotes from a subsurface gold mine as revealed by analysis of genome fragments. **Environmental Microbiology**, 7(12):1967-1984.
- OCHSENREITER, T.; SELEZI, D.; QUAISER, A.; BONCH-OSMOLOVSKAYA, L.; SCHLEPER, C. (2003) Diversity and abundance of Crenarchaeota in terrestrial habitats studied by 16S RNA surveys and real time PCR. **Environmental Microbiology**, 5(9): 787-797.
- OGRAM, A.; SAYLER, G.S.; BARKAY, T. (1987) The extraction and purification of microbial DNA from sediments. *Journal of Microbiological Methods*, 7: 57-66. Apud GABOR, E.M.; VRIES, E.J. de; JANSSEN, D.B. (2003) Efficient recovery of environmental DNA for expression cloning by indirect extraction methods. **FEMS Microbiology Ecology**, 44: 153-163.
- OLSEN, J.G.; LANE, D.J.; GIOVANNONI, S.J.; PACE, N.R.; STAHL, D.A. (1986) Microbial ecology and evolution: a ribosomal RNA approach. **Annual Review of Microbiology**, 40: 337-365.
- OROS-SICHLER, M.; COSTA, R.; HEUER, H.; SMALLA, K. Molecular fingerprinting techniques to analyze soil microbial communities, In: *Modern Soil Microbiology*

- (Elsas, J.D.; Jansson, J.K.; Trevors, J.T.; eds). USA: CRC Press, 2ª edição, 2006. Capítulo 14, p.355-386.
- PACE, N.R. (1997) A molecular view of microbial diversity and the biosphere. **Science**, 276: 734-740.
- PACE, N.R. (2009) Mapping the Tree of Life: Progress and Prospects. **Microbiology and Molecular Biology Reviews**, 73(4): 565-576.
- PAZINATO, J.M.; PAULO, E.N.; MENDES, L.W.; VAZOLLER, R.F.; TSAI, S.M. (2010) Molecular Characterization of the Archaeal Community in an Amazonian Wetland Soil and Culture-Dependent Isolation of Methanogenic Archaea. **Diversity**, 2: 1026-1047.
- PEREIRA, B.A.S. Reserva Ecológica do IBGE, Brasília (DF). Disponível em: <<http://www.recor.org.br/cerrado/solos.html>> . Acesso em 18 de março de 2011.
- PESTER, M. SCHLEPER, C.; WAGNER, M. (2011) The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. **Current Opinion in Microbiology**, 14:1-7.
- PRESTON, C.; WU, K.Y.; MOLINKSKI, T.F.; DELONG, E.F. (1996) A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp.nov. **Proceedings of the National Academy of Sciences of USA**, 93: 6241-6246
- QUIRINO, B.F.; PAPPAS, G.J.; TAGLIAFERRO, A.C.; COLLEVATTI, R.G.; NETO, E.L.; SILVA, M.R.S.S. da; BUSTAMANTE, M.M.C.; KRÜGER, R.H. (2009) Molecular phylogenetic diversity of bacteria associated with soil of the savanna-like Cerrado vegetation. **Microbiological Research**, 164: 59-70.
- REATTO, A.; CORREIA, J.R.; SPERA, S.T. Solos do bioma Cerrado: aspectos pedológicos, In: **Cerrado: ambiente e flora** (SANO, S.M. ; ALMEIDA, S.P.; Eds.) Planaltina: Embrapa-CPAC, 1998. Capítulo 2, p. 47-86.
- RIBEIRO, J.F.; WALTER,B.M.T.; Fitofisionomias do bioma Cerrado. In: **Cerrado: ambiente e flora** (SANO, S.M.; ALMEIDA, S.P.; Eds.) Planaltina: Embrapa-CPAC, 1998. Capítulo 3, p. 90-166.

- RIVERA, M.C; JAIN, R.; MOORE, J.E.; LAKE, J.A. (1998) Genomic evidence for two functionally distinct gene classes. **Proceedings of the National Academy of Sciences of USA**, 95: 6239-6244.
- ROBERTSON, C.E.; HARRIS, J.K.; SPEAR, J.R.; PACE, N.R. (2005) Phylogenetic diversity and ecology of environmental Archaea. **Current Opinion in Microbiology**, 8: 638-642.
- ROBERTSON, C.E.; SPEAR, J.R.; HARRIS, J.K.; PACE, N.R. (2009) Diversity and stratification of archaea in a hypersaline microbial mat. **Applied and Environmental Microbiology**, 75(7): 1801-1810.
- ROESCH, L.F.W.; FULTHORPE, R.R.; RIVA, A.; CASELLA, G.; HADWIN, A.K.M.; KENT, A.D.; DAROUB, S.H.; CAMARGO, F.A.O.; FARMERIE, W.G.; TRIPLET, E.W. (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. **The ISME Journal**, 1: 283-290.
- ROOSE-AMSALEG, C.L.; GARNIER-SILLAM, E.; HARRY, M. (2001) Extraction and purification of microbial DNA from soil and sediment samples. **Applied Soil Ecology**, 18: 47-60.
- ROUX, S.; ENAULT, F.; BRONNER, G.; DEBROAS, D. (2011) Comparison of 16S rRNA and protein-coding genes as molecular markers for assessing microbial diversity (Bacteria and Archaea) in ecosystems. **FEMS Microbiology Ecology**, 78: 617-628.
- SAITOU N. and NEI M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution** 4:406-425.
- SCHLEPER, C.; JURGENS, G.; JONUSCHEIT, M. (2005) Genomic studies of uncultivated Archaea. **Nature Reviews Microbiology**, 3: 479-488.
- SCHLOSS, P. D., LARGET, B. R. ;HANDELSMAN, J. 2004. Integration of microbial ecology and statistics: a test to compare gene libraries. **Applied Environmental Microbiology**, 70:5485–5492.
- SCHLOSS, P.D. e HANDELSMAN, J. (2004) Status of the Microbial Census. **Microbiology and Molecular Biology Reviews**, 68 (4): 686-691.

- SCHLOSS, P.D.; WESTCOTT, S.L.; RYABIN, T.; HALL, R.J.; HARTMANN, M.; HOLLISTER, E.B.; LESNIEWSKI, R.A.; OAKLEY, B.B.; PARKS, D.H.; ROBINSON, C.J.; SAHL, J.W.; STRES, B.; THALLINGER, G.G.; VAN HORN, D.J.; WEBER, C.F. (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. **Applied and Environmental Microbiology**, 75(23): 7537-7541.
- SILVA, M.R.S.S. **Produção de serrapilheira, biomassa e diversidade de comunidades bacterianas do solo em áreas de Cerrado sob diferentes usos e manejos**. 2004. Dissertação de mestrado, Departamento de Ecologia, Instituto de Ciências Biológicas, Universidade de Brasília.
- SKILLMAN, L.C.; EVANS, P.N.; STRÖMPL, C.; JOBLIN, K.N. (2006) 16S rDNA directed PCR primers and detection of methanogens in the bovine rumen. **Letters in Applied Microbiology**, 42: 222-228.
- SPANG, A.; HATZENPICHLER, R.; BROCHIER-ARMANET, C.; RATTEI, T.; TISCHLER, P.; SPIECK, E.; STREIT, W.; STAHL, D.A.; WAGNER, M.; SCHLEPER, C. (2010) Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. **Trends in Microbiology**, 18(8): 331-340.
- STACKEBRANDT E, GOEBEL BM. (1994) Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. **International Journal of Systematic and Evolutionary Microbiology**, 44:846-849.
- STACKEBRANDT, E. (2006) Exciting times: the challenge to be a bacterial systematist. In: Molecular Identification, Systematics, and Population Structure of Prokaryotes (Stackebrandt, E., Ed.) Germany: Springer-Verlag Berlin Heidelberg, p.1-21.
- STALEY, J.T.; KONOPKA, A. (1985) Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. **Annual Review of Microbiology**, 39: 321-46.

- STANDING, D.; KILLHAM, K. The Soil Environment. In: *Modern Soil Microbiology* (Elsas, J.D.; Jansson, J.K.; Trevors, J.T.; eds). USA: CRC Press, 2a edição, 2006. Capítulo 1, p.1-22.
- SWIFT, M.J; BIGNELL, D.E.; MOREIRA, F.M.S.; HUISING, E.J. The inventory of soil biological diversity: concepts and general guidelines. In: *A handbook of Tropical Soil Biology: Sampling & Characterization of Below-ground Biodiversity* (MOREIRA,F.M.S.; HUISING,E.J.; Bignell, D.E.; Eds). Sterling,USA: Earthscan, 2008. Páginas 1-16.
- TAGLIAFERRO, A.C. 2005. **Análise molecular da diversidade bacteriana de solos do Cerrado utilizando bibliotecas de rDNA 16S - uma perspectiva biotecnológica.** 125f. Dissertação (Mestre em Ciências Genômicas e Biotecnologia) – Universidade Católica de Brasília, Brasília.
- TAMURA, K.; PETERSON, D.; PETERSON, N.; STECHER, G.; NEI, M.; KUMAR, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance and Maximum Parsimony Methods. **Molecular Biology and Evolution**, 28: 2731—2739.
- TORSVIK, V.; OVREAS, L. Microbial phylogeny and diversity in soil. In: **Modern Soil Microbiology** (Elsas, J.D.; Jansson, J.K.; Trevors, J.T.; eds). USA: CRC Press, 2a edição, 2006. Capítulo 2, p.23-52.
- TORSVIK, V.e OVREAS, L. (2002) Microbial diversity and function in soil: from genes to ecosystems. **Current Opinion in Microbiology**, 5(3): 240-245.
- TOURNA, M.; STIEGLMEIER, M.; SPANG, A.; KONNEKE, M.; SCHINTLMEISTER, A.; URICH, T.; ENGEL, M.; SCHLOTTER, M.; WAGNER, M.; RICHTER, A.; SCHLEPER, C. (2011) *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. **Proceedings of the National Academy of Sciences of USA**, 108(20): 8420-8425.
- VAN ELSAS, J.D.; TORSVIK, V.; HARTAMANN, A.; OVREAS, L; JANSSON, J. The bacteria and archaea in soil. In: **Modern Soil Microbiology** (Elsas, J.D.; Jansson, J.K.; Trevors, J.T.; eds). USA: CRC Press, 2a edição, 2006. Capítulo 4, p.83-105.

- WALTER, B.M.T.; CARVALHO, A.M. de; RIBEIRO, J.F. O conceito de savana e de seu componente Cerrado, In: **Cerrado: ecologia e flora - Volume 1**. (SANO, S.M.; ALMEIDA, S.P.; RIBEIRO, J.F.; eds) Embrapa Cerrados, Brasília, DF: Embrapa Informação Tecnológica, 2008, capítulo 1, 21-45.
- WANG, Q.; GARRITY, G.M.; TIEDJE, J.M.; COLE, J.R. (2007) Naïve Bayesian Classifier for Rapid Assignment of rRNA sequences into the new bacterial taxonomy. **Applied Environmental Microbiology**, 73 (16): 5261-5267.
- WEGLEY, L.; YU, Y.; BREITBART, M.; CASAS, V.; KLINE, D.I.; ROHWER, F. (2004) Coral-associated Archaea. **Marine Ecology Progress Series**, 273: 89-96.
- WHITE, J.R.; NAVLAKHA, S.; NAGARAJAN, N.; GODSI, M-P.; KINGSFORD, C.; POP, M. (2010) Alignment and clustering of phylogenetic markers – implications for microbial diversity studies. **BMC Bioinformatics**, 11: 152.
- WHITTAKER, R.H. (1969) New concepts of kingdoms of organisms. **Science**, 163:150-160.
- WINTZINGERODE, F.V.; GÖBEL, U.B.; STACKEBRANDT, E. (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. **FEMS Microbiology Reviews**, 21:213-229.
- WIT, R. de; BOUVIER, T. (2006) “Everything is everywhere, but, the environment selects”; what did Baas Becking and Beijerinck really say? **Environmental Microbiology**, 8(4): 755-758.
- WOESE, C.R.; FOX, G.E. (1977) Phylogenetic structure of the prokaryotic domain: The primary kingdoms. **Proceedings of the National Academy of Sciences of USA**, 74(11): 5088-90.
- WOESE, C.R.; KANDLER, O.; WHEELIS, M.L. (1990) Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria and Eucarya. **Proceedings of the National Academy of Sciences of USA**, 87: 4576-79.
- WOESE, C.R., MAGRUM, L.J., GUPTA, R., SIEGEL, R.B., STAHL, D.A., KOP, J., CRAWFORD, N., BROSIUS, J., GUTELL, R. et al. (1980) Secondary structure model for bacterial 16S ribosomal RNA: phylogenetic, enzymatic and chemical evidence. **Nucleic Acids Research**, 8, 2275–2293.

- YASSIN, A.F., KROPPESTEDT, R.M., LUDWIG W.(2003) *Corynebacterium glaucum* sp. nov. **International Journal of Systematic and Evolutionary Microbiology**, 53:705-709.
- ZHANG, L.M.; OFFRE, P.R.; HE, J.Z.; VERHAMME, D.T.; NICOL, G.W.; PROSSER, J.I. (2010) Autotrophic ammonia oxidation by soil thaumarchaea. **Proceedings of the National Academy of Sciences of USA**, 107(40): 17240-17245.
- ZUCKERKANDL, E.; PAULING, L. (1965) Molecules as documents of evolutionary history. **Journal of Theoretical Biology**, 8(2): 357-366.

ANEXO I. Trabalhos apresentados em congresso

Trabalho apresentado na forma de pôster no 3º Congresso sobre Diversidade Microbiana da Amazônia e XII Encontro Nacional de Microbiologia Ambiental (ENAMA) realizado em Manaus em Dezembro de 2010

Estudo da diversidade filogenética de *Archaea* no solo e no sedimento do bioma Cerrado

Elisa Catão, Julianna B. Peixoto, Thiago Rodrigues, Vitor Valdez, Helena Magaldi, Cynthia Kyaw

A microbiologia ambiental atual procura descrever os microorganismos a partir de técnicas independentes de cultivo dado que apenas uma pequena fração dos mesmos é capaz de formar colônias em meios sólidos. O Domínio *Archaea* vem sendo descrito principalmente através dessas técnicas moleculares e atualmente é dividido em dois filos: Crenarchaeota e Euryarchaeota. Dado que sequências do gene rRNA 16S de *Archaea* foram descritos de diversas amostras, as archaeas passaram a ser consideradas ubíquas. No entanto pouco se conhece sobre sua importância nos ciclos biogeoquímicos e sua diversidade filogenética e funcional ainda é subestimada. O nosso objetivo é o de caracterizar a diversidade e a riqueza de archaeas em amostras de solo e de sedimento do Cerrado - a principal vegetação do Brasil Central, sendo o segundo maior do país. O cerrado constitui um dos biomas de maior diversidade biológica de nosso planeta; é uma das savanas mais ricas do mundo; é considerado um dos 25 "hotspots" mundiais de importância para a conservação. O DNA genômico total de amostras compostas de solo e de sedimentos foi extraído diretamente com o *PowerSoil DNA Isolation Kit* (MO Bio Laboratories Inc.). Primers *Archaea*-específicos foram utilizados para amplificar a região do gene marcador 16S rRNA do DNA extraído. Após a amplificação, os fragmentos foram clonados no vetor pGEM®-T Easy Vector System I (Promega, Madison, USA) e transformados. Clones foram aleatoriamente selecionados e sequenciados no sequenciador automático ABI3730 (Applied Biosystems). Em um screening inicial, todas as sequências de boa qualidade apresentaram correspondência de 100% com o Domínio *Archaea* segundo alinhamento com sequências do banco de dados *Ribosomal Database Project*. Ao alinhá-las com o *GenBank* da *National Center for Biotechnology* todas correspondiam a clones de archaea não cultiváveis. A análise dessas sequências nos permite identificar as archaeas presentes naquelas amostras ambientais e, posteriormente, acessar metabolismos ainda desconhecidos e com potencial biotecnológico.

Trabalho apresentado na forma de pôster no 26º Congresso Brasileiro de Microbiologia realizado em Foz do Iguaçu em Outubro de 2011

Riqueza do Domínio *Archaea* em solo do bioma Cerrado

Elisa Catão, Ricardo H. Kruger, Cynthia M. Kyaw

O Domínio *Archaea* vem sendo descrito principalmente através de técnicas moleculares, independentes de cultivo e, atualmente, é dividido em dois filos formalmente aceitos, *Crenarchaeota* e *Euryarchaeota*, além de três outros propostos – *Korarchaeota*, *Nanoarchaeota* e *Thaumarchaeota*. Uma vez que sequências do gene de rRNA 16S de *Archaea* foram descritas a partir de amostras de diferentes ambientes, as archaeas passaram a ser consideradas organismos ubíquos na natureza. Este trabalho tem como objetivo caracterizar a riqueza de archaeas em amostras de solo de Cerrado - a principal vegetação do Brasil Central e a segunda maior do país. O Cerrado constitui uma das savanas mais ricas do mundo, sendo considerado um dos 25 “hotspots” mundiais de importância para a conservação. O DNA metagenômico de amostras de solos provenientes de duas fitofisionomias distintas, campo denso e mata de galeria, com réplica biológica de cada, foi extraído empregando-se o *PowerSoil DNA Isolation Kit* (MOBio Laboratories Inc.). Iniciadores *Archaea*-específicos foram utilizados em experimentos de PCR visando amplificar regiões do gene de rRNA16S a partir do DNA metagenômico de cada solo. Os fragmentos foram ligados ao vetor pGEM®-T Easy Vector System I (Promega, Madison, USA) e transformados em células de *Escherichia coli* DH5 . Clones recombinantes foram aleatoriamente selecionados e seqüenciados no sequenciador automático ABI3730 (Applied Biosystems). Os resultados obtidos a partir das amostras de solo revelaram sequências com correspondência de 90-100% com o domínio *Archaea*. A maioria das sequências foram classificadas no Filo *Crenarchaeota*, Classe *Thermoprotei*, segundo alinhamento com o banco de dados *Ribosomal Database Project*.

Apoio financeiro: FAP/DF