Almir José Ferreira

**DIVERSIDADE E ESTRUTURA DA COMUNIDADE
BACTERIANA ASSOCIADA ÀS ARMADILHAS DA
PLANTA CARNÍVORA *Utricularia gibba*
(LENTIBULARIACEAE) E AO AMBIENTE AQUÁTICO**

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

São Paulo
2011

Almir José Ferreira

**DIVERSIDADE E ESTRUTURA DA COMUNIDADE
BACTERIANA ASSOCIADA ÀS ARMADILHAS DA PLANTA
CARNÍVORA *Utricularia gibba* (LENTIBULARIACEAE) E
AO AMBIENTE AQUÁTICO**

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

Área de concentração: Biotecnologia

Orientador:

Prof. Dr. Welington Luiz Araújo

Versão Original

São Paulo

2011

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

© reprodução total

Ferreira, Almir José.

Diversidade e estrutura da comunidade bacteriana associada às armadilhas da planta carnívora *Utricularia gibba* (Lentibulariaceae) e ao ambiente aquático / Almir José Ferreira. -- São Paulo, 2011.

Orientador: Welington Luiz de Araújo.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação Interunidades em Biotecnologia USP/IPT/Instituto Butantan. Área de concentração: Biotecnologia. Linha de pesquisa: Genética molecular e Ecologia Microbiana.

Versão do título para o inglês: Diversity and structure of bacterial communities associated to the traps of the carnivorous plant *Utricularia gibba* (Lentibulariaceae) and aquatic environment

Descritores: 1. Diversidade 2. Interação bactéria-planta 3. *Polynucleobacter* 4. *Acidobacterium* 5. Pirosequenciamento I. Araújo, Welington Luiz de II. Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação Interunidades em Biotecnologia USP/IPT/Instituto Butantan III. Título.

ICB/SBIB0220/2011

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

Candidato(a): Almir José Ferreira.

Título da Dissertação: Diversidade e estrutura de comunidade bacteriana associada às armadilhas da planta carnívora *Utricularia gibba* (Lentibulariaceae) e ao ambiente aquático

Orientador(a): Welington Luiz de Araújo.

A Comissão Julgadora dos trabalhos de Defesa da **Dissertação de Mestrado**,
em sessão pública realizada a/...../....., considerou

() Aprovado(a)

() Reprovado(a)

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Presidente: Assinatura:
Nome:
Instituição:

DEDICO

A minha mãe e irmãs,

Pelo amor, carinho e o altruísmo para eu chegasse aqui

OFEREÇO

*Aos meus amigos,
principalmente do laboratório que
me ajudaram a seguir em frente
nestes quatro anos*

OBRIGADO!

AGRADECIMENTOS

Ao Prof. Dr. Welington Luiz de Araújo, pela oportunidade dada, pela confiança depositada em meu trabalho, os ensinamentos e por me ajudar nos momentos mais difíceis, atuando muitas vezes até como um pai para mim.

Ao Felipe pela amizade, companheirismo e cobranças para fazer o melhor dentro e fora do laboratório.

Aos meus queridos amigos que convivo dentro e fora do laboratório, os quais contribuíram não só para meu trabalho, mas também para o meu engrandecimento tanto profissional como humano:

À Aline Neves pela confiança inicial no meu trabalho e auxílio nos primeiros passos.

À Ana Marie pelo suporte e confiança

À Carol Bertini pela amizade e cumplicidade mesmo em momentos difíceis;

Ao Deibs pelas conversas e tantos galhos quebrados;

À Emy Mano pela inspiração, força e cumplicidade;

À Fernanda Caravieri pelo carinho e bom humor impulsionadores;

À Flávinha Holanda pelos isolamentos de endófitos e gargalhadas;

Ao Fábio Moura pela força nos momentos difíceis e admirável prestatividade;

À Léia Fávaro pelos conselhos e apoio;

À Marília Sanchez pela amizade de carinho tão caro e valioso;

À Maristela Ciraulo pela inspiração, força e esperança no amanhã;

À Thaís Prado pelo apoio nas atividades no laboratório e amizade.

Também quero agradecer aos amigos e colegas que além da amizade me proporcionaram sempre que precisei toneladas de reagentes, centenas de horas em equipamentos e principalmente as portas abertas e idéias:

À Ana Claudia, Cris bombom, Daiene Santos, Débora e Kelly Chen pelos inúmeros galhos quebrados e momentos de descontração no laboratório de Genômica Estrutural;

À Corina, Cris Castilho, Fabi, Gustavo, Jaque, Juliana, Juliana Viana, Laura, Pâmela Resende, Paola, Sara da Mata, Suhaila, Tamara Cruz, Tati Okazaki, Vânia, por milhares de favores e amizade no Lagoaa;

À Aline Oliveira, Binho, Caio César, Clarissa, Daniel Borges, Gabriel Honorato, Joel Meira, Kelly Rosa, Mari Miranda, Rafael Campos, Renata Vieira e Saura, Vanessa Nessner do NCA pelo apoio;

Ao Anderson Ferreira, Andréa Bogas, Armando, Carol Quecine, Digue, Fernando Andreote, Francisco Andreote, Joelma e Manuela Dourado da ESALQ pela Luz em todos os momentos que precisamos;

À Aline Lopes, Adriana, Ana Carla, Cris Brum, Fernandinha Helena, Fernanda Storte, Gabi Delphino, Janaina Mendes, Juliana Crescente, Luciana Francisco, Monick Aquino, Renato, Sara Burela e Silvia Umezaki pela companhia, risadas e momentos muito especiais no LABMEM.

Aos meus queridos amigos que sempre me apoiaram, mesmo nos momentos mais complicados e vibraram nos de sucesso:

Danilo, Evelyne, Juninho, Jefferson, Paulinha, Solange e Tânia.

À minha família que fez o melhor por mim se esforçando ao máximo para me dar estrutura, amor e consolo nas horas que precisei.

Às Instituições, Universidade Mogi das Cruzes pela estrutura, à Universidade de São Paulo pela oportunidade de mestrado e à Fapesp pelo financiamento do projeto.

Muito Obrigado!

*“Aprender é a única coisa de que a mente
nunca se cansa, nunca tem medo e nunca se
arrepende”.*

Leonardo da Vinci

RESUMO

FERREIRA, A. J. **Diversidade e estrutura de comunidade bacteriana associada às armadilhas da planta carnívora *Utricularia gibba* (Lentibulariaceae) e ao ambiente aquático**, 2011. 71 f. Dissertação (Mestrado em Biotecnologia) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2011.

A manutenção dos ecossistemas é dependente da interação dos diversos organismos presentes, dentre os quais estão as comunidades microbianas. Neste contexto, o Brasil é tido como um dos países com a maior biodiversidade, porém ainda existem poucos estudos sobre a diversidade microbiana em ambientes aquáticos e a associação desta comunidade com plantas carnívoras. Apesar dos avanços da metagenômica utilizando a clonagem combinada à técnica de sequenciamento, ainda existe uma grande lacuna no conhecimento a respeito da diversidade de microrganismos presentes no ambiente. Dessa forma, o objetivo do presente trabalho foi avaliar por meio do sequenciamento em larga escala a diversidade e estrutura da comunidade bacteriana associada às armadilhas (utrículos) da planta carnívora *Utricularia gibba* e ao ambiente aquático. Para alcançar os objetivos foi realizado o pirosequenciamento 454 parcial do gene 16S rRNA de bactérias associadas aos utrículos de plantas representantes da espécie *U. gibba* e da água de dois locais na várzea do Alto Tietê. Foram avaliadas 14.020, 14.520 e 12.371 sequências para *U. gibba*, local 1 (onde está a planta analisada) e o local 2, respectivamente. Os resultados das análises de OTU's indicaram que ocorreu um agrupamento das comunidades bacterianas presentes na água (independente do local amostrado), sendo significativamente diferente daquela presente nos utrículos. Foi observado que a comunidade bacteriana da água é composta principalmente por membros dos filos Proteobacteria, Actinobacteria, Firmicutes e Verrucomicrobia, enquanto a comunidade associada à *U. gibba* é composta por membros dos filos Proteobacteria, Firmicutes, Cyanobacteria e Acidobacteria. Foi observado que o gênero *Polynucleobacter* foi dominante nos dois ambientes aquáticos, visto que em média 13% das sequências obtidas da água apresentam similaridade a este grupo. Entretanto, este gênero não foi observado nos utrículos, onde *Acidobacterium* e *Methylococcus*, os quais não foram detectados na água foram dominantes. Tendo em vista que a comunidade bacteriana presente nos utrículos deve ter sido selecionada a partir de bactérias da água, os resultados observados sugerem que o ambiente presente no interior dos utrículos selecione bactérias específicas, independente da sua frequência na água. O papel destas bactérias no interior dos utrículo é ainda incerto, mas pode ser sugerido que podem participar da degradação das presas capturadas por *U. gibba*.

Palavras-chave: Diversidade. Interação bactéria-planta. *Polynucleobacter*. *Acidobacterium*. Pirosequenciamento.

ABSTRACT

FERREIRA, A. J. **Diversity and structure of bacterial communities associated to the traps of the carnivorous plant *Utricularia gibba* (Lentibulariaceae) and aquatic environment**, 2011. 71 p. Master Thesis (Biotechnology) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2011.

The maintenance of ecosystems is dependent of the various organisms interactions, among which are the microbial communities. In this context, Brazil is considered one of the countries with the greatest biodiversity, but there are still few studies about microbial diversity in aquatic environments and the carnivorous plants association communities. Despite of advances in metagenomics using cloning combined with sequencing technique, there is still a large gap in knowledge about the microorganisms diversity in the environment. Thus, the aim of this study was to evaluate by large-scale sequencing the bacterial community diversity and structure associated to the traps (utricles) of the carnivorous plant *Utricularia gibba* and its aquatic environment. To achieve the aim was performed partial 454 pyrosequencing of rRNA 16S gene from bacteria associated to *U. gibba* utricles and water from two sites in the floodplain of the Alto Tietê, where this plant is found. We evaluated 14020 sequences from bacteria associated to *U. gibba*, 14520 and 12371 from two samples water, Site 1 (where the plant is analyzed) and Site 2, respectively. The results of the OTU's analysis indicated that there was a group of bacterial communities in water (regardless of location sampled), being significantly different from that present in the utricles. It was observed that the water bacterial community is composed primarily of the Proteobacteria, Actinobacteria, Firmicutes and Verrucomicrobia phyla members, while the community associated to *U. gibba* comprises the Proteobacteria, Firmicutes, Cyanobacteria, and Acidobacteria phyla. It was also observed that the genus *Polynucleobacter* were dominant in both environments aquatics, since on average 13% of the sequences obtained from the water, have similarity to this group. However, this genus was not observed in utricles, where *Acidobacterium* and *Methylococcus*, which were not detected in the water were so dominant. Since the bacterial community present in the utricles should have been selected from water, the observed results suggest that the present environment inside the utricles select specific bacteria, independent of their frequency in water. The o role of these bacteria inside the utricle is still uncertain, but may be suggested that may participate in the prey degradation captured by *U. gibba*.

Key words: Diversity. Bacteria-plant interaction. *Polynucleobacter*. *Acidobacteria*. Pyrosequencing.

LISTA DE FIGURAS

- Figura 1-** Histograma obtido pelo RDP (Ribosomal database Project) indicando o resultado da seleção de sequências com qualidade superior para ser utilizada nas análises.....31.
- Figura 2-** Histograma obtido do RDP indicando o tamanho dos fragmentos através das posições das bases de início e fim no alinhamento.....33.
- Figura 3-** Análise de *Clusters* indicando o número de UTO's formadas de acordo com a similaridade.....33.
- Figura 4-** Curvas de rarefação construídas a partir de análises realizadas no programa Mothur indicando o efeito do esforço no sequenciamento para estimar o número de UTO's.35.
- Figura 5-** Diagrama de Venn apresentando UTOs (Unidades Taxonômicas Operacionais) de acordo com os níveis de similaridade.....36.
- Figura 6-** Heatmap apresentando UTO's de acordo com os níveis de similaridade, onde as linhas coloridas representam as UTO's e a gradação de cores indica a quantidade de mais sequências nas UTO's.....38.
- Figura 7-** Percentual de Filos bacterianos encontrados. Em destaque, filos presentes com maior percentual nos tratamentos.....41.
- Figura 8-** Classes de Proteobacteria identificadas e responsáveis por 54% da comunidade bacteriana total analisada.43.
- Figura 9-** Classes de Firmicutes identificadas e responsáveis por 8% da comunidade bacteriana total analisada.43.
- Figura 10-** Classes de Acidobacteria identificadas, comunidade bacteriana dominante associada aos utrículos de *Utriculária gibba*.44.
- Figura 11-** Proporção de gêneros de maior similaridade com as sequências obtidas de *U. gibba* e dos locais 1 e 2.....47.

LISTA DE TABELAS

Tabela 1. Primers adaptados capazes de flanquear a região 968-1401 do gene 16S rRNA do domínio bactéria.....	29.
Tabela 2. Distribuição de UTOs de acordo com a similaridade e tratamento.....	35.
Tabela 3. Estimadores de riqueza Chao-1 e Ace aplicados de acordo com o nível taxonômico inferido para os valores de agrupamento por similaridade das bibliotecas analisadas.....	39.
Tabela 4. Índices de diversidade Shannon-Winner e Simpson aplicados de acordo com o nível taxonômico inferido para os valores de agrupamento por similaridade das bibliotecas analisadas.....	40.
Tabela 5. Significância das diferenças encontradas na comparação dos filos.....	42.
Tabela 6. Significância encontrada na comparação das diferentes classes.....	45.

LISTA DE SIGLAS E ABREVIATURAS

APG	Angiosperm Phylogeny Group
NCBI	Nacional Center of Biotechnology Information
PCR	Polymerase Chain Reaction
RDP	Ribosomal Database Project
SIMO	Sapelo Island Microbial Observatory
UTO	Unidade Taxonômica Operacional

SUMÁRIO

1 INTRODUÇÃO	15.
2 REVISÃO DA LITERTURA.....	16.
2.1 Microrganismos e o Ecossistema Aquático.....	16.
2.2 Diversidade e distribuição de <i>Utricularia</i> (Lentibulariaceae)	17.
2.3 Caracterização da Comunidade Microbiana por Métodos Independentes de Cultivo..	
.....	17.
2.3.1 Análise da Diversidade Microbiana.....	17.
2.3.2 Estimativas da Composição da Comunidade Microbiana.....	20.
2.4 Tecnologia de Pirosequenciamento.....	21.
2.5 Dados Metagenômicos.....	22.
2.5.1 Pré-processamento.....	22.
2.5.2 Ferramentas aplicadas aos dados de pirosequenciamento.....	23.
2.5.3 Ferramentas RDP - Pyrosequencing Pipeline.....	24.
2.5.4 Alinhamento.....	25.
3 OBJETIVOS.....	27.
3.1 Objetivo Geral.....	27.
3.2 Objetivos Específicos.....	27.
4 MATERIAIS E MÉTODOS.....	28.
4.1 Área de coleta e Extração de DNA total bacteriano.....	28.
4.2 Amplificação do gene 16S rRNA de bactérias.....	28.
4.3 Purificação, quantificação e pirosequenciamento do gene 16S rRNA.....	29.
4.4 Análise e Processamento de Dados.....	29.
4.4.1 Filtragem de sequências de acordo com a qualidade.....	29.
4.4.2 Estimativa de riqueza e diversidade.....	30.
4.4.3 Classificação das sequências.....	30.
4.4.4 Cálculo das diferenças entre as bibliotecas.....	30.
5. RESULTADOS	31.
5.1 Preparo das amostras.....	31.
5.2 Análise e seleção qualitativa das sequências	31.
5.3 Alinhamento.....	32.
5.4 Construção de Clusters.....	33.

5.5 Rarefação.....	34.
5.6 Distribuição de UTO's entre os locais.....	35.
5.7 Riqueza e diversidade.....	40.
5.8 Identificação das sequências.....	40.
6 DISCUSSÃO.....	48.
7 CONCLUSÕES.....	59.
REFERÊNCIAS	60.

1 INTRODUÇÃO

A descoberta da dimensão da diversidade bacteriana em ambientes complexos como o solo, água, rizosfera, interior (endófitos) e superfícies (epífitas) de plantas, junto com as técnicas que utilizam a recuperação de DNA e RNA bacterianos diretamente do ambiente tem aumentado o interesse no estudo de comunidades microbianas de ambientes pouco avaliados. Essas técnicas podem ter várias aplicações, entre elas a identificação de novas espécies bacterianas, descoberta de agentes para biorremediação, determinação da qualidade do solo por meio de populações indicadoras e o estudo de impactos sobre comunidades bacterianas, como por exemplo, os diferentes tipos de manejo, a presença de poluentes e o cultivo de diferentes genótipos de plantas. Além disso, durante a interação com a planta, algumas espécies bacterianas podem produzir compostos com grande potencial biotecnológico, como fármacos e enzimas.

As plantas carnívoras apresentam relação ainda pouco conhecida com microrganismos como algas, fungos, protozoários e bactérias. Para alguns grupos taxonômicos, sabe-se que a presença de bactérias é de fundamental importância para as atividades de digestão das presas capturadas. A família Lentibulariaceae, composta por espécies de plantas carnívoras, é formada pelos gêneros *Utricularia*, *Genlisea* e *Pinguicula*. As espécies de *Utricularia* apresentam uma intrigante modificação em sua porção vegetativa, resultando em uma das mais complexas estruturas foliares existentes no Reino Plantae, conhecida como utrículo. Essa armadilha possui forma ovalada, diferentes tipos de tricomas (tectores e glandulares) além da presença de apêndices próximos à entrada do utrículo. Adaptados para a absorção de nutrientes, os utrículos se apresentam como órgãos de sucção, absorvendo pequenos organismos e os digerindo tanto com suas próprias enzimas como também por comunidades de microrganismos viventes em seu interior. Dessa forma, é de grande interesse avaliar a diversidade da comunidade bacteriana associada a essas plantas e o entendimento do papel destas bactérias para o desenvolvimento da planta hospedeira. Entretanto, para isso se faz necessário o conhecimento da diversidade bacteriana do ambiente onde estas plantas se desenvolvem, permitindo assim estudar de que forma estas plantas selecionam a comunidade bacteriana no interior dos utrículos.

2 REVISÃO DA LITERATURA

2.1 Microrganismos e o Ecossistema Aquático

A presença dos mais diversos seres vivos está ligada à atividade metabólica proveniente da diversidade de microrganismos presentes na natureza (STOLZ et al., 1989; TRUPER, 1992). Os microrganismos são fundamentais para a manutenção dos ecossistemas, atuando principalmente em processos como a ciclagem de carbono, a fixação biológica do nitrogênio atmosférico, desnitrificação, produção de metano, redução de sulfato, transformação de metais e de diferentes moléculas (AZEVEDO, 1998; HAINES et al., 2002). Assim, a ciclagem de nutrientes pode ser afetada pela comunidade de microrganismos presentes no local, resultando em uma alteração de compostos nutricionais disponíveis para as plantas e afetando seu crescimento (REBER, 1992). Além disso, os microrganismos participam de inúmeros processos complexos, dentre eles, a fotossíntese, ciclos biogeoquímicos, conservação da fertilidade e estrutura de solos (STOLZ et al. 1989; TRUPER, 1992). Apesar das bactérias serem os organismos mais diversos do planeta, estima-se que somente uma pequena fração foi identificada (KENNEDY, 1999). Esta carência de informações não é diferente com a microbiota aquática dulcícola. À medida que diversas alterações ambientais ocorrem, esta diversidade microbiana pode variar em riqueza de espécies e quantidade de indivíduos, uma vez que a densidade populacional é regulada por fatores como fotoperíodo, temperatura e demanda por nutrientes. Essas variações podem ser determinantes para a distribuição de espécies de microrganismos (RIKHVANOV et al., 1999). Dessa forma, impactos ambientais como descarte de efluentes não tratados contribuem para a rápida degradação desses ecossistemas. Além disso, ocorre um acúmulo de metais pesados no ambiente, prejudicando a qualidade das águas e atividades dependentes do seu uso, desde domésticas e agrícolas até as atividades industriais (MOHANTY et al., 2000). A crescente atividade antropogênica tem implicado na mudança da diversidade microbiana nos ecossistemas, bem como seu funcionamento, podendo levar à extinção de espécies capazes de promover a manutenção destes ambientes e o equilíbrio ecológico (AZEVEDO, 1998). Por esse motivo, a bioindicação através de microrganismos tem sido amplamente utilizada para monitorar e auxiliar na recuperação ambiental (HAINES et al., 2002).

2.2 Diversidade e distribuição de *Utricularia* (Lentibulariaceae)

O gênero *Utricularia* possui a maior riqueza de espécies carnívoras quando comparado aos demais táxons que apresentam o mesmo hábito. É representado por aproximadamente 150 espécies, das quais em torno de 60 ocorrem no Brasil. O gênero representa a maior parte de espécies existentes na família Lentibulariaceae L.C. Richard (Lamiales, Euasterídeas) (APG II, 2003) na qual se encontram também os gêneros *Pinguicula* e *Genlisea*, todos com distribuição cosmopolita, principalmente em regiões tropicais e subtropicais, sendo ausentes apenas em regiões áridas e algumas ilhas oceânicas (TAYLOR, 1989).

As espécies deste gênero apresentam hábitos diversos, ocupando ambientes aquáticos ou terrestres, adaptadas nas formas rupícolas, reofíticas ou mesmo epífitas (TAYLOR, 1989). *Utricularia gibba* L. e *U. neottioides* A. St.-Hil.; Girard são espécies aquáticas, ambas encontradas em corpos d'água (riachos, lagos) na forma submersa, porém a primeira sendo livre-flutuante enquanto que a segunda ocorre adpressa às rochas (hábito reofítico). *U. gibba* e *U. subulata* apresentam distribuição pantropical e ampla no Brasil, enquanto que *U. neottioides* ocorre do norte à região sudeste brasileira e *U. reniformes* apresenta distribuição restrita ao sul e sudeste (TAYLOR, 1989; CORRÊA; MAMEDE, 2002), sendo que em alguns casos, estão sobre ameaça de desaparecer em algumas regiões devido ao desaparecimento das condições ideais de crescimento.

2.3 Caracterização da Comunidade Microbiana por Métodos Independentes de Cultivo

2.3.1 Análise da Diversidade Microbiana

O sequenciamento de DNA pode ser considerado como uma das mais importantes ferramentas de estudo dos sistemas biológicos (RONAGHI, 2001). Este método é cada vez mais empregado para a análise de bibliotecas de fragmentos de DNA obtidos diretamente do ambiente e utilizados para determinação da diversidade e da função de comunidades microbianas. A estratégia metagenômica constitui em uma derivação da genômica microbiana convencional, pois não há necessidade de se obter culturas puras para sequenciamento, revelando, portanto, os genomas contidos na comunidade e não em populações isoladas. Como resultado constitui uma importante ferramenta para determinação de hipóteses a respeito das interrelações dos membros da comunidade.

Além disso, também é possível investigar a maioria dos microrganismos que não podem ser cultivados em laboratório, assim é possível entender as interações e evolução dos ecossistemas microbianos por meio de um mapa mais detalhado. Estas investigações são possíveis atualmente graças à diminuição dos custos do sequenciamento e da evolução da bioinformática, a qual condiciona o processamento de um enorme número de dados (KUNIN et al. 2008).

A metagenômica utilizando a técnica de Sanger aliada à análise de bibliotecas de clones constitui uma importante ferramenta para detecção de microrganismos que caracterizam a comunidade de amostras ambientais. Entretanto, em ecossistemas mais diversificados os resultados das análises podem expressar resultados superficiais. Essa superficialidade reflete o tamanho das bibliotecas de fragmentos amplificados do gene 16S rRNA, onde são normalmente utilizados de 100 a 500 clones, permitindo somente a identificação das populações predominantes (DUNBAR et al., 2002; AXELROOD et al., 2002). Além disso, o tamanho pequeno das amostras gera erros em modelos virtuais, divergindo dos verdadeiros valores reais de índices relacionados à riqueza (AXELROOD et al., 2002; ROESCH et al., 2007). Dessa forma, a utilização de técnicas mais modernas como pirosequenciamento aplicado à investigação da diversidade proporciona uma amostragem muito maior que as abordagens tradicionais, permitindo identificar espécies pouco frequentes, mas que podem ter importância crucial para a manutenção do ambiente.

A aplicação da técnica de pirosequenciamento constitui-se na mais avançada ferramenta para caracterização da diversidade presente em ecossistemas complexos (MARGULIES et al., 2005; YOUSSEF et al., 2009). Paralelamente ao desenvolvimento da técnica, estão surgindo diversos estudos focando a precisão da técnica, onde apesar da limitação em identificar poucas centenas de bases, os estudos procuram investigar as vantagens e implicações da utilização da técnica de pirosequenciamento. Neste contexto, YOUSSEF et al. (2009) por meio de comparações entre bibliotecas de fragmentos do gene 16S rRNA e sequências curtas, simularam regiões obtidas por pirosequenciamento e constaram que não há diferenças significativas entre as regiões V3 e V6 do gene 16S rRNA para distinguir taxonomicamente os grupos bacterianos, sendo possível a utilização de ambas as regiões para a determinação taxonômica sem prejuízo de dados na escolha. Assim, o aumento das regiões sequenciadas por meio do sistema 454 *titanium* que pode chegar até 500 bases sequenciadas, associada à escolha de regiões estratégicas de genes importantes para a descrição taxonômica podem

auxiliar a evitar distorções nos resultados como os encontrados por Sundquist et al. (2007). Neste estudo, os autores observaram inconformidades nas comparações entre bibliotecas de fragmentos do gene 16S rRNA e os dados obtidos por pirosequenciamento, sendo estas diferenças atribuídas às disparidades entre o tamanho dos fragmentos sequenciados que foram de aproximadamente 700pb para os clones e 200pb para os fragmentos gerados por pirosequenciamento (20 system 454 Life Sciences).

Existe uma grande dificuldade em calcular a quantidade de Unidades Operacionais Taxonômicas (UTO's) por grama de solo. Schloss; Handelsman (2006) estimaram entre 2000 e 5000 UTO's por meio de duas bibliotecas provenientes de solos do Alasca e Minnesota. Hong et al. (2006) em estudo também com clones estimaram que no sedimento marinho há entre 2000 e 3000 UTO's, e Torsvik et al. (1998) estimaram que pode existir até 10.000 genomas por grama de solo. Apesar disso, por meio de técnicas mais modernas é possível afirmar que estas estimativas subestimam a verdadeira quantidade de espécies no solo. Roesch et al. (2007) utilizaram o método de pirosequenciamento do gene 16S rRNA aliado à inferências estatísticas para estimar diversidade bacteriana presente em solos do Canadá, Illinois, Flórida e Brasil. Seus resultados mostraram predominância de grupos de Alphaproteobactérias, Bacteriodetes e Betaproteobactérias, enquanto que Archaea representou apenas 0,009% de todas as sequências nos quatro locais avaliados, apesar dos mesmos serem de solos agrícolas, de floresta boreal e urbanos. Por este método os autores identificaram de 26.140 a 53.533 sequências de DNA bacteriano por grama de solo. É estimado que um grama de solo possa conter até 8,3 milhões de espécies de bactérias (GANS; WOILINSKY; DUNBAR 2005; SCHLOSS; HANDELSMAN, 2006).

Adaptando a técnica de pirosequenciamento à investigação da diversidade, Roesch et al. (2007) isolaram DNA do solo e amplificaram parte do gene 16S rRNA usando *primers* específicos para a região V9, esta região é altamente variável. Para a sua amplificação foram realizados poucos ciclos com o objetivo de reduzir possíveis distorções resultantes do processo. Os dados obtidos passam por tratamento estatístico, o qual estabelece o número de UTO's com base em transformação de curvas de rarefação, onde a riqueza de espécies associada às curvas pode ser utilizada como indicador de UTO's, resultando na caracterização da comunidade (KUNIN et al. 2008).

2.3.2 Estimativas da Composição da Comunidade Microbiana

Após as primeiras análises de processamento padrão das sequências (pré-processamento, processamento, montagem e acabamento), é realizada uma reavaliação dos dados metagenômicos, a fim de estimar a composição da comunidade. Esses tratamentos permitem evitar possíveis distorções dos dados de PCR (SUZUKI; GIOVANNONI 1996; VON et al., 1997), identificando grupos com sequências similares, bem como, sua abundância relativa. Com o objetivo de identificar diferentes integrantes de uma comunidade microbiana, são utilizadas sequências de regiões variadas do DNA comum a toda a comunidade alvo, visto que estas sequências funcionam como marcadores, sendo importantes na determinação da abundância. Para fazer estimativas quanto a composição das comunidades são utilizados principalmente marcadores para o gene 16S rRNA (DESANTIS et al. 2006), além do gene 23S rRNA (RNA ribossomal), EF-Tu, EF-G (fatores de alongamento), rpoB (subunidade da RNA polimerase), RecA (proteína de reparo do DNA), e HSP70 (proteína de choque térmico) (MARTÍN et al. 2006; VENTER et al. 2004; VON et al. 2007). Estas sequências são comparadas com as sequências presentes em um banco de dados disponível por meio de alinhamento e realizadas inferências filogenéticas por meio de árvores calculadas.

Entretanto, surgem problemas para determinar a composição da comunidade, pois sua inferência por meio dessas árvores pode ser tendenciosa ou incompleta, uma vez que frequentemente são usados poucos grupos, principalmente Proteobacteria, Firmicutes e Actinobacteria, ou seja, apenas três filos, enquanto há pelo menos 50 filos (HUGENHOLTZ, 2002). Outros problemas surgem do fato de que os genes derivados de séries de dados metagenômicos são provenientes frequentemente de sequências fragmentadas, podendo ocorrer sobreposição de proteínas da mesma família, impedindo uma correta matriz de distância evolutiva e comprometendo a árvore (BROCCHIERI et al., 2001). Além disso, alinhamentos incompletos, e genes de cópia única podem ser sub-representados em bibliotecas de *shotgun* em comparação aos com mais cópias. Também nos casos de clonagem (Sanger) podem não ser representados genes tóxicos para a célula hospedeira (*E. coli*). Porém esta toxicidade varia de acordo com a capacidade da bactéria em transcrever e traduzir o gene introduzido (SOREK et al., 2007). Outro problema ocorre durante a análise dos resultados obtidos, onde a potencial transferência horizontal de genes entre populações simpátricas pode gerar informações conflituosas tanto a respeito do organismo receptor quanto do doador.

Atualmente, o maior problema para a análise está no fato de que o processo é baseado na estimativa da composição da comunidade microbiana, a qual demonstra uma fraca representação da diversidade de isolados sequenciados (HUGENHOLTZ, 2002; JANSSEN et al., 2003), resultando em agrupamentos de organismos filogeneticamente distantes ou ausência de *hits*. Além disso, o método BLAST (o qual é correntemente utilizado para análises) superestima a abundância dos táxons mais cobertos, como Firmicutes e Proteobacteria. O método de MEGAN (HUSON et al., 2007), procura corrigir tal problema, reduzindo falsos resultados por meio da atribuição de fragmentos de sequência para o menor ancestral comum do conjunto de táxons obtidos na comparação. Entretanto, grande parte das sequências pode ser atribuída como pertencentes a altos níveis taxonômicos como, por exemplo, bactéria, constituindo em informação irrelevante.

2.4 Tecnologia de Pirosequenciamento

A necessidade atual pela redução no custo do par de base sequenciada, juntamente com a demanda de um maior número de dados em curto espaço de tempo propiciaram o desenvolvimento de novas técnicas capazes de suprir parcialmente tais exigências. Apesar da grande evolução na técnica de Sanger originalmente proposta há mais de 30 anos, ainda há limitações no custo e velocidade de sequenciamento de DNA, tornando-se cada vez menos compatível às atuais necessidades. A partir do exposto, surgem diferentes técnicas como sequenciamento por hibridização (BAINS; SMITH, 1988;), técnicas de detecção de uma única molécula, ressonância molecular de sequenciamento, enzima-dependente, etc., mas até o presente momento apenas a técnica de pirosequenciamento (RONAGHI et al., 1996) foi realmente empregada em maior escala para a análise da diversidade microbiana de diferentes ambiente. Este fato deve-se à capacidade de montar em um único *chip* centenas de milhares de reações ao mesmo tempo.

As reações de pirosequenciamento ocorrem em uma sequência planejada e são coordenadas pelas enzimas DNA polimerase, Sulfurilase ATP, Luciferase e Apirase. A sequência de nucleotídeos pode ser identificada devido ao fato de que as classes dos mesmos são adicionados e testados individualmente, mediante a polimerização e liberação de um pirofostato inorgânico (PPi), o qual é utilizado pela enzima Sulfurilase ATP para produzir ATP que é usado pela enzima Luciferase para emissão de fluorescência. O sinal e sua intensidade são captados por uma câmara de CCD e

indicam que determinado nucleotídeo do fluxo emitido foi incorporado à cadeia de DNA. Após tal reação a Apirase contida no sistema degrada possíveis dNTPs não utilizados na reação evitando falsos positivos durante as próximas etapas envolvendo outros nucleotídeos (ROCHE. 454 Sequencing, Branford, CT, USA). Estes processos devem ser muito bem coordenados, sendo adicionados e liberados em quantidades molares dos reagentes e enzimas (AGAH et al.2004). A emissão de luz que é detectada e registrada na forma de um pico denominado pirograma, o qual é diretamente proporcional ao número de nucleotídeos incorporados na sequência. Pode-se resumir a tecnologia de Pirosequenciamento como técnica que se baseia no método de sequenciamento por detecção do pirofosfato luminométrico produzido pela incorporação de nucleotídeos na reação (RONAGHI; UHLEN; NYREN 1998). A maior vantagem do pirosequenciamento em relação a Sanger é devido ao fato de não ser necessária a clonagem do fragmento a ser sequenciado, reduzindo uma etapa dispendiosa de tempo e recursos financeiros, uma vez que a técnica tem como vantagens a precisão, flexibilidade, processamento paralelo e pode ser facilmente automatizada. Além disso, há outras vantagens como a não necessidade de primers rotulados, nucleotídeos marcados e eletroforese em gel (RONAGHI, 2001).

Entretanto, uma limitação na técnica de pirosequenciamento é que o sequenciamento só pode ser realizado em sentido unidirecional, diferentemente de Sanger. Além disso, outras críticas quanto o tamanho dos fragmentos sequenciados, os quais chegavam até 500 bases no sistema *Titanium* podem diminuir com o lançamento do sistema *Gold* que pode chegar até 1000 bases (ROCHE Systems, Branford, CT, USA). Este aumento já era esperado, pois existe uma tendência de aumento desse tamanho nos próximos anos a medida que a técnica evolui, fato este já observado antes, pois inicialmente a técnica gerava no máximo apenas sequências com até 100 bases (KUNIN et al. 2008). Atualmente é reconhecível que a técnica de pirosequenciamento pode ser amplamente aplicada, sendo uma alternativa para a caracterização detalhada de DNA (RONAGHI, 2001), bem como de comunidades microbianas de ambientes complexos.

2.5 Dados Metagenômicos

2.5.1 Pré-processamento

Por meio de programas de processamento de dados é possível realizar um pré-processamento dos dados e identificação de populações dominantes a fim de

caracterizar uma comunidade existente. No pré-processamento, dos dados “brutos” que saem das máquinas de sequenciamento são removidas bases de baixa qualidade, além de contaminantes. A identificação das bases de DNA, ou *Basecalling* pode ser realizada por *basecallers* como KB ABI (APPLIED BIOSYSTEMS by life technologies,) e Phred (EWING et al. 1998), todos comportando-se de forma muito similar. O software mais usado é o Phred, no qual são determinados como índices de probabilidade de erros por base, onde o escore de qualidade q atribuído a uma base está relacionado com a probabilidade p estimado erroneamente chamado de base pela seguinte fórmula:

$$q = -10 * \log_{10}(p)$$

Dessa forma, um índice de qualidade Phred de 20 corresponde a uma probabilidade de erro de 1% (EWING et al. 1998). Outro problema a ser corrigido por meio de programas é a contaminação, passível de ocorrer pela presença de sequências do vetor de clonagem com fragmentos de DNA da célula hospedeira (*E. coli*). Apesar de não ser necessária clonagem na técnica de pirosequenciamento, pode ocorrer contaminações e, portanto são necessário cuidados especiais nas análises. Uma estratégia é o BLAST ou ainda a verificação do teor de CG nas sequências, contudo, estes métodos devem ser realizados com muito critério, julgando caso a caso as suspeitas de contaminantes (KUNIN et al. 2008).

O pirosequenciamento de amostras ambientais gera enormes quantidades de dados, necessitando de gerenciamento. A técnica de Sanger pode produzir 700kb de dados de sequência, enquanto, Illumina ao mesmo tempo pode produzir aproximadamente 1 Gb de dados de sequência. Os resultados do processamento dos dados podem ser obtidos pelos pesquisadores sem o uso de super computadores em seus centros de pesquisas, podendo ser utilizados programas disponíveis em computadores de outras instituições que disponibilizam o acesso facilmente em rede (KUNIN et al., 2008).

2.5.2 Ferramentas aplicadas aos dados de pirosequenciamento

A utilização de sequências de genes dos RNAs ribossômicos obtidos de amostras ambientais tornou-se um método padrão para determinar a composição da comunidade bacteriana. Dessa forma, a utilização de novas tecnologias, como pirosequenciamento de amplicons de rRNA mistos permitem uma análise detalhada dessa comunidade de

forma rápida e barata (COLE et al., 2009). A partir disso, surgiram diversas ferramentas para análise de metadados que utilizam o gene 16S rRNA para a caracterização de comunidades bacterianas. As ferramentas mais usadas para este tipo de análise são os projetos SILVA (SILVA, Garching, Germany), Greengenes (DESANTIS et al., 2006), MOTHUR (SCHLOSS, et al., 2009), SIMO (SIMO, Athens, GA, USA) e RDP (RDP Project, East Lansing MI, USA). Dentre estes projetos o RDP (Ribosomal Database Project) possui a interface mais amigável e prática para a utilização de dados metagenômicos. Outro projeto de análise de dados que vem ganhando destaque é o MG-RAST (MEYER et al., 2008), capaz de aceitar dados de 454 e Sanger. Os dados são trabalhados em linguagem *Perl*, realizando o processo em diferentes estágios, onde são realizadas buscas utilizando BlastX (ALTSCHUL et al. 1997) e algoritmos criteriosos de seleção contra diferentes bancos de dados como o International Nucleotide Sequence Database Collaboration (INSDC, <http://insdc.org>) por sequências codificadoras de proteínas (FIGFams), em bancos de sequências de RNA ribossomal como RDP-II (COLE et al., 2008), Green Genes (DESANTIS et al., 2006), European RNA (WUYTS; PERRIÈRE; PEER, 2004). Além disso, são realizadas pesquisas em banco de dados de sequências de cloroplastos, mitocôndria e elementos móveis, Programa ACLAME (Leplae; Lima-Mendez; Toussaint, 2010). Após os dados serem analisados por meio tanto de informações filogenéticas quanto fenéticas, podem ser descritas em heatmaps e pontuações (MEYER, et al., 2008).

2.5.3 Ferramentas RDP - *Pyrosequencing Pipeline*

O RDP (Ribosomal Database Project) possui mais de 33.082 sequências do gene 16S rRNA do domínio Archaea e 643.916 do domínio Bacteria, sendo atualizado mensalmente a partir da colaboração da INSDC (International Nucleotide Sequence Database Collaboration) (BRUNAK et al., 2002). Estas sequências são submetidas a filtros que visam determinar a qualidade (programa *Pintail*), alinhar (*infernai*) e marcar como sequências de alta qualidade ou suspeitas (menos de 10%). Além disso, o site conta com várias ferramentas de análise direcionadas a usuários. Estas ferramentas estão voltadas para análise da qualidade das sequências e investigação da diversidade em processamentos rápidos. A análise de sequências obtidas por pirosequenciamento tornou-se dinâmica devido a diversas ferramentas de análise (*Pyrosequencing pipeline*), capaz de processar grande volume de dados de forma rápida como o alinhamento *infernai*, também disponível aos usuários. Este alinhador utiliza como estratégia a

correlação das estruturas primária e secundária do gene 16S rRNA para obter alinhamentos consistentes com qualidade superior em um menor tempo (COLE et al., 2009). Ainda no *Pyrosequencing pipeline*, são encontradas ferramentas como o *Pipeline Initial Process*, que realizam uma triagem das sequências brutas, retirando sequências e setores com qualidade inferior. Outras ferramentas realizam cluster, análise da riqueza, diversidade, rarefação, classificação taxonômica, além de comparações entre as bibliotecas submetidas e análise de apoio a outros programas de análise como o *Mothur* (SCHLOSS et al., 2009), *SPADE*, *R* (R, Nashville TS, USA) e *Estimate* (ESTIMATE, STORS, CT, USA).

A análise da comunidade bacteriana conta com uma abordagem diferenciada, uma vez que existem diferenças significativas entre as estruturas secundárias conservadas do gene 16S rRNA de bactérias e arqueias, dessa forma, cada conjunto de dados submetido aos filtros e alinhamento deve ser indicado para maximizar o número de posições comparáveis. Além disso, sequências utilizadas como grupo externo são fornecidas em cada alinhamento, sendo *Escherichia coli* para arqueia e *Methanocaldococcus jannaschii* para alinhamentos envolvendo bactérias, possibilitando a reconstrução filogenética. Outra mudança na abordagem filogenética consiste na maneira como a taxonomia bacteriana é tratada, possuindo uma abordagem mais filogenética, onde são utilizados trabalhos taxonômicos e opiniões de grupos que trabalham com microbiologia ambiental. No banco de dados foram reavaliados além de pequenos grupos, também grupos maiores como Firmicutes, Cyanobacteria, Acidobacteria, Verrucomicrobia, OP11, seguindo respectivamente os trabalhos propostos por Ludwig et al. (2008), Wilmotte e Herdman (2001), Barns et al. (2007), Sangwan et al. (2004) e Harris et al. (2004). Estas modificações melhoram o trabalho do classificador, a ferramenta utilizada para classificar taxonomicamente as sequências (COLE et al., 2009).

2.5.4 Alinhamento

Os RNAs ribossomais como 16S rRNA possui em regiões funcionais mais conservadas capazes de formar estruturas secundárias. A utilização de informações sobre essas estruturas secundárias pode determinar o pareamento dessas regiões homólogas em métodos de alinhamento. Dessa forma, o pareamento utilizando essas regiões, enquanto correlaciona a estrutura primária à secundária propicia a realização de alinhamentos com maiores índices de qualidade (COLE et al., 2009). Os alinhamentos

com maior probabilidade de sucesso podem ser realizados de maneira mais rápida, utilizando sequências já caracterizadas em banco de dados. Essas ferramentas de busca fazem uma comparação da sequência linear (ALTSCHUL et al., 1997), onde o alinhamento por similaridade, como o Blast (BLAST, Bethesda, MD, USA), classifica os resultados utilizando pontuações (GAUTHERET; LAMBERT., 2001; ZHANG et al., 2005) e assim melhoram a probabilidade de resultados mais corretos. Neste contexto, programas probabilísticos como SCFGs (*Stochastic Context-Free Grammars*) realizam uma combinação de sequências de tRNA e sua estrutura secundária conservada (DURBIN et al., 1998). Utilizando modelos ocultos de Markov (HMM) de forma mais generalizada é possível a aplicação na modelagem de sequências lineares, realizando a busca de genes, alinhamentos múltiplos, motivos de proteínas, além da busca por sequências mais similares (DURBIN et al., 1998). Assim, a partir da aplicação de uma arquitetura particular do software SCFG, os modelos de covariância (CMs) permitiram a utilização de ferramentas especificamente para a busca de RNAs por similaridade (EDDY; DURBIN 1994). Apesar disso, os softwares baseados em CMs possuem grande complexidade para a análise computacional, impossibilitando sua aplicação prática. Este problema tem sido contornado pelo *Query-Dependent Banding* (QDB), este é capaz de uma rápida busca de RNA por similaridade graças ao seu método que combina propriedades específicas dos CMs que predefinem as regiões independentemente da sequência alvo e em seguida procura um consenso entre as estruturas secundárias. Graças a essas modificações o pacote de softwares Infernal disponível no RDP (RDP PYROSEQUENCING, East Lansing, MI, USA), tornou-se uma ferramenta quatro vezes mais rápida, além de mais sensível e específica para a identificação de sequências homólogas de RNAs estruturais (NAWROCKI; EDDY, 2007). O alinhador *Infernal* fornece vantagens em relação ao RNAcad (utilizado até a versão 9 do RDP). Além disso, o alinhador é 25 vezes mais rápido, oferecendo um tratamento muito mais intuitivo quanto aos possíveis erros presentes nas sequências que causam perda na qualidade e tempo durante o processo de alinhamento. Quando está ausente uma base em um dos lados da hélice, o RNAcad interrompe o alinhamento do outro lado da hélice. Entretanto no alinhador *Infernal* o processo de alinhamento continua mesmo com apenas uma base de um dos lados, sendo este evento penalizado pelo algoritmo (COLE et al., 2009).

3 OBJETIVOS

3.1 Objetivo Geral

Tendo em vista que este trabalho está associado ao projeto “Diversidade genética de microrganismos associados às plantas carnívoras terrestres e aquáticas do gênero *Utricularia* (Lentibulariaceae)” (processo FAPESP 07/58277-7), o objetivo geral foi avaliar por meio do sequenciamento em larga escala a diversidade e estrutura da comunidade bacteriana associada aos utrículos de *Utricularia gibba* (Lentibulariaceae) e duas lagoas onde são encontradas espécimes desta planta carnívora.

3.2 Objetivos Específicos

Para alcançar o objetivo geral, os objetivos específicos que compõem este trabalho foram:

- a) obter sequência parcial do gene 16S rRNA de bactérias associadas à *Utricularia gibba* e à água de duas lagoas da Várzea do Rio Tietê em Mogi das Cruzes;
- b) identificar as comunidades bacterianas por meio de pirosequenciamento do gene 16S rRNA;
- c) estimar a diversidade e a riqueza de bactérias associadas à *Utricularia gibba* e às duas áreas de várzea amostradas, bem como o padrão de dominância;
- d) correlacionar a diversidade bacteriana observada na água da várzea com aquela presente em *Utricularia gibba*.

4 MATERIAIS E MÉTODOS

4.1 Área de coleta e Extração de DNA total bacteriano

As amostras de água foram obtidas de dois trechos da APA da várzea do rio Tietê. Os locais compreendem a estrada Mogi-Salesópolis-SP (23°33'28.85"S , 46° 8'12.65"O) (local 1) e ao bairro de César de Souza, Mogi das Cruzes-SP (23°32'4.97"S, 46° 8'42.47"O) (local 2). As amostras de plantas de *U. gibba* também foram obtidas no trecho de várzea próximo à estrada Mogi-Salesópolis-SP (local 1), onde foram coletadas plantas de 4 pontos específicos e não contínuos (lagoas diferentes). Este fato foi necessário, para evitar que a mesma planta fosse coletada em diferentes pontos, visto que *U. gibba* apresenta crescimento por meio de estolões. Na coleta de água, para cada ponto, foram coletados de 50mL de água a uma profundidade de aproximadamente 30 cm e a uma distância de pelo menos um metro da margem, levando em conta a posição das plantas. Logo após coletadas, as amostras de água foram levadas ao laboratório, onde em recipientes estéreis, as amostras de cada local foram misturadas, e filtradas em membrana com poros de 0,22µm (Millipore® - Billerica, NY, USA). Foram utilizadas 4 membranas para cada local (Locais 1 e 2), totalizando assim 200mL de água por local. Após a filtração, cada membrana foi separadamente recortada e os fragmentos submetidos à extração do DNA total das bactérias utilizando o kit “UltraClean Soil DNA Isolation Kit®” (MoBio Labs, Inc. Solana Beach, USA) seguindo as recomendações do fabricante. Para a extração do DNA total presente nos utrículos foram utilizados 0,5g de utrículos jovens e sadios de cada planta, totalizando quatro extrações, onde foi utilizado o mesmo kit usado para extração do DNA total presente na água. Os resultados das extrações foram avaliados em gel de agarose (1%).

4.2 Amplificação do gene 16S rRNA de bactérias

O gene 16S rRNA foi amplificado utilizando os primers forward BAC4A MID7 para o local 1, BAC4A MID8 para o local 2 e BAC4A MID9 para *U. gibba* (tabela 1). O primer reverse BAC5rB foi utilizado para todas bibliotecas. Estes primers são específicos para o domínio Bacteria e compreendem a região 968-1401 do gene, além disso, eles possuem sequências *barcode* necessárias no preparo das amostras para pirosequenciamento. A PCR foi realizada em um volume final de 50 µL contendo tampão da enzima 1 X, 5 mM de MgCl₂, 10 mM de dNTPs, 0,1 µM de cada primer, 0,1U/µL de Taq DNA Polimerase. A reação da amplificação foi realizada em

termociclador, programado para realizar uma desnaturação inicial a 94 °C por 4 minutos, 25 ciclos de desnaturação a 94 °C por 30 segundos, anelamento a 62,5 °C por 1 minuto e extensão a 72 °C por 1 minuto, seguida de extensão final a 72 °C por 10 minutos. Após a amplificação, a reação de PCR foi avaliada por eletroforese em gel de agarose (1%).

Tabela 1. Primers adaptados capazes de flanquear a região 968-1401 do gene 16S rRNA do domínio bactéria.

Primer	Sequência (5' - 3')	Sentido
BAC4AMID7	CGTATCGCCTCCCTCGCGCCATCAGCGTGTCTCTAAACGCGAAGAACCTTAC	Forward
BAC4AMID8	CGTATCGCCTCCCTCGCGCCATCAGCTCGCGTGTCAACGCGAAGAACCTTAC	Forward
BAC4AMID9	CGTATCGCCTCCCTCGCGCCATCAGTAGTATCAGCAACGCGAAGAACCTTAC	Forward
BAC5rB	CTATGCGCCTTGCCAGCCCGCTCAGCGGTGTGTACAAGGCCCGGGAACG	reverse

* Sequências sublinhadas compreendem ao *barcode* e sequências em negrito correspondem aos primers específicos para o domínio bactéria.

Fonte: (FERREIRA A. J., 2011).

4.3 Purificação, quantificação e pirosequenciamento do gene 16S rRNA

O produto de PCR foi purificado utilizando o Kit QIAquick PCR purification (QIAGEN GmbH, Hilden, Germany), seguindo as instruções do fabricante. Após a purificação, a pureza e quantidade de *amplicons* foi verificada em gel de agarose (1,5%) e espectrofotometria (NanoDrop® - Wilmington, DE, USA). Os produtos de cada uma das quatro reações de PCR resultantes de cada tratamento (local 1, local 2 e *U. gibba*) foram misturados proporcionalmente de acordo com a quantificação e enviados para Sequencing Center - Roche Applied Science (Branford, CT, USA), onde foram sequenciados pelo método de pirosequenciamento.

4.4 Análise e Processamento de Dados

4.4.1 Filtragem de sequências de acordo com a qualidade

Primeiramente foi necessário uma seleção das sequências quanto à qualidade das bases e o tamanho, para isso foi utilizada a ferramenta “*Pipeline Initial Process*” disponível no Pipeline do RDP. Assim, foram utilizados como parâmetros de seleção o arquivo de qualidade de bases, o tamanho dos fragmentos, a identificação das sequências dos primers e a presença de bases não identificadas. Com o objetivo de obter um conjunto de sequências com qualidade satisfatória para as análises, somente sequências com bases de pontuação superior a 20 foram utilizadas, ou seja, bases com a

probabilidade máxima 1% de erro (EWING et al., 1998). Também não foram utilizadas sequências com bases não identificadas e com tamanho inferior a 380 pares de bases. Houve uma tolerância de 20% de distância para a similaridade com os primers (sem as sequências adaptadoras), sendo os mesmos mantidos ao final do processo, pois as sequências pertenciam ao domínio Bactéria. Assim, fora da região flanqueada pelos primers foram cortadas as bases dos adaptadores, eliminando o perigo de sequências truncadas que reduzem a qualidade e interferem na classificação das sequências (KUNIN et al. 2008).

4.4.2 Estimativa de riqueza e diversidade

O alinhamento e o cluster foram gerados utilizando as ferramentas “*Aligner*” e “*Complete Linkage Clustering*” respectivamente, ambas disponíveis no RDP (Ribosomal Database Project). Ainda no RDP foram construídas curvas de rarefação e os índices de diversidade (Shannon-Winner) e riqueza (Chao-1). Além disso, foi construída uma matriz triangular, necessária para o programa *Mothur* (SCHLOSS, et al., 2009). Neste programa foram calculadas as curvas de rarefação, diagramas de Venn, *Heatmaps*, os índices de diversidade Shannon-Winner e Simpson, além dos estimadores de riqueza Ace e Chao-1, sendo possível comparar estes resultados aos obtidos do RDP.

4.4.3 Classificação das sequências

A classificação das sequências foi realizada por meio do programa RDPquery, disponível no SIMO (Sapelo Island Microbial Observatory). O SIMO faz a análise de similaridade com base no banco de dados do RDP, onde existem apenas sequências do gene 16S rRNA testadas e avaliadas quanto à qualidade, além da identificação dos isolados microbianos. Os parâmetros de busca foram baseados na similaridade, onde quanto maior a similaridade, menor foi o nível taxonômico identificado. A análise de sequências de similaridade foi realizada de acordo com as recomendações do sistema (Domínio >0%, Filo >75%, Classe >85%, Ordem > 91%, Família >92%, Gênero >95%, Espécie > 97% e linhagem =100%) (JANGID et al. 2008; STACKEBRANDT; GOEBEL, 1994).

4.4.4 Cálculo das diferenças entre as bibliotecas

As diferenças entre os grupos taxonômicos dos tratamentos foram calculadas utilizando a ferramenta *Libcompare* presente no *pipeline* do RDP.

5 RESULTADOS

5.1 Preparo das amostras

Por meio do método utilizado foi possível extrair quantidade suficiente de DNA das amostras ambientais, o qual foi utilizado em PCR com primers específicos para a região 968-1401 o gene 16S rRNA.

5.2 Análise e seleção qualitativa das sequências

O pirosequenciamento gerou 71.042 sequências, entretanto, nem todas apresentaram qualidade e tamanho necessários para serem empregados nas análises taxonômicas e ecológicas, sendo utilizadas após os métodos de filtragem aproximadamente 41.404 sequências (58,28%). A seleção de sequências quanto à qualidade demonstrou uma boa eficiência (figura 1), onde inicialmente os arquivos “crus” de *U. gibba* e dos locais 1 e 2 contendo 24.513, 25.766 e 20.763 sequências respectivamente, passaram a 14.020, 14.520 e 12.371 após os cortes dos filtros. As descartadas não possuíam critérios mínimos como pontuação do Phred inferior a 20, tamanho do *read* inferior a 380 bases, sequências com menos de 80% de similaridade aos primers, além de pelo menos uma base não identificada. A maior causa para a reprovação das *reads* foi a ausência de sequências similares aos primers, onde aproximadamente 40,6% (28.871) foram descartadas mesmo com uma tolerância de 80% na similaridade. O filtro seguinte removeu 1.175 (1,6%) sequências que continham pelo menos uma base não identificada (“N”). Em seguida, foram reprovadas 200 (0,28%) sequências com menos de 380 bases de comprimento.

Figura 1. Histograma de seleção de sequências

No histograma pode ser observado o resultado da seleção de sequências com qualidade superior para ser utilizada nas análises. Em A observa-se o 1 com as amostras de 432 bases em média e em C histograma do local 2 com comprimento médio de 431 bases.

Fonte: (FERREIRA A. J., 2011).

O processo de seleção de sequências levando em conta a qualidade das sequências obtidas a partir das bactérias associadas à *U. gibba* (figura 1A) garantiu a retirada de aproximadamente 47% das sequências desse tratamento, sendo 37% (9.993) das sequências cortadas da análise devido à ausência das sequências correspondentes aos primers, enquanto 1,8% (479) foram excluídas devido a presença de bases não identificadas e apenas 0,08% (21) foram desconsideradas da análise devido ao seu tamanho inferior a 380 bases. Como resultado, as 14.020 sequências de qualidade satisfatória resultantes do procedimento continham em média 432 bases. As sequências aprovadas do local 1 (figura 1B) (estrada Mogi-Salesópolis-SP) também apresentaram um comprimento médio de 432 bases. Ainda no tratamento local 1, aproximadamente 42% (10.791) das sequências foram removidas devido à ausência de correspondência com os primers. Em menor proporção, 1,6.% (432) das leituras foram removidas devido à presença de bases não identificadas e 0,09% não continham pelo menos 380 bases de comprimento. Já o tratamento local 2 (figura 1C) (bairro Cezar de Souza – Mogi das Cruzes-SP), sofreu uma menor quantidade de cortes para a correspondência de primers (38% - 7.967) e presença de bases não identificadas (1,2% - 264). Apenas o corte de sequências com comprimento inferior a 380 pares de bases foi maior (0,77% - 161).

A observação da proporção decrescente de cortes devido aos filtros de correspondência dos primers, presença de bases não identificadas e comprimento mínimo do fragmento são normais, uma vez que no processo de seleção estes filtros são empregados nesta ordem, assim é natural que sempre ocorra mais exclusões de sequências que não possuam sequência similar a dos primers.

5.3 Alinhamento

O alinhamento realizado com as sequências obtidas pelo filtro indicou o início na posição de base número 968 (*Escherichia coli* J01695 como referência), corroborando com a sequência do primer utilizado. Já a posição final dos fragmentos variou em aproximadamente 68,78% das sequências na base número 1.401 (figura 2), sendo o comprimento médio de 411 bases. O tamanho desses fragmentos pode variar de acordo com os microrganismos devido à taxa evolutiva do mesmo.

Figura 2. Histograma de tamanho dos fragmentos.

O histograma obtido do RDP indicando o tamanho dos fragmentos através das posições das bases de início e fim no alinhamento. Em A, está apresentado o alinhamento do tratamento *U. gibba*, em B o tratamento Local 1 e em C o tratamento Local 2 (escala baseada em *E. coli*).

Fonte: (FERREIRA A. J., 2011).

5.4 Construção de *Clusters*

A construção dos *Clusters* realizada com a ferramenta “*Complete Linkage Clustering*” do RDP, utilizou os alinhamentos do tratamento *U. gibba* e dos locais 1 e 2 (figura 3). Foi observado um grande número de agrupamentos na análise, chegando a 28.688 grupos com 100% de similaridade e 1.404 grupos com 91% de similaridade. O maior intervalo no número de grupos ocorreu entre 100% e 99%, podendo indicar uma ampla variedade dentro das espécies bacterianas.

Figura 3. de *Cluster*.

Número de UTO's formadas de acordo com a similaridade. A análise seguiu como parâmetros os valores de similaridade de >91% para ordem; >92% família; >95% gênero; >97% espécie e =100% (único) para linhagem.

Fonte: (FERREIRA A. J., 2011).

A construção do arquivo de *Cluster* é essencial para as análises, pois a partir dele é possível realizar outras análises como a construção de curvas de rarefação, estimadores de riqueza, índices de diversidade, diagramas de Venn e *Heatmaps*.

5.5 Rarefação

As curvas de rarefação construídas tanto pelo RDP (dados não mostrados) quanto pelo *Mothur* apresentaram tendência muito similar. Isto pode ser considerado natural, uma vez que ambas as análises utilizaram o mesmo arquivo de *cluster*. Também foi necessária uma pequena adaptação nos resultados, pois o programa *Mothur* é mais flexível, considerando duas ou mais sequências como 100% similares a partir de 99,9951%. Diferentemente, a ferramenta *pipeline* do *RDP II Project* é mais rigorosa, considerando 100% como nível absoluto. Desta forma, para equiparar as duas análises o nível único, considerado pelo programa *Mothur* para diferenciar grupos com distância igual a zero foi adotado como 100% e chamado neste trabalho como “100% (único)”.

A amostragem das curvas de rarefação não indica estabilização de sequências ao nível de 100% (únicas) de similaridade, ilustrando grande diversidade de linhagens bacterianas (figura 4A). Entretanto, a tendência das curvas abaixo desses índices apontaram para uma estabilização próxima, resultado de uma boa amostragem ainda no nível de espécie para o local 2. A análise das sequências indicou uma maior riqueza no local 1 (estrada Mogi-Salesópolis-SP), seguida por *U. gibba* e por último o local 2 (bairro Cezar de Souza – Mogi das Cruzes-SP) em todos os níveis (figura 2A e B). Em todos os tratamentos no nível de similaridade 100% (único) (linhagem) as curvas mantiveram-se muito acentuadas. Esta tendência modifica-se bruscamente ao nível de 97% (espécie), onde começa a estabilização, principalmente no local 2 e em 95% (gênero) no local 1 e *U. gibba* (Figura 4B). O local 1 e *U. gibba* só alcançam a estabilização a 91% (ordem) e o local 2 a 95%, indicando uma alta diversidade nos três tratamentos.

Figura 4. Curvas de rarefação.

Construídas a partir de análises realizadas no programa Mothur, as curvas indicam o efeito do esforço no sequenciamento para estimar o número de UTO's. À esquerda, estão as curvas construídas ao nível de similaridade de 100% (únicas) e 97%. À direita, estão as curvas as construídas ao nível de similaridade de 95% e 91% para os três tratamentos. Os níveis de 100% (único), 95, 92 e 91%, correspondem aos níveis taxonômicos de linhagem, espécie, gênero, família e ordem respectivamente.

Fonte: (FERREIRA A. J., 2011).

5.6 Distribuição de UTO's entre os locais

A análise de distribuição das 41.404 sequências obtidas após a filtragem gerou até 28.688 UTOs (Unidades Taxonômicas Operacionais) com 100% (único) quando observados os 3 tratamentos juntos (tabela 2). A observação dos tratamentos separadamente revelou que o local 1 gerou mais UTO's, chegando a 10.450 UTOs, seguido de *U. gibba* (10.414) e o local 2 (8.541). Tal tendência no número de UTOs foi mantida nos outros níveis de similaridade observados (97, 95 e 91%).

Tabela 2. Distribuição de UTOs de acordo com a similaridade e tratamento.

Similaridade (%)	Nível Taxômico	<i>U. gibba</i>	Local 1	Local 2	Total dos tratamentos
100	Linhagem	10.414	10.450	8.541	28.688
97	Espécie	1.934	2.401	1.089	3.547
95	Gênero	1.356	1.822	882	2.752
91	Ordem	797	1.040	591	1.404

Fonte: (FERREIRA A. J., 2011).

A análise revelou que apenas 0,17% das UTO's delimitadas ao nível 100% (único) são compartilhadas pelos três tratamentos (figura 5A). Neste nível, o local 1 apresentou o maior número de UTO's, seguido pelo tratamento *U. gibba* e o menor número de UTO's presentes no local 2. De acordo com a análise por curvas de rarefação esta proporção de UTO's está diretamente ligada ao número amostral de sequências, o que é natural, uma vez que neste nível as curvas não apresentaram

estabilização, indicando que a amostragem não foi suficiente. Neste nível de similaridade nota-se um compartilhamento maior das UTO's do local 1 e 2, seguido pelo local 1 e *U. gibba*. Entretanto, esta tendência não é mantida, pois quando comparadas as sequências agrupadas com 97% de similaridade (figura 5B) entre si, foi observado que os tratamentos Local 1 e *U. gibba* apresentam compartilhamento maior que *U. gibba* e o Local 2. A análise a 95 e 91% de similaridade (figura 5C e D) também acompanharam esta tendência. Dessa forma, os resultados da análise de UTO's nos diagramas de Venn demonstram que os ocorre um maior compartilhamento entre *U. gibba* e o Local 1, indicando que há uma maior influência do local sobre a comunidade associada aos utrículos.

Figura 5. Diagramas de Venn.

UTOs (Unidades Taxonômicas Operacionais) de acordo com os níveis de similaridade. Em A, UTOs agrupadas por critério em 100% (único); B, 97%, C, 95% e D, 91% de similaridade. Fonte: (FERREIRA A. J., 2011).

A análise por *Heatmap* foi utilizada para caracterizar melhor a quantidade de sequências presentes em cada UTO (figura 6), podendo realizar assim inferências quanto à densidade de determinados grupos bacterianos em comparação com as diferentes amostras. Linhas escuras (pretas) mostram um número pequeno de sequências, enquanto linhas vermelhas mostram um grande número de sequências. Sequências similares são observadas na mesma altura em tratamentos diferentes. Assim nas 50 UTO's compartilhadas com 100% (único) de similaridade foi observado que apesar do número grande de UTOs, existe um número pequeno de sequências representantes de cada UTOs, não sendo observadas linhas vermelhas para este nível de similaridade. A comparação entre os locais 1 e 2, mostrou uma UTO presente em maior densidade (linha vermelha), e compartilhada entre estes locais. Esta mesma UTO também esteve presente na amostra proveniente de utrículos de *U. gibba*, mas com intensidade menor (coloração preta), mostrando menor quantidade de sequências nesta UTO. O mesmo fenômeno foi observado em UTOs obtidas a partir de todos os níveis de similaridade (figura 6A, B, C e D). Assim, a análise a 97% indicou que nas 267 UTO's (4,7%) compartilhadas há pelo menos uma espécie com o mesmo padrão de dominância presente nos locais 1 e 2, enquanto em *U. gibba* a mesma UTO possui uma menor densidade de sequências, este padrão é mantido nos níveis de gênero (95%) e ordem (91%), indicando que a comunidade dominante em *U. gibba* apesar de presente também na água, não obtém o mesmo sucesso na colonização do ambiente externo.

Figura 6. Análise por *Heatmap*.UTO's de acordo com os níveis de similaridade

Linhas coloridas representam as UTO's e a gradação de cores indica a densidade de sequências nas UTO's. Em A, UTO's agrupadas por critério em 100% (único); B, 97%, C, 95% e D, 91% de similaridade.

Fonte: (FERREIRA A. J., 2011).

5.7 Riqueza e diversidade

A riqueza e a diversidade analisadas pelo *Mothur* e pela ferramenta de análise no *pipeline* do RDP apresentaram índices similares (tabela 3). A riqueza do local 1 apresentou maiores índices em relação ao local 2, sendo todos desviados significativamente de acordo com a faixa de similaridade adotada. Entretanto, os maiores valores foram alcançados para as amostras de *U. gibba*, mas ainda similares com o local 1. No nível de 100% o valor máximo de riqueza alcançado para *U. gibba* foi de 99.102 em Chao-1 e 260.111 em Ace. O segundo maior valor foi estimado no local 1 em Chao-1 (88.366) e Ace (223.111), enquanto o local 2 apresentou os menores valores de riqueza (Chao-1 – 68.011 e Ace – 152.523). Acompanhando a tendência, o local 2 apresentou os menores níveis para a faixa de similaridade de 91% (Ordem) em ambos os estimadores (Chao-1 - 834 e Ace – 917).

Tabela 3. Estimadores de riqueza Chao-1 e Ace aplicados de acordo com o nível taxonômico inferido para os valores de agrupamento por similaridade das bibliotecas analisadas.

Biblioteca	Similaridade	# sequências	# UTO	Nível taxonômico	Chao-1	Ace
Local 1	100% (único)		10.450	Linhagem	88.366 (81.638-95.730)	223.111 (216.020-230.447)
	97%	14.520	2.401	Espécie	4.505 (4.212-4.845)	6.230 (6.070-6.397)
	95%		1.822	Gênero	3.034 (2.836-3.270)	3.857 (3.746-3.975)
	91%		1.040	Ordem	1.447 (1.351-1.572)	1.446 (1.394-1.505)
Local 2	100% (único)		8.541	Linhagem	68.011 (62.448-74.148)	152.523 (147.441-157.790)
	97%	12.371	1.089	Espécie	2.383 (2.127-2.703)	4.068 (3.870-4.281)
	95%		882	Gênero	1.797 (1.596-2.055)	2.824 (2.672-2.988)
	91%		591	Ordem	834 (766-927)	917 (894-941)
<i>U. gibba</i>	100% (único)		10.414	Linhagem	99.102 (91.012-108.005)	260.111 (252.900-267.537)
	97%	14.020	1.934	Espécie	3.356 (3.125-3.632)	4.306 (4.187-4.431)
	95%		1.416	Gênero	2.299 (2.127-2.511)	2.770 (2.684-2.862)
	91%		797	Ordem	1.127 (1.037-1.250)	1.097 (1.060-1.140)

Fonte: (FERREIRA A. J., 2011)

Os resultados gerados pelos índices de Shannon-Winner e Simpson (tabela 4) apresentaram uma diferença menor entre os dois locais amostrados. Entretanto, o local 1 também possui uma maior diversidade em relação ao local 2. Em contraste, os índices de diversidade de *U. gibba* foram semelhantes aos do local 1, não divergindo significativamente em nenhum dos níveis taxonômicos (espécie, gênero e ordem) apresentados. Além disso, os índices de diversidade no local 2 em relação à planta e ao local 1 foram menores, mostrando diferenças significativas. Embora, o índice de Shannon-Winner tenha indicado uma diversidade com dimensão semelhante entre a comunidade do local 1 e *U. gibba*, a diversidade calculada pelo índice de Simpson apresentou diferenças significativas entre os tratamentos. Os valores para a diversidade foram maiores em *U. gibba* em todas as faixas taxonômicas analisadas.

Tabela 4. Índices de diversidade Shannon-Winner e Simpson aplicados de acordo com o nível taxonômico inferido para os valores de agrupamento por similaridade das bibliotecas analisadas.

Biblioteca	Similaridade	# sequências	# UTO	Nível taxonômico	Shannon (H')	Simpson
Local 1	100% (único)	14.520	10.450	Linhagem	8,42 (8,9-8,94)	0,002 (0,002-0,002)
	97%		2.401	Espécie	5,00 (5,18-5,82)	0,02 (0,02-0,021)
	95%		1.822	Gênero	5,93 (5,12-5,73)	0,026 (0,025-0,027)
	91%		1.040	Ordem	4,43 (4,39-4,46)	0,044 (0,042-0,045)
Local 2	100% (único)	12.371	8.541	Linhagem	8,33 (8,10-8,56)	0,006 (0,005-0,006)
	97%		1.089	Espécie	3,64 (3,5-3,84)	0,083 (0,08-0,086)
	95%		882	Gênero	3,38 (3,05-3,72)	0,091 (0,088-0,094)
	91%		591	Ordem	3,58 (3,40-3,76)	0,1 (0,097-0,103)
<i>U. gibba</i>	100% (único)	14.020	10.414	Linhagem	8,45 (8,29-8,62)	0,001 (0,001-0,001)
	97%		1.934	Espécie	5,39 (5,29-5,48)	0,013 (0,013-0,014)
	95%		1.416	Gênero	5,74 (5,59-5,89)	0,018 (0,017-0,019)
	91%		797	Ordem	4,59 (4,30-4,89)	0,022 (0,021-0,023)

(FERREIRA A. J., 2011).

5.8 Identificação das sequências

A análise de Filos bacterianos via *RDPquery* demonstrou que as populações predominantes das duas bibliotecas de água possuem grande similaridade entre si. Enquanto que a comunidade de *U. gibba* apresentou maior similaridade com as sequências do local 1, corroborando com os dados da análise por UTO's (figura 7). Os grupos com maior representação nos tratamentos foram Proteobacteria, Actinobacteria, Verrucomicrobia, Firmicutes, Cyanobacteria, Acidobacteria e Chlamydiae. Os grupos com menor frequência foram Nitrospira, Bacteroidetes, Chlorofexi, Deferribacteres, Gemmatimonadetes, Aquificae, Fusobacteria, Thermodesulfobacteria, Deinococcus-Thermus, Spirochaetes, Synergistetes, Tenericutes, Chlorobi, Planctomycetes, Bacteria_incertae_sedis e Lentisphaerae. Desses filos, foi observado um maior número de sequências similares as do filo Proteobacteria nos três tratamentos amostrados. Essas sequências correspondem a aproximadamente 54% do total analisado, sendo maior no local 1 (58%), seguido de *U. gibba* (57%) em relação ao local 2 (45%), o qual, segundo a análise da ferramenta *Libcompare* possui diferença significativa em relação aos outros tratamentos (tabela 5). O segundo maior grupo é representado por Actinobacteria (20%). Entretanto, a proporção em relação aos locais se altera, sendo menor no local 1 (21%) e *U. gibba* (20%) em relação ao local 2 (39%), tendo este quase o dobro e também significância nas diferenças. Também foi possível observar uma diferença expressiva no filo Chlamydiae, no qual o local 1 apresentou 3,19% das sequências com maior similaridade, enquanto no local 2 e *U. gibba* foi de apenas 0,5% e 0,6% respectivamente. Essas diferenças podem ocorrer em função das diferenças ambientais,

onde determinados grupos de Proteobacteria e Chlamydiae podem ter seus nichos ocupados por Actinobacteria ou vice-versa. Outros grupos que apresentaram diferenças significativas dentre os tratamentos foram Acidobacteria, Firmicutes e Cyanobacteria, pois estes filos apresentaram maior proporção na planta em relação ao meio aquático.

A quantidade de sequências resgatadas da planta e similares à Acidobacteria foi mais de dez vezes maior que as da água, indicando este como um forte nicho para este grupo. Já Firmicutes e Cyanobacteria apresentaram respectivamente em média proporções de 3 a 4 vezes maiores em relação ao ambiente externo. Também foram observados grupos com proporção similar entre a comunidade das plantas e a da água, sendo os mesmos pertencentes aos filos, Bacteria_incertae_sedis, Lentisphaerae, Aquificae, Gemmatimonadetes, Deferribacteres, Thermodesulfobacteria e Deinococcus-Thermus. Alguns filos como Tenericutes e Chlorobi somente foram encontrados local 1, enquanto Spirochaetes e Synergistetes foram encontrados no local 1 e *U. gibba*. Por outro lado, Fusobacteria foi encontrada exclusivamente no local 2. No entanto, a não observação dessas sequências nos dois locais ou na planta pode ter sido devido a limitações no método, uma vez que nesses casos menos de uma dezena de sequências foram encontradas nos tratamentos.

Figura 7. Percentual de Filos bacterianos.

Em destaque, filos presentes com maior percentual nos tratamentos
 Fonte: (FERREIRA A. J., 2011).

Tabela 5. Significância das diferenças encontradas na comparação dos filos.

Filo	Local 1 X Local2	Local 1 X <i>U. gibba</i>	Local2 X <i>U. gibba</i>
Acidobacteria	6,74 ⁻⁰¹	6,00^{-14*}	6,00^{-14*}
Actinobacteria	6,00^{-14*}	6,00^{-14*}	6,00^{-14*}
Bacteria_incertae_sedis	3,15 ⁻⁰¹	4,25 ⁻⁰¹	8,46 ⁻⁰²
Bacteroidetes	6,00^{-14*}	5,74^{-07*}	6,00^{-14*}
BRC1	-	2,37 ⁻⁰¹	3,00 ⁻⁰¹
Chlamydiae	6,00^{-14*}	6,00^{-14*}	1,74 ⁻⁰¹
Chlorobi	9,18 ⁻⁰²	6,82 ⁻⁰²	-
Chloroflexi	1,80^{-04*}	6,00^{-14*}	6,00^{-14*}
Cyanobacteria	2,67 ⁻⁰¹	6,00^{-14*}	6,00^{-14*}
Deinococcus-Thermus	5,83 ⁻⁰¹	5,18 ⁻⁰¹	-
Firmicutes	2,99^{-10*}	4,80^{-04*}	6,00^{-14*}
Fusobacteria	1,95 ⁻⁰¹	-	2,06 ⁻⁰¹
Gemmatimonadetes	7,80^{-03*}	1,24 ⁻⁰¹	4,28^{-05*}
Lentisphaerae	4,90 ⁻⁰¹	7,83 ⁻⁰²	1,66 ⁻⁰¹
Nitrospira	1,78 ⁻⁰¹	3,47 ⁻⁰²	1,02^{-03*}
OP10	5,83 ⁻⁰¹	9,74 ⁻⁰¹	5,64 ⁻⁰¹
OP11	8,80 ⁻⁰¹	5,18 ⁻⁰¹	4,39 ⁻⁰¹
Planctomycetes	5,94 ⁻⁰¹	2,25^{-05*}	2,38^{-06*}
Proteobacteria	6,00^{-14*}	3,47 ⁻⁰¹	6,00^{-14*}
Tenericutes	1,70 ⁻⁰¹	1,34 ⁻⁰¹	-
TM7	6,00^{-14*}	5,24^{-12*}	2,88^{-03*}
Verrucomicrobia	6,00^{-14*}	6,00^{-14*}	2,93 ⁻⁰²
WS3	2,68 ⁻⁰²	2,88^{-03*}	1,81^{-06*}

* valores significativos, onde $p < 0,01$

Fonte: (FERREIRA A. J., 2011)

A análise do grupo Proteobacteria demonstrou predominância da classe Betaproteobacteria em todos os tratamentos, chegando a 52% das sequências obtidas (figura 8). Entretanto, ao comparar os tratamentos separadamente, foram observadas diferenças entre todos, onde a classe Betaproteobacteria possui um percentual maior no local 1 (62%) em relação ao local 2 (41%) e *U. gibba* (28%), sendo todas as diferenças consideradas significativas (tabela 6). Alphaproteobacteria mostrou-se predominante em *U. gibba* (33%) no Local 2 (48%) ($p < 0.01$). Este resultado contrasta com a população de Alphaproteobacteria do local 1, onde apenas 12% das sequências alinharam com sequências deste filo. Também foi possível verificar a presença dos grupos de Deltaproteobacteria, Gammaproteobacteria e Epsilonproteobacteria. Ao analisar a proporção destes grupos entre os tratamentos, pode-se notar uma colonização similar nos três ambientes, onde a proporção da população de Gammaproteobacteria e Deltaproteobacteria são maiores ($p < 0.01$) no local 2, seguida por *U. gibba* e o Local 1 ($p < 0.01$). Epsilonproteobacteria ($p > 0.01$) foi encontrado apenas no local 1, este grupo geralmente é considerado raro, não sendo normalmente encontrado em trabalhos envolvendo diversidade bacteriana de água doce.

Figura 8. Classes de Proteobacteria.

As classes foram responsáveis por 54% da comunidade bacteriana total analisada.

Fonte: (FERREIRA A. J., 2011).

O Filo Firmicutes foi representado principalmente pelas classes Bacilli (42%) e Clostridia (41%) (figura 9). Clostridia foi observada proporcionalmente nos locais 1 (64%) e 2 (59%), sendo significativamente menor em *U. gibba* (25%). A classe Bacilli foi observada em maior percentual em *U. gibba* (46%) do que em relação aos locais 1 (33%) e 2 (39%). A classe Thermolithobacteria apresentou grande percentual em *U. gibba* (27%), estando em uma frequência menor que 2% nas amostras de água. Já Erysipelotrichi só foi encontrada no local 1, podendo ser considerada rara neste ambiente, uma vez que apenas uma sequência teve similaridade com a classe.

Figura 9. Classes de Firmicutes.

As Classes foram responsáveis por 8% da comunidade bacteriana total analisada.

Fonte: (FERREIRA A. J., 2011).

Apesar de Acidobacteria representar apenas pouco mais de 4% do total de sequências analisadas, ao observar sua distribuição como classe, foi observado que ela está restrita a poucos grupos. Enquanto os filos Proteobacteria, Firmicutes e Actinobacteria estão subdivididos em dezenas de grupos, a população de Acidobacteria está concentrada principalmente na classe Acidobacteria_Gp1 (figura 10) com mais de 80% das sequências do filo. O número de sequências similares à Acidobacteria foram significativamente maiores em *U. gibba* com as classes Acidobacteria_Gp1, Acidobacteria_Gp3, Acidobacteria_Gp6 e Holophagae. Acidobacteria_Gp1 teve a distribuição mais heterogênea, contribuindo com mais de 72% em *U. gibba*, enquanto o local 1 contribuiu com 3,5% e o local 2 com 4,8%. Acidobacteria_Gp3 representou 5,3% em *U. gibba*, 0,7% no local 1 e 0,4% no local 2. A comunidade de Acidobacteria_Gp6 foi ligeiramente maior em *U. gibba*, com 3,8% do que nos locais 1 (2,2%) e no Local 2 (1%). Holophagae também apresentou o mesmo padrão, tendo a sua frequência de 1,8%, 1,2% e 1,3% para *U. gibba*, local 1 e local 2. Apenas no local 1 foi detectada uma sequência similar à Acidobacteria_Gp5, enquanto Acidobacteria_Gp7 foi detectada apenas no local 1 e *U. gibba*, sendo uma sequência resgatada de cada tratamento.

Figura 10. Classes de Acidobacteria.

Comunidade bacteriana dominante associada aos utrículos de *Utriculária gibba*.
Fonte: (FERREIRA A. J., 2011).

Tabela 6. Significância encontrada na comparação das diferentes classes.

Filo	Classe	Local 1 X Local2	Local1 X <i>U.gibba</i>	Local2 X <i>U. gibba</i>
Acidobacteria	Acidobacteria_Gp1	3,98^{-03*}	6,00^{-14*}	6,00^{-14*}
	Acidobacteria_Gp2	1,09 ⁻⁰¹	7,57^{-03*}	3,02 ⁻⁰¹
	Acidobacteria_Gp3	3,73 ⁻⁰¹	2,58^{-12*}	1,68^{-16*}
	Acidobacteria_Gp4	2,28 ⁻⁰¹	9,62 ⁻⁰¹	2,51 ⁻⁰¹
	Acidobacteria_Gp5	5,83 ⁻⁰¹	5,18 ⁻⁰¹	0
	Acidobacteria_Gp6	4,66⁻⁰²	8,54^{-03*}	2,67^{-05*}
	Acidobacteria_Gp7	0	2,37 ⁻⁰¹	5,64 ⁻⁰¹
	Acidobacteria_Gp10	1,95 ⁻⁰¹	4,83 ⁻⁰¹	5,34 ⁻⁰¹
	Holophagae	8,21^{-04*}	1,12 ⁻⁰¹	2,38^{-06*}
Actinobacteria	Actinobacteria	6,00^{-14*}	6,00^{-14*}	6,00^{-14*}
Bacteria_incertae_sedis	Ktedonobacteria	3,15 ⁻⁰¹	4,25 ⁻⁰¹	8,46 ⁻⁰²
Bacteroidetes	Bacteroidia	8,80 ⁻⁰¹	5,18 ⁻⁰¹	4,39 ⁻⁰¹
	Flavobacteria	1,04^{-04*}	3,57^{-04*}	5,64 ⁻⁰¹
	Sphingobacteria	6,00^{-14*}	2,60^{-06*}	6,00^{-14*}
Chlamydiae	Chlamydiae	6,00^{-14*}	6,00^{-14*}	1,74 ⁻⁰¹
Chlorobi	Chlorobia	9,18 ⁻⁰²	6,82 ⁻⁰²	0
Chloroflexi	Anaerolineae	4,21^{-03*}	2,67^{-05*}	2,05^{-11*}
	Caldilineae	5,44 ⁻⁰¹	1,86^{-03*}	2,52 ⁻⁰²
Deinococcus-Thermus	Deinococci	5,83 ⁻⁰¹	5,18 ⁻⁰¹	0
Firmicutes	Bacilli	5,29 ⁻⁰¹	5,29 ⁻⁰¹	2,15 ⁻⁰¹
	Clostridia	2,80^{-13*}	6,00^{-04*}	6,00^{-14*}
	Erysipelotrichi	5,83 ⁻⁰¹	5,18 ⁻⁰¹	0
Gemmatimonadetes	Gemmatimonadetes	7,80^{-03*}	1,24 ⁻⁰¹	4,28^{-05*}
Lentisphaerae	Lentisphaeria	4,90 ⁻⁰¹	7,83 ⁻⁰²	1,66 ⁻⁰¹
Planctomycetes	Planctomycetacia	5,94 ⁻⁰¹	2,25^{-05*}	2,38^{-06*}
Proteobacteria	Alphaproteobacteria	1,24 ⁻⁰¹	6,00^{-14*}	6,00^{-14*}
	Betaproteobacteria	1,16^{-07*}	6,00^{-14*}	6,00^{-14*}
	Deltaproteobacteria	6,00^{-14*}	6,00^{-14*}	6,00^{-14*}
	Epsilonproteobacteria	3,15 ⁻⁰¹	1,34 ⁻⁰¹	0
	Gammaproteobacteria	6,00^{-14*}	6,00^{-14*}	6,00^{-14*}
Tenericutes	Mollicutes	1,70 ⁻⁰¹	1,34 ⁻⁰¹	0
Verrucomicrobia	Opitutae	1,71^{-05*}	2,71 ⁻⁰²	1,98^{-09*}
	Spartobacteria	6,00^{-14*}	6,00^{-14*}	6,00^{-14*}
	Subdivision3	4,13^{-05*}	5,74^{-07*}	6,00^{-14*}
	Subdivision5	9,18 ⁻⁰²	6,82 ⁻⁰²	0
	Verrucomicrobiae	4,24 ⁻⁰¹	5,03 ⁻⁰¹	8,26 ⁻⁰¹

* valores significativos, onde $p < 0,01$.

Fonte: (FERREIRA A. J., 2011).

As sequências também foram analisadas quanto sua similaridade ao nível de gênero presentes no banco de dados (figura 11). Para esta análise, foram consideradas apenas sequências que alcançaram pelo menos 95% de similaridade com as sequências do banco de dados. Os resultados revelaram a presença de 255 gêneros diferentes, sendo a maior parte dessas sequências (90%) distribuída entre vários gêneros com presença inferior a 1,5% em relação ao total amostrado. Dessa forma, ocorreram diferenças nas proporções observadas de gêneros em relação às análises levando em conta classes e filos. Além disso, a busca também revelou que 28% das sequências continham similaridade com *Polynucleobacter*, constituindo-se assim, no grupo com maior frequência. Além disso, outras sequências apresentaram maior similaridade com

gêneros como *Acidobacterium* (7,8%), *Microcella* (6,0%), *Variovorax* (5,7%), *Cupriavidus* (5,7%), *Curvibacter* (5,2%), *Hydrogenophaga* (4,9%), *Delftia* (4,8%), *Yonghaparkia* (4,7%), *Methylococcus* (3,9%), *Comamonas* (3,4%), *Cryobacterium* (3,1%), *Geobacter* (3,0%), *Pseudomonas* (2,0%), *Undibacterium* (1,6%), *Herbaspirillum* (1,4%), *Methylobacillus* (1,2%), *Arcicella* (1,2%), *Azonexus* (1,2%), *Anaeromyxobacter* (1,1%), *Caulobacter* (1,0%), *Aquicella* (1,0%), *Tepidimonas* (0,9%) e *Chromobacterium* (0,8%)

A comparação das sequências de acordo com o tratamento indicou certa similaridade entre as frequências dos grupos dominantes tanto no local 1 quanto no local 2. O local 1 apresentou uma distribuição mais homogênea, além de uma menor ($p < 0,01$) concentração de bactérias similares a *Polynucleobacter* (6,7%) em relação ao local 2 no qual *Polynucleobacter* corresponde a 15% das sequências identificadas ao nível de gênero. Também foi observado que poucas sequências obtidas de *U. gibba* apresentaram similaridade com *Polynucleobacter*, sendo este percentual menor que 0,12% ($p < 0,01$). Assim este grupo pertencente à Betaproteobacteria é responsável pela maior parte da comunidade desta classe encontrada nos locais 1 e 2. Além disso, a densidade muito inferior desse gênero em associação aos utrículos indica um maior sucesso na colonização da água. Ao invés de *Polynucleobacter*, foi encontrada uma maior densidade de *Acidobacterium* em *U. gibba*, perfazendo pouco mais de 23% ($p < 0,01$) do total identificado. Entretanto não foram encontradas sequências com similaridade à *Acidobacterium* nos locais 1 e 2. Outros gêneros da comunidade de *U. gibba* mostraram diferenças significativas ($p < 0,01$) em relação aos gêneros presentes nos dois locais, dentre os gêneros pode-se destacar *Methylococcus* e *Geobacter*. Em *U. gibba* o total de sequências similares à *Methylococcus* chegou a 11%, contra 0,3% no local 1 e 0,02% no local 2. Ainda em *U. gibba* foi observado um percentual grande de 7% de *Geobacter*, sendo maior que 1,5% e 0,06% observado nos locais 1 e 2, respectivamente. Além desses gêneros, em *U. gibba* também foram encontrados outros gêneros em proporção superior, tais como *Aquicella*, *Methylobacillus*, *Anaeromyxobacter*, *Tepidimonas* e *Azonexus*, não sendo os dois últimos gêneros detectados no local 2. Além disso, outros grupos como *Microcella*, *Cupriavidus*, *Hydrogenophaga*, *Curvibacter*, *Comamonas*, *Variovorax*, *Yonghaparkia*, *Delftia* e *Cryobacterium* apresentaram uma maior proporção na água em relação à planta.

Figura 11. Relação Gênero.

Proporção de gêneros de maior similaridade com as sequências obtidas de *U. gibba* e dos locais 1 e 2. Linhas de fundo representando o grupo (filó ou classe) ao qual pertencem os gêneros.

Fonte: FERREIRA A. J., 2011.

6 DISCUSSÃO

O procedimento de seleção de sequências é extremamente importante para que as análises de diversidade estejam mais próximas da realidade. Uma seleção deficiente pode provocar um aumento expressivo no número de UTOs, pois podem ser criados vários *clusters* diferentes com as mesmas sequências. Caso este problema persista em alinhamentos múltiplos como *clustalw* e *muscle*, a matriz pode tornar-se distorcida, indicando que sequências de baixa qualidade ocasionam também um alinhamento de qualidade inferior, causando impactos significativos sobre a análise (SCHLOSS, 2010). Além disso, a acurácia da classificação taxonômica seria prejudicada devido à alta dependência da qualidade das sequências (KUNIN et al., 2008). Dessa forma, a utilização de sequências de qualidade foi fundamental para que os resultados não se desviassem radicalmente da realidade. Neste contexto, o alinhamento de referência utilizado pelo RDP-II considera a estrutura secundária formada pelo fragmento do gene 16S rRNA, posicionando melhor as bases no alinhamento (COLE et al., 2009). A qualidade do alinhamento é fundamental, pois pode afetar significativamente as distâncias calculadas, implicando em uma maior ou menor diversidade genética em regiões variáveis. Schloss (2010) verificou que de acordo com o método de alinhamento utilizado é possível interpretar diferentes graus de diversidade genética na mesma região. O emprego da ferramenta *Aligner Infernal* (RDP-II) nessas regiões variáveis indicou uma maior diversidade genética em relação ao método SILVA (PRUESSE et al., 2007), com exceção das regiões V3 e V4 (não utilizadas neste estudo). Além da análise da estrutura secundária formada pelo rRNA, outras medidas são importantes para melhorar a precisão do alinhamento. Tais medidas como a observação de bases identificadas, qualidade das mesmas, presença de regiões conservadas e tamanho mínimo dos fragmentos sequenciados podem ser úteis para caracterizar sequências dentro de parâmetros mínimos de qualidade. Entretanto, apesar dos esforços em melhorar a qualidade, estes métodos de controle podem subestimar a diversidade, não detectando determinados organismos do ambiente (SCHLOSS, 2010).

A utilização de métodos não paramétricos também pode estimar a riqueza. Entretanto, esta abordagem pode extrapolar as estimativas de UTOs e a estrutura da população, pois estes métodos também variar em decorrência do tamanho das amostras. Por outro lado, a construção de curvas de rarefação foi importante para dimensionar o esforço amostral e comparar a riqueza entre comunidades. Sabe-se que o número de organismos detectados em uma determinada amostra é fortemente influenciado pelo

número de sequências analisadas em qualquer nível filogenético (SCHLOSS; HANDELSMAN, 2005). Dessa forma, a utilização dessas curvas também ajudou a eliminar distorções nos resultados. Outra vantagem na utilização de curvas de rarefação é que os dados gerados são semelhantes às análises do EstimateS, porém são bem mais rápidos (SCHLOSS; HANDELSMAN, 2005).

Neste trabalho foi observado que a riqueza de alguns grupos pode divergir significativamente de acordo com o ambiente. Mesmo em trechos de várzea comunicantes pelo rio foi observada uma riqueza muito similar e uma diversidade ligeiramente diferente, ocorrendo diferenças significativas entre poucos grupos. As diferenças entre grupos nos locais 1 e 2 podem ter se dado em decorrência de várias razões como o relevo, força da corrente, deposição do sedimento, contaminação das águas, tipo de solo e flora local. Em diversos trabalhos foi observada uma menor riqueza na população microbiana aquática em relação aos solos, provavelmente porque o meio aquoso apresenta um desafio hidrostático para muitas bactérias. Além disso, o meio aquoso fornece um ambiente mais homogêneo, diferentemente do solo que pode ser interpretado como diversos nichos que podem ser colonizados por diferentes espécies bacterianas. Dentre as características que podem estar atuando para o perfil da comunidade observada, provavelmente a maior interferência seja a contaminação por águas residuais que são despejadas no rio Tietê e seu afluentes. Em trabalhos envolvendo águas dulcícolas empregando técnicas independentes de cultivo, já foi observado que a proporção de Betaproteobacteria em águas com poucos nutrientes costuma ser menor, corroborando com as observações feitas nos locais deste trabalho, onde o número grande dessa classe de bactérias pode ser consequência do excesso de nutrientes provindo do descarte de esgoto no rio. Além disso, o ambiente apresenta características de eutrofização, indicando que a comunidade microbiana nessas áreas de várzea do Alto Tietê pode sofrer mudanças devido ao aumento da concentração de nutrientes e/ou moléculas específicas.

Xing, Hahn e Wu (2009) analisaram lagos oligotróficos de água doce e salgada e verificaram que ambos tinham predominância de Bacteroidetes e Cyanobacteria, seguidos de Actinobacteria, Alphaproteobacteria e Betaproteobacteria. Em trabalho utilizando a comunidade presente em lagos de água doce eutrofizados Tamaki et al. (2005) também constataram predominância do grupo Proteobacteria, sendo predominantes Deltaproteobacteria e Betaproteobacteria. Já grupos como Cyanobacteria e Verrucomicrobia tiveram pequena participação, similarmente ao ocorrido no presente

trabalho. Outro grupo com maior frequência neste trabalho foi a classe Clostridia, esta é conhecida pelo fato de muitos de seus representantes se adaptarem em condições anaeróbicas. Seu habitat é predominante no solo, mas já foram encontrados em poeira, esgotos, rios, lagos, água do mar, leite, verduras, carne fresca, peixes, insetos e no trato intestinal (SATHISH; SWAMINATHAN, 2009), podendo se adaptar à água de diferentes locais. Ainda, observando as possíveis alterações que a eutrofização causada pelo despejo de águas residuais pode exercer sobre a comunidade bacteriana, está a reprodução descontrolada de cianobactérias e algas eucarióticas. Nesse ambiente, as comunidades bacterianas utilizam nutrientes, como o nitrogênio e fósforo presentes em níveis elevados, sendo a presença destes nutrientes considerada um fator limitante, e interferindo assim no crescimento destes organismos fotossintéticos (CONLEY, 2000). Em estações de tratamento de águas residuais, por exemplo, a principal fonte de nitrogênio e fósforo provém de matéria fecal, resíduos industriais e produtos de limpeza como detergentes sintéticos (ARVIN; JENKINS, 1985). Devido ao fato de muitos membros do filo Cyanobacteria serem diazotróficos a demanda por nitrogênio não é tão limitante quanto a concentração de fósforo (KORTSTEE et al., 1994; NIXON et al., 1996), pois os organismos diazotróficos são capazes de fixar o nitrogênio atmosférico, suprimindo suas necessidades. O crescimento descontrolado leva ao esgotamento dos nutrientes e conseqüentemente à morte dos organismos fototróficos, queda nos níveis de oxigênio e a morte estagnação do ambiente.

Dessa forma, a comunidade dos locais de estudo pode estar sofrendo de um processo de alterações nos níveis de oxigênio. Além disso, o aumento na turbidez da água reduz a penetração da luz, e conseqüentemente a geração de oxigênio provinda da atividade fotossintética. Assim, a alteração da comunidade microbiana está diretamente ligada às mudanças nos processos ambientais. Em trabalhos envolvendo águas residuais normalmente são identificadas muitas bactérias envolvidas na remoção biológica do fósforo, nitrificantes e desnitrificantes (SEVIOUR; MINO; ONUKI, 2003). Outro problema é a presença de Cyanobacteria, pois essas bactérias produzem substâncias tóxicas, podendo seu consumo causar ameaça à saúde pública e levando à morte em alguns casos (CODD, 1995). O tratamento da água altera a comunidade microbiana, entretanto, estudos têm demonstrado que alguns patógenos oportunistas podem sobreviver em sistemas de distribuição de água potável. Essa resistência pode ser atribuída a mecanismos como a permeabilidade da parede celular e formação de biofilme (FREEMAN et al., 2006; GROBE; WINGENDER; FLEMMING., 2001), assim o

tratamento da água à jusante do rio pode servir em parte como processo seletivo de alguns microrganismos. Hong et al. (2010) constataram em seu estudo que a comunidade em sistemas de distribuição de água potável pode ser constituída principalmente por Betaproteobacteria (53%) e Alphaproteobacteria (43%), outros grupos como Gammaproteobacteria, Actinobacteria dentre outras constituíram uma menor parcela da comunidade, similarmente à proporção encontrada neste trabalho.

Inicialmente *Acinetobacter* spp., uma Gammaproteobacteria, era tida como principal removedora de fósforo, entretanto ao aplicar técnicas independentes de cultivo, Streichan, Golecki e Schön (1990) não confirmaram este fato, visto que estes estudos apontaram para membros da classe Betaproteobacteria como responsáveis pela remoção biológica do fósforo (HIRAISHI; UEDA; ISHIHARA, 1998). Bond et al. (1999) constataram que a remoção biológica do fósforo em lodos de águas residuais é realizada principalmente por membros do filo Actinobacteria e Proteobacteria da classe Betaproteobacteria. Assim como neste trabalho, eles também observaram que as populações dessas bactérias foram dominantes nas amostras analisadas, sendo constituídas por 55% de Betaproteobacteria e 35% de Actinobacteria (proporções semelhantes). Outros trabalhos utilizando diferentes métodos como análise de quinonas, construção de bibliotecas 16S rRNA e hibridação *in situ* identificaram como principais membros na comunidade Betaproteobacteria seguido por Actinobacteria. Também foi observado que a classe Betaproteobacteria tinha vínculo com o grupo *Rhodocyclus* (HIRAISHI; UEDA; ISHIHARA, 1998; WAGNER et al., 1994; BOND et al., 1999). Outra característica observada em trabalhos envolvendo densidade predominante de Betaproteobacteria é a capacidade de oxidar a amônia, promovendo a nitrificação (OTAWA et al., 2006). Estudos com base no gene 16S rRNA e genes funcionais revelaram que a oxidação do nitrogênio está relacionada principalmente a duas betaproteobactérias, *Nitrosomonas* spp. e *Nitrosospira* spp.

Este nitrogênio originado de efluentes pode provocar a contaminação das águas doce, subterrâneas e marinhas, levando à eutrofização (OTAWA et al., 2006). Os microrganismos nitrificantes podem reduzir a eficiência do tratamento da água, diminuindo o efeito da monoclорamina e aumento posterior do crescimento microbiano (EICHLER et al., 2006). Entretanto, não foram detectados *Nitrosomonas* spp. e *Nitrosospira* spp. na presente análise, sugerindo que possa existir uma característica diferenciada deste ambiente, onde outros microrganismos podem utilizar este nutriente.

Polynucleobacter foi o gênero mais frequente tanto no local 1 como no local 2. A espécie *P. necessarius* foi descrita inicialmente como uma bactéria endossimbionte obrigatória do citoplasma de ciliados do gênero *Euplotes*, e comumente chamado de omicron ou partícula parecida com omicron (HECKMANN; SCHMIDT, 1987). Posteriormente, diferentes autores demonstraram que esta espécie não é encontrada exclusivamente no interior de *Euplotes* spp., sendo um táxon planctônico de grande importância numérica e funcional em sistemas de água (HAHN, 2003; HAHN *et al.*, 2009; JEZBEROVÁ *et al.*, 2010). Experimentos de hibridação *in situ* indicaram que esse endossimbiose seria pouco provável, devido a alta complementaridade com bactérias de vida livre (BURKERT *et al.*, 2003), sendo levantadas duas hipóteses. A primeira, seria que *P. necessarius* é um membro de vida livre do bacterioplâncton de água doce. Outra hipótese, é que um ancestral de *P. necessarius* teria invadido ciliados, adaptando-se e perdido a capacidade de viver livre. Entretanto, *P. necessarius* atualmente é descrita como uma bactéria cosmopolita de vida livre com distribuição relatada em ambientes de água doce na Europa central, China e leste da África. Estes ambientes variam em diferentes zonas climáticas, sendo relatado em águas doce alcalinas e ácidas, lagos oligotróficos em montanhas, lagoas eutrofizadas, rios, em sistemas de turfeiras rasas e poças temporárias em estradas florestais. Sua colonização pode ser permanente ou temporária, além disso, sua presença já foi descrita em zonas anóxicas de sistemas estagnados (JEZBEROVÁ *et al.*, 2010; JEZBERA *et al.*, 2011). Sua principal contribuição no meio ambiente é na constituição do bacterioplâncton, mas apesar de sua ampla distribuição, seu cultivo em laboratório não é fácil, sendo crucial um processo de aclimatização (HAHN, 2003). Bruns *et al.* (2003) observaram que bactérias de tamanho diminuto (0,2 a 0,5 µm de diâmetro) como Actinomycetales (quase 100% das actinobactérias encontradas na água analisada deste trabalho) teriam grande densidade em ambientes aquáticos, pois seria mais difícil para predadores de bactérias “pastejarem” células não filamentosas. Neste contexto, *P. necessarius* além da adaptação a variações de pH e temperatura, também possui célula de tamanho reduzido, sendo este um dos motivos para explicar a grande disseminação do grupo (HAHN, 2003). Um fator limitante para seu crescimento populacional seria a oferta de carbono orgânico na forma de substâncias húmicas (BURKERT *et al.*, 2003). Apesar disso, o grupo parece não ter uma preferência específica por certos ambientes de água doce, embora *P. necessarius* parece variar sazonalmente (HAHN, 2003).

Espera-se encontrar grande diversidade bacteriana nos rios, uma vez que a mesma é resultado da mistura de diversos táxons de bactérias aquáticas e terrestres que se misturam pelas vias de escoamento (LILES et al., 2003; TOPP, 2003; ZWART et al., 2002). Dessa forma, este trabalho é importante, uma vez que a comunidade bacteriana presente em rios e outros ecossistemas de água doce tem sido pouco estudada através de técnicas independentes de cultivo. As bactérias ribeirinhas são as menos conhecidas entre as de outros sistemas aquáticos (COTTRELL et al., 2005). Essas bactérias agem na decomposição da matéria orgânica, desempenhando um papel fundamental transformação do material transportado para os oceanos (OPSAHL; BENNER; AMON, 1999; HERNES; BENNER, 2002). Porém atividades antropogênicas alteram este mecanismo, causando impactos sobre o consumo de matéria orgânica e, consequentemente sobre o ecossistema (COVERT; MORAN, 2001).

Caravieri (2011) ao trabalhar na mesma área (estrada Mogi-Salesópolis-SP) e comparar com a comunidade presente em plantas presentes no estado do Mato Grosso-Brasil, detectou por meio de técnicas independentes de cultivo a comunidade bacteriana associada aos utrículos e estolões de *Utricularia breviscapa*. Seus resultados indicam que a densidade da comunidade presente na planta varia conforme o ambiente, ressaltando a capacidade diferenciada da planta hospedeira em suportar microrganismos conforme o local. Nas plantas, tanto em utrículos quanto estolões foi encontrado predominantemente Betaproteobacteria, sendo esta fração da comunidade de aproximadamente 47% em utrículos e 66% em estolões. Entretanto, ainda no trabalho de Caravieri (2011) quando observados os gêneros encontrados, incluindo Betaproteobacteria, nota-se que não foi detectado *P. necessarius*, nem de outros gêneros apresentados em uma densidade superior a 1% neste presente trabalho (*Anaerococcus*, *Comamonas*, *Cryobacterium*, *Cupriavidu*, *Curvibacter*, *Delftia*, *Frontridians*, *Hydrogenophaga*, *Kluyvera*, *Leifsonia*, *Microcella*, *Parachlamydia*, *Pseudoclavibacter*, *Sporichthya*, *Variovorax*, *Verrucomicrobium* e *Yonghaparkia*). Curiosamente alguns gêneros descritos na planta não foram detectados na água (*Defluviicoccus*, *Kitasatospora*, *Rahnella*, *Ralstonia* e *Rhodococcus*), sugerindo que a planta desempenha um ambiente seguro e estável principalmente para *Ralstonia* que na planta contribui com aproximadamente 6,3% da comunidade do utrículo e 21,8% do estolão. Este fato indica a importância da interação das bactérias com a planta hospedeira e sua capacidade de selecionar a comunidade presente em seu interior. Dessa forma, essa relação também pode ser benéfica para planta, que poderia ser auxiliada nos processos

de digestão da presa, empregando enzimas providas da comunidade bacteriana associada e obtendo os nutrientes necessários para seu desenvolvimento. Assim o ambiente atuaria fornecendo os microrganismos e estes os que fossem capazes de se adaptarem à mesma seriam mantidos.

As análises estatísticas quanto ao compartilhamento de UTOs e a densidade de sequências presentes nas mesmas indicaram que apesar da grande oferta de microrganismos da água à planta, ocorre influência do ambiente (água) sobre comunidade de *U. gibba*. Entretanto, esta influência não é o fator principal para determinar a população bacteriana dominante nos utrículos da planta, ficando a maior influência em função do nicho criado pela planta hospedeira. Corroborando com esta análise, a comparação da comunidade presente nos utrículos com a presente na água por identificação das sequências revelou também a maior frequência de grupos considerados raros na água. Ao comparar a comunidade presente na água com a comunidade presente nos utrículos de *U. gibba* foi observada a ocorrência de aproximadamente 59% dos gêneros compartilhados entre a planta e a água de pelo menos um dos locais, porém a densidade variou bastante, reforçando ainda mais a hipótese de que os utrículos são capazes de fornecer um microambiente no qual uma comunidade bacteriana pode ser selecionada devido ao estabelecimento de uma relação diferenciada, na qual ocorre uma função biológica benéfica tanto para as bactérias quanto para a planta hospedeira.

A predominância de gêneros de Acidobacteria na comunidade associada aos utrículos pode estar envolvida na degradação de presas e matéria orgânica presente na água. Entretanto há poucas informações sobre os indivíduos que compõe o filo, sendo pouco descrito, apesar de ser considerado um grupo com alta taxa de diversidade filogenética, chegando a ser próxima à Proteobacteria (HUGENHOLTZ; GOEBEL; PACE, 1998), além de amplamente distribuído. As acidobactérias já foram identificadas em solos e sedimentos (JANSSEN, 2006), nascentes de água quente (BRYANT *et al.*, 2007), regiões de turfeira (DEDYSH *et al.*, 2006), lagos de mineração (KLEINSTEUBER *et al.*, 2008), sistemas de distribuição de água (MARTINY *et al.*, 2005), no mar (QUAISER *et al.*, 2008), cavernas (MEISINGER *et al.*, 2007) e ambientes poluídos (WARD *et al.*, 2009). Pankratov *et al.* (2008) observaram uma versatilidade do grupo testando a capacidade de metabolizar diferentes fontes de carbono, sendo destacada a capacidade de crescer em meios contendo xilana, pectina, amido, etanol e metanol. Assim, a grande distribuição do filo, sua abundância e sua capacidade de resistir a ambientes poluídos e extremos, sugere um importante papel

desse grupo no ambiente. Apesar disso, sabemos muito pouco sobre sua fisiologia individual, funções e contribuições metabólicas para o meio ambiente (WARD et al., 2009).

Observando Acidobacteria como comunidade dominante associada aos utrículos, pode-se inferir que acidobactérias podem colonizar a armadilha com sucesso, mesmo que estes sejam caracterizados por extremos de pH, concentração de O₂ e etc., provavelmente devido ao fato de que este filo apresenta membros capazes de se estabelecer em ambientes desafiadores. Considerando que foram detectadas 1047 sequências similares a *Acidobacterium*, e que o volume médio de fluido em dez utrículos é de 20µL (SIROVÁ; ADAMEC; VRBA, 2003) e foram utilizados 200mL de água por local, a frequência de *Acidobacterium* nos utrículos é pelo menos 1,04x10⁶ vezes maior que no mesmo volume de água. Assim, os utrículos se mostram um ambiente extremamente benéfico aos representantes deste gênero. A estratégia de sobrevivência destas bactérias nos utrículos é muito eficiente e pode estar ligada à tolerância e à utilização de fontes de carbono provenientes da planta e matéria orgânica diversa que adentra a armadilha (WARD et al., 2009). Okamura et al. (2011) observaram a plasticidade do metabolismo de membros do filo testando com sucesso o crescimento com diversas fontes de carbono como L-arabinose, celobiose, D-frutose, D-galactose, D-glicose, glicerol, mio-inositol, D-lactose, maltose, D-manose, sacarose, trealose, D-xilose, gluconato, L-glutamato, histidina, ácido casamino, extrato de levedura e pepitona. O interior dos utrículos pode ter a temperatura baixa e ainda ácido que muitas bactérias desse filo crescem ativamente, uma vez que os membros do filo são caracterizados por habitar ambientes com pH baixo, crescendo mesmo em temperaturas baixas. Pankratov et al. (2008) verificaram que alguns membros das subdivisões 3 (presentes neste estudo) crescem em pH entre 5,5 a 6,5 e os representantes da subdivisão 1, a maior fração da comunidade de Acidobacteria presente neste estudo, chegam a crescer sob condições de laboratório em níveis de pH de até 3,5. Neste contexto, Sirová, Adamec e Vrba (2003) mediram o pH do fluido presente dentro dos utrículos de quatro espécies do gênero *Utricularia* (*U. aurea*, *U. foliosa*, *U. australis* e *U. vulgaris*) e constataram que o pH se mantinha ácido, mesmo estando a armadilha ocupada ou vazia. Os níveis de pH variaram de 5,2 nas armadilhas jovens e vazias de *U. aurea* a 4,2 nas armadilhas mais antigas e vazias de *U. vulgaris*. O pH também é um fator importante para a atividade de enzimas como fosfatases, as quais possuem uma atividade diferenciada em relação ao meio externo. Sirová, Adamec e Vrba (2003)

verificaram que o pH baixo no interior dos utrículos pode aumentar a atividade dessas enzimas, o que consequentemente pode disponibilizar este nutriente para a planta, principalmente em ambientes oligotróficos. Este fato também poderia facilitar a colonização do utrículo por membros do filo Acidobacteria, em detrimento a outros grupos.

Além da capacidade de crescer em meio ácido, membros de Acidobacteria também possuem a capacidade de crescer ativamente em baixas temperaturas, sendo consistente em trabalhos que avaliam a diversidade bacteriana de pântanos frios e ácidos. Outra característica do organismo é a capacidade de metabolizar substâncias complexas como hemicelulose, celulose e quitina (WARD et al., 2009). Neste contexto, há indícios que essas bactérias tenham aparato enzimático para auxiliar na digestão de presas como pequenos artrópodes por meio de quitinases, além da matéria orgânica vegetal. Ward *et al.* (2009), por meio de estudos do genoma, observaram evidências quanto a capacidade de redução de nitrato e nitrito, porém não foi observada a fixação do nitrogênio atmosférico, nem de desnitrificação. No genoma também foram encontrados muitos genes de receptores de sideróforos, mas não de produção, indicando a necessidade de associação com outros microrganismos para a aquisição de ferro. Além disso, existem diversos genes codificantes de síntese de celulose, e de proteínas de alto peso molecular que podem ser utilizadas para resistir à desidratação e produção de biofilme. Isto pode contribuir quando a armadilha da planta carnívora se rearma, expulsando a água e estabelecendo uma pressão hidrostática negativa.

Os alinhamentos com baixa similaridade às sequências tipo do banco de dados (RDPquery) observados neste trabalho podem ser considerados normais, pois apesar do filo ser muito diverso, o mesmo está atualmente dividido em apenas cinco gêneros válidos taxonomicamente, *Acidobacterium*, *Holophaga*, *Geothrix*, *Terriglobus* e *Edaphobacter* (KOCH et al., 2008; PANKRATOV et al., 2008). Em 1997, o filo continha 4 a 5 subdivisões (LUDWIG et al., 1997), chegando a 26 atualmente (EICHORST, 2011). Apesar do grande desenvolvimento, a taxonomia do filo Acidobacteria é extremamente pobre, sendo necessários mais esforços para cultivar e caracterizar fisiologicamente o grupo. Evidências por meio do sequenciamento do genoma indicam que Acidobacteria compartilha características com Proteobacteria, Cyanobacteria, e fungos (WARD et al., 2009). Assim, apesar da biologia molecular ser uma ferramenta fundamental para identificação de acidobactérias no ambiente, ainda faltam dados a respeito de seu papel no ambiente.

Methylococcus e *Geobacter* também apresentaram diferenças significativas entre a comunidade associada aos utrículos e a comunidade presente na água. *Geobacter* é constituído por bactérias anaeróbias pertencentes à Gammaproteobacteria e constituem um grupo com importância em ambientes contaminados, onde podem prover a biorremediação. A característica mais marcante do gênero é a capacidade de suas espécies em utilizar inúmeros doadores e receptores de elétrons como H₂, Fe(III), acetato fumarato e malato (CACCAVO et al., 1994). *Geobacter* spp. podem ser utilizadas como agente de biorremediação em águas subterrâneas, degradando petróleo (SNOEYENBOS-WEST et al., 2000) e reduzindo o urânio (VI) a urânio (IV), sua forma insolúvel, a qual pode ser removida da água (WILKINS et al., 2006). Dessa forma, seu metabolismo pode interferir diretamente em ciclos biogeoquímicos, atuando na mobilidade de metais (KING et al., 2009). Além disso, as espécies de *Geobacter* possuem uma alta expressão de genes de fixação de nitrogênio, sugerindo que estejam envolvidas nesse processo. Outras bactérias como *Rhodospirillum rubrum* (também presente neste trabalho) são incapazes de realizar a fixação de nitrogênio, competindo com *Geobacter* por este nutriente (ZHUANG et al., 2011). Esta interação ilustra a importância da associação entre vários grupos de bactérias para a comunidade como um todo e a planta hospedeira. O fato dos resultados indicarem que *Geobacter*, um grupo de bactérias anaeróbias, serem densos nos utrículos e conviverem com outras aeróbias como acidobactérias indica que a comunidade bacteriana é diversa metabolicamente. Esta diversidade metabólica deve coexistir, atuando em um ambiente dinâmico e instável. Adamec (2007) analisou a concentração de oxigênio presente no fluido dos utrículos e constatou que os utrículos apresentam variações drásticas nos níveis de oxigênio. As armadilhas desarmadas possuem níveis de oxigênio comparáveis ao do meio externo, entretanto quando estas armadilhas disparam e a maior parte da água é bombeada para fora, a concentração de oxigênio no interior das armadilhas decai, chegando a zero.

O gênero *Methylococcus* possui duas espécies, *M. capsulatus* e *M. thermophilus*, são bactérias Gram negativas em forma de cocos e aeróbias obrigatórias. Além disso, são capazes de viver em diversos ambientes, resistindo e crescendo melhor em altas temperaturas (45 °C) (WALTERS et al., 1999). São bactérias importantes ambientalmente por serem metanotróficas, ou seja, podem utilizar o metano como única fonte de carbono, algumas conseguem utilizar o metanol (WARD et al., 2004). Dentro do domínio Bacteria, a distribuição de metanotróficas está limitada a apenas 11 gêneros de Proteobacteria, sendo divididos nas classes Alphaproteobacteria e

Gammaproteobacteria à qual pertence *Methylococcus*. Esta divisão não é apenas decorrente da filogenia, mas também da forma fisiológica como a bactéria assimila metano. As gamaproteobactérias utilizam ribulose monophosphate (RuMP) na via de assimilação de formaldeído, enquanto as alfavproteobactérias utilizam a via da serina (WARD et al., 2004). Graças a esta capacidade de utilizar o metano para a produção de energia esta bactéria tem atraído o interesse no combate ao efeito estufa, remediando locais contaminados por materiais sanitários (WARD et al., 2004). A capacidade de organismos metanotróficos oxidarem metano é devido à presença de enzimas monooxigenases (MMOs), estas podem catalizar a reação criando partículas ligadas à membrana (pMMO) e solúveis (sMMO) (WARD et al., 2004). Além da capacidade de oxidar metano, *M. capsulatus* possui a capacidade de fixar o nitrogênio atmosférico, sendo importante em ambientes com falta desse nutriente. A fixação é possível graças aos genes *nif*, os quais são compartilhados com outras espécies de bactérias fixadoras (BENDER; CONRAD, 1994). Além da capacidade de fixar o nitrogênio, *Methylococcus* pode realizar desnitrificação e nitrificação, convertendo o nitrogênio consideravelmente (WARD et al., 2004). A alta frequência de *Methylococcus* em *U. gibba* pode estar envolvida na disponibilização de nitrogênio para a planta e degradação do metano resultante da ação de microrganismos metanogênicos que decompõe a matéria orgânica no interior dos utrículos.

No presente trabalho foi observado que os utrículos de *U. gibba* apresenta uma grande diversidade bacteriana, similar àquela presente no ambiente aquático onde esta planta é encontrada. Entretanto, existe uma variação significativa nas espécies presentes e na densidade de grupos compartilhados entre o ambiente externo (água) e os utrículos. Dessa forma, é possível sugerir que a comunidade bacteriana no interior dos utrículos é proveniente do meio externo (água), mas que seleciona grupos específicos independente da sua densidade no meio aquático. A comunidade bacteriana dos utrículos de *U. gibba* deve ser adaptada a uma variação extrema nas condições ambientais, e utilizar a matéria orgânica proveniente das presas capturadas pela planta.

7 CONCLUSÕES

Os resultados obtidos permitem concluir que:

- A amostragem, de acordo com as curvas de rarefação, foi satisfatória, principalmente até o nível de gênero;
- O local de amostragem e não a seleção feita pela planta é o principal fator associado à diversidade bacteriana, visto que a diversidade da planta e da água no local 1 foi similar entre si e diferiu significativamente do local 2;
- A composição da comunidade é determinada pelo nicho ocupado, visto que ocorreu maior similaridade entre o local 1 e o local 2, do que entre o utrículo (obtido do local 1) e a água do local 1;
- Proteobacteria representa mais de 50% da comunidade total, sendo a classe Betaproteobacteria, representada pelo gênero *Polynucleobacter*, o mais importante numericamente nas duas lagoas amostradas;
- A comunidade associada aos utrículos é diferente da comunidade presente na água, sendo os representantes do gênero Gp1 de Acidobacteria a comunidade dominante neste ambiente.

REFERÊNCIAS*

- ADAMEC, L. Oxygen Concentrations Inside the Traps of the Carnivorous Plants *Utricularia* and *Genlisea* (Lentibulariaceae). **Ann. Botany**, v. 100, p. 849–856, 2007.
- ALTSCHUL, S. F.; MADDEN, T. L.; SCHAFFER, A. A.; ZHANG, J.; ZHANG, Z.; MONTEIRO, W.; LIPMAN D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucl. Acids Res.**, v. 25, p. 3389-3402, 1997.
- APPLIED BIOSYSTEMS BY LIFE TECHNOLOGIES. **ABI PRISM® 377 Genetic Analyzer**. Disponível em: < <http://www.appliedbiosystems.com/absite> > Acesso em: 03 mar. 2010.
- ARVIN, E.; JENKINS, D. Biological removal of phosphorus from wastewaters. **Crit. Rev. Environ. Control.**, v. 15, p. 25-64, 1985.
- AXELROOD, P. E., M. L. CHOW, C. C. RADOMSKI, J. M. MCDERMOTT, AND J. DAVIES. Molecular characterization of bacterial diversity from British Columbia forest soils subjected to disturbance. **Can. J. Microbiol.**, v. 48, p. 655– 674, 2002.
- AZEVEDO, J. L. Biodiversidade Microbiana e Potencial Biotecnológico. In: MELO, I. S.; AZEVEDO, J. L. **Ecologia Microbiana**. Embrapa-CNPMA. p. 488, 1998.
- BAINS, W.; SMITH, G. C. A novel method for nucleic acid sequence determination. **J. Theor. Biol.**, v. 135, p. 303–307, 1988.
- BARNS, S. M.; CAIN, E. C.; SOMMERVILLE, L.; KUSKE, C. R. Acidobacteria phylum sequences in uranium-contaminated subsurface sediments greatly expand the known diversity within the phylum. **Appl. Environ. Microbiol.**, v. 73, p. 3113–3116, 2007.
- BENDER, M.; CONRAD, R. Microbial Oxidation of Methane, Ammonium and Carbon Monoxide and Turnover of Nitrous Oxide and Nitric Oxide in Soils. **Biogeochemistry**, v. 27, p. 97-112, 1994.
- BLAST. **Basic Local Alignment Search Tool**. Disponível em: <<http://blast.ncbi.nlm.nih.gov/Blast.cgi>>. Acesso em: 08 jun. 2010.
- BOND, P. L.; ERHART, R.; WAGNER, M.; KELLER, J.; BLACKALL, L. L. Identification of Some of the Major Groups of Bacteria in Efficient and Nonefficient Biological Phosphorus Removal Activated Sludge Systems. **Appl. Environ. Microbiol.**, v. 65, p. 4077–4084, 1999.
- BROCCHIERI, L. Phylogenetic inferences from molecular sequences: review and critique. **Theor. Popul. Biol.**, v. 59, p. 7-40, 2001.
- BRUNS, A.; NUBEL, U.; CYPIONKA, H.; OVERMANN, J. Effect of Signal Compounds and Incubation Conditions on the Culturability of Freshwater Bacterioplankton. **Appl. Environ. Microbiol.**, v. 69, p. 1980-1989, 2003.

*De acordo com:

BRUNAK, S.; DANCHIN, A.; HATTORI, M.; NAKAMURA, H.; SHINOZAKI, H.; MATISE, T.; PREUSS, D. Nucleotide Sequence Database Policies. **Science**, v. 298, p. 1333, 2002.

BRYANT, D. A.; COSTAS, A. M. G.; MARESCA, J. A.; CHEW, A. G.; KLATT, C. G.; BATESON M. M.; TALLON, L. J.; HOSTETLER, J.; NELSON, W. C.; HEIDELBERG, J. F.; WARD, D. M. *Candidatus Chloracidobacterium thermophilum*: an aerobic phototrophic *Acidobacterium*. **Science**, v. 317, p. 523–526, 2007.

BURKERT, U.; WARNECKE, F.; BABENZIEN, D.; ZWIRNMANN, E.; PERNTHALER, J.; Members of a Readily Enriched β -Proteobacterial Clade Are Common in Surface Waters of a Humic Lake. **Appl. Environ. Microbiol.**, v. 69, p. 6550-6559, 2003.

CACCAVO, F. J.; LONERGAN, D.; LOVLEY, D. R.; DAVIS, M.; STOLZ, J. F.; MCINERNEY, M. J. *Geobacter sulfurreducens* sp. nov., a hydrogen- and acetate-oxidizing dissimilatory metal-reducing microorganism. **Appl. Environ. Microbiol.**, v. 60, p. 3752-3759, 1994.

CARAVIERI, F. A. **Diversidade genética de bactérias associadas às plantas do gênero *Utricularia* (Lentibulariaceae)**. 2011. 62f. Dissertação (Mestrado em Biotecnologia) - Universidade de Mogi das Cruzes, Mogi das Cruzes, 2011.

CHAO, A.; SHEN, T.-J. **Program SPADE (Species Prediction And Diversity Estimation)**. Disponível em: <<http://chao.stat.nthu.edu.tw>>. Acesso em: 06 jun. 2010.

CODD, G.A. Cyanobacterial toxins: Occurrence, properties and biological significance. **Water Sci. Technol.**, v. 32, p. 149-156, 1995.

COLE, J. R.; WANG, Q.; CARDENAS E.; FISH, J.; CHAI, B.; FARRIS, R. J. KULAM-SYED-MOHIDEEN, A. S.; MCGARRELL, M. S.; MARSH, T.; GARRITY, G. M.; TIEDJE J. M.; The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. **Nucl. Acids Res.**, v. 37, p. D141–D145, 2009.

CONLEY, D. J. Biogeochemical nutrient cycles and nutrient management. **Hydrobiologia**. v. 410, p. 87-96, 2000.

CORRÊA, M. A.; MAMEDE, M. C. H. **Lentibulariaceae**. São Paulo, Instituto de Botânica, 2002.

COTTRELL, M. T.; KIRCHMAN, D. L. Community composition of marine bacterioplankton determined by 16S rRNA gene clone libraries and fluorescence in situ hybridization. **Appl. Environ. Microbiol.**, v. 66, p. 5116–5122, 2000.

COTTRELL, MT.; WAIDNER, LA.; YU, L.; KIRCHMAN DL. Bacterial diversity of metagenomic and PCR libraries from the Delaware River. **Environ. Microbiol.**, v 7, p 1883-1895, 2005.

COVERT, J. S.; MORAN, M. A. Molecular characterization of estuarine bacterial communities that use high- and low-molecular weight fractions of dissolved organic carbon. **Aquat. Microbial. Ecol.**, v. 25, p. 127–139, 2001.

DEDYSH, S. N.; PANKRATOV, T. A.; BELOVA, S. E.; KULICHEVSKAYA, I. S.; LIESACK, W. Phylogenetic analysis and in situ identification of Bacteria community composition in an acidic Sphagnum peat bog. **Appl. Environ. Microbiol.**, v. 72, p. 2110–2117, 2006.

DESANTIS, T. Z. P.; HUGENHOLTZ, N.; LARSEN, M.; ROJAS, E. L.; BRODIE, K.; ELLER, T.; HUBER, D.; DALEVI, P.; HU, E. ANDERSEN, G. L.. Greengenes, a chimerachecked 16S rRNA gene database and workbench compatible with ARB. **Appl. Environ. Microbiol.**, v. 72, p. 5069-5072, 2006.

DESANTIS, T. Z., P. HUGENHOLTZ, N. LARSEN, M. ROJAS, E. L. BRODIE, K. KELLER, T. HUBER, D. DALEVI, P. HU; G. L. ANDERSEN. 2006. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. **Appl Environ Microbiol.**, 72:5069-72.

DUNBAR, J.; BARNS, S. M.; TICKNOR, L. O.; KUSKE, C. R.. Empirical and theoretical bacterial diversity in four Arizona soils. **Appl. Environ. Microbiol.**, v. 68, p. 3035–3045, 2002.

DURBIN, R.; EDDY, S. R.; KROGH, A.; MITCHISON, G. J. RNA structure analysis In: DURBIN, R.; EDDY, S. R.; KROGH, A.; MITCHISON, G. J. **Biological sequence analysis: Probabilistic models of proteins and nucleic acids**. Cambridge (United Kingdom): Cambridge University Press. p. 350, 1998.

EDDY, S. R.; DURBIN, R. RNA sequence analysis using covariance models. **Nucl. Acids Res.**, v. 22, p. 2079–2088, 1994.

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucl. Acids Res.**, v. 32, p. 1792–1797, 2004.

EICHLER, S.; CHRISTEN, R.; HOLTJE, C.; WESTPHAL, P.; BOTEL, J.; BRETTAR, I.; MEHLING, A.; HOFLE, M. G. Composition and dynamics of bacterial communities of a drinking water supply system as assessed by RNA- and DNA-based 16S rRNA gene fingerprinting. **Appl. Environ. Microbiol.**, v. 72, p. 1858-1872, 2006.

EICHORST, S. A.; KUSKE, C. R.; SCHMIDT, T. M. Influence of Plant Polymers on the Distribution and Cultivation of Bacteria in the Phylum *Acidobacteria*. **Appl. Environ. Microbiol.**, v. 77, p. 586–596, 2011.

ESTIMATES. **Statistical Estimation of Species Richness and Shared Species from Samples**. Disponível em: <<http://viceroy.eeb.uconn.edu/estimates>>. Acesso em: 06 jun. 2010.

EWING, B.; HILLIER, L.; WENDL, M. C.; GREEN, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. **Genome Res.**, v. 8, p.175-185, 1998.

- FALSENSTEN, J. PHYLIP-Phylogeny inference package (version 3.2). **Cladistics**, v. 5 p. 164-166, 1989.
- FREEMAN, R.; GEIER, H.; WEIGEL, K. M.; DO, J.; FORD, T. E.; CANGELOSI, G. A. Roles for cell wall glycopeptidolipid in surface adherence and planktonic dispersal of *Mycobacterium avium*. **Appl. Environ. Microbiol.**, v. 72, p.7554-7558, 2006.
- GANS, J.; WOILINSKY, M.; DUNBAR, J. Computational improvements reveal great bacterial diversity and high metal toxicity in soil. **Science**, v. 309, p. 1387–1390, 2005.
- GAUTHERET, D.; LAMBERT, A. Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. **J. Mol. Biol.**, v. 313, p. 1003–1011, 2001.
- GREENGENES . **Green genes**. Disponível em: <<http://greengenes.lbl.gov/>>. Acesso em: 05 mar. 2010.
- GROBE, S.; WINGENDER, J.; FLEMMING, H. C. Capability of mucoid *Pseudomonas aeruginosa* to survive in chlorinated water. **Int. J. Hyg. Environ. Health**, v. 204, p. 139-142, 2001.
- HAHN, M. W. Isolation of Strains Belonging to the Cosmopolitan *Polynucleobacter necessarius* Cluster from Freshwater Habitats Located in Three Climatic Zones **Appl. Environ. Microbiol.**, v. 69, p. 5248-5254, 2003.
- HAHN, M. W.; LANG, E.; BRANDT, U.; W. U, Q. L; SCHEUERL, T. Emended description of the genus *Polynucleobacter* and the species *Polynucleobacter necessarius* and proposal of two subspecies, *P. necessarius* subsp. *necessarius* subsp. nov. and *P. necessarius* subsp. *asymbioticus* subsp. nov. **Int. J. Syst. Evol. Microbiol.**, v. 59, p. 2002-2009, 2009.
- HAINES, J. R.; HERRMANN, R.; LEE, K.; COBANLI, S.; BLAISE, C. Microbial population analysis as a measure of ecosystem restoration. **Biorremediat. J.**, v. 6, p. 283-296, 2002.
- HARRIS, J. K.; KELLEY, S. T.; PACE, N. R. New perspective on uncultured bacterial phylogenetic division OP11. **Appl. Environ. Microbiol.**, v. 70, p. 845–849, 2004.
- HECKMANN, K.; SCHMIDT, HJ. *Polynucleobacter necessarius* gen. nov., sp. nov., an obligately endosymbiotic bacterium living in the cytoplasm of *Euplotes aediculatus*. **Int. J. Syst. Bacteriol.**, v 37, p 456–457. 1987.
- HERNES, P. J.; BENNER, R. Transport and diagenesis of dissolved and particulate terrigenous organic matter in the North Pacific Ocean. **Deep Sea Res. Part I Oceanogr. Res. Pap.**, v. 49, p. 2119–2132, 2002.
- HIRAISHI, A.; UEDA, Y.; ISHIHARA, J. Quinone profiling of bacterial communities in natural and synthetic activated sewage for enhanced phosphate removal. **Appl. Environ. Microbiol.**, v. 64, p. 992–998, 1998.

HONG S-H, BUNGE J, JEON S-O, Epstein SS. Predicting microbial species richness. **Proc. Nat. Acad. Sci. USA**, v. 103, p. 117–122, 2006.

HONG, P.; HWANG, C.; LING, F.; ANDERSEN, G. L.; LECHEVALLIER, M. W.; LIU, W. T. Pyrosequencing Analysis of Bacterial Biofilm Communities in Water Meters of a Drinking Water Distribution System. **Appl. Environ. Microbiol.**, v. 76, p. 5631-5635, 2010.

HUBER, T.; FAULKNER, G.; HUGENHOLTZ, P. Bellerophon; a program to detect chimeric sequences in multiple sequence alignments. **Bioinformatics**, v. 20, p. 2317-2319, 2004.

HUGENHOLTZ, P. Exploring prokaryotic diversity in the genomic era. **Genome Biol.**, v. 3, p. 31-38. 2002.

HUGENHOLTZ, P.; GOEBEL, B. M.; PACE, N. R. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. **J. Bacteriol.**, v. 180, p. 4765–4774, 1998.

HUSON, D. H.; AUCH, A. F.; QI, J.; SCHUSTER, S. C. MEGAN analysis of metagenomic data. **Genome Res.**, v.17, p. 377-386, 2007.

INSDC. **International Nucleotide Sequence Database Collaboration**. Disponível em: <<http://insdc.org>>. Acesso em: 05 mar. 2010.

JANGID, K.; WILLIAMS M. A.; FRANZLUEBBERS A. J.; SANDERLIN, J. S.; REEVES, J. H.; JENKINS, M. B.; ENDALE, D. M. COLEMAN, D. C.; WHITMAN, W. B. Relative impacts of land-use, management intensity and fertilization upon soil microbial community structure in agricultural systems. **Soil Biol. Biochem.**, v. 40, p. 2843-2853, 2008.

JANSSEN, P. H. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. **Appl. Environ. Microbiol.**, v. 72, p. 1719–1728, 2006.

JANSSEN, P.; AUDIT, B.; CASES, I.; DARZENTAS, N.; GOLDOVSKY, L.; KUNIN, V.; LOPEZ-BIGAS, N.; PEREGRIN-ALVAREZ, J.; PEREIRA-LEAL, J. B. TSOKA, S.; OUZOUNIS, C. A. Beyond 100 genomes. **Genome Biol.**, v. 4, p. 402, 2003.

JEZBERA, J.; JEZBEROVÁ, J.; BRANDT, U.; HAHN, M.W. Ubiquity of *Polynucleobacter necessarius* subspecies *asymbioticus* results from ecological diversification. **Environ. Microbiol.**, v. 13, p. 922-931, 2011.

JEZBEROVÁ, J.; JEZBERA J.; BRANDT, U.; LINDSTRÖM, E. S.; LANGENHEDER, S.; HAHN, M. W. Ubiquity of *Polynucleobacter necessarius* subspecies *asymbioticus* in lentic freshwater habitats of a heterogenous 2000 km² area. **Environ. Microbiol.**, v. 12, p. 658–669, 2010.

KENNEDY, A. C. Bacterial diversity in agroecosystems. **Agric. Ecosyst. Environ.**, v. 74, p. 65-76, 1999.

KING, E. L.; TUNCAY, ORTOLEVA; E. P. MEILE, C. In Silico *Geobacter sulfurreducens* Metabolism and Its Representation in Reactive Transport Model. **Appl. Environ. Microbiol.**, v 75, p 83–92, 2009.

KLEINSTEUBER, S.; MÜLLER, F. D.; CHATZINOTAS, A.; WENDT-POTTHOFF, K.; HARMS, H. Diversity and *in situ* quantification of *Acidobacteria* subdivision 1 in an acidic mining lake. **FEMS Microbiol. Ecol.**, v. 63, p. 107–117, 2008.

KOCH, I. H.; GICH, F.; DUNFIELD, P. F.; OVERMANN, J. *Edaphobacter modestus* gen. nov., sp. nov., and *Edaphobacter aggregans* sp. nov., acidobacteria isolated from alpine and forest soils. **Int. J. Syst. Evol. Microbiol.**, v. 58, p. 1114–1122, 2008.

KORTSTEE, G. J. J.; APPELDOORN, K. J.; BONTING, C. F. C.; NIEL, E. W. J.; VEEN, H. W. Biology of polyphosphate-accumulating bacteria involved in enhanced biological phosphorus removal. **FEMS Microbiol. Rev.**, v. 15, p. 137–153, 1994.

KUNIN, V.; COPELAND, A.; LAPIDUS, A.; MAVROMATIS, K.; HUGENHOLTZ, P. A Bioinformatician's Guide to Metagenomics. **Microbiol. Mol. Biol. Rev.**, v. 72, p. 557–578, 2008.

LEPLAE, R.; LIMA-MENDEZ G.; TOUSSAINT, A. ACLAME: A Classification of Mobile genetic Elements, update 2010. **Nucl. Acids Res.**, v 38, p D57–D61, 2010.

LILES, M. R.; MANSKE, B. F.; BINTRIM, S. B.; HANDELSMAN, J.; GOODMAN, R. M. A census of rRNA genes and linked genomic sequences within a soil metagenomic library. **Appl. Environ. Microbiol.**, v. 69, p. 2684–2691, 2003.

LUDWIG, W.; SCHLEIFER, K. H.; WHITMAN, W.B. Revised Road Map to the Phylum Firmicutes. In: VOS, P. DE GARRITY, G. JONES, D. KRIEG, N.R. LUDWIG, W. RAINEY, F.A. SCHLEIFER, K.-H.; WHITMAN, W.B. **Bergey's Manual of Systematic Bacteriology**. v. 3, 2ed. New York, Springer-Verlag. 2008, p. 01–08.

LUDWIG, W.; STRUNK, O.; WESTRAM, R.; RICHTER, L.; MEIER, H.; YADHUKUMAR.; BUCHNER, A.; LAI, T.; STEPPI, S.; JOBB, G.; FÖRSTER, W.; BRETTSCHE I.; GERBER, S.; GINHART, AW.; GROSS, O.; GRUMANN, S.; HERMANN, S.; JOST, R.; KÖNIG, A.; LISS, T.; LÜBMANN, R.; MAY, M.; NONHOFF, B.; REICHEL, B.; STREHLOW, B.; STAMATAKIS, A.; STUCKMANN, A.; VILBIG, A.; LENKE, M.; LUDWIG, T.; BODE, A.; SCHLEIFER, KH. Detection and *in situ* identification of representatives of a widely distributed new bacterial phylum. **FEMS Microbiol Lett.**, v. 153, p. 181–190, 1997.

MARGULIES, M., M. EGHOLM, W. E. ALTMAN, S. ATTIYA, J. S. BADER, L. A. BEMBEN, Genome sequencing in microfabricated high-density picolitre reactors. **Nature**, v. 437, p. 376–380, 2005.

MARTÍN, H. G.; IVANOVA, N.; KUNIN, V. WARNECKE, F.; BARRY, K.W.; MCHARDY, A. C.; YEATES, C.; SALAMOV, A. A.; SZETO, E.; DALIN, E.; PUTNAM, N. H.; SHAPIRO, H. J.; PANGILINAN, J. L.; RIGOUTSOS, I.;

KYRPIDES, N. C.; BLACKALL, L. L.; MCMAHON, K. D.; HUGENHOLTZ, P. Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities. **Nature Biotechnol.** v. 24, p. 1263-1269, 2006.

MARTINY, A. C.; ALBRECHTSEN, H-J.; ARVIN, E.; MOLIN, S. Identification of bacteria in biofilm and bulk water samples from a nonchlorinated model drinking water distribution system: detection of a large nitrite-oxidizing population associated with *Nitrospira* spp. **Appl. Environ. Microbiol.**, v. 71, p. 8611–8617, 2005.

MEISINGER, D. B.; ZIMMERMANN, J.; LUDWIG, W.; SCHLEIFER, K. H.; WANNER, G.; SCHMID, M.; PHILIP, C. B.; ENGEL, A. S.; NATUSCHKA, L. LEE, M. *In situ* detection of novel *Acidobacteria* in microbial mats from a chemolithoautotrophically based cave ecosystem (Lower Kane Cave, WY, USA). **Environ. Microbiol.**, v. 9, p. 1523–1534, 2007.

MEYER, F.; PAARMANN, D.; D'SOUZA, M.; OLSON, R.; GLASS, E. M.; KUBAL, M.; PACZIAN, T.; RODRIGUEZ, A.; STEVENS, R.; WILKE, A.; WILKENING J.; EDWARDS, R. A. The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. **BMC Bioinformatics**, v. 9, p. 386-394, 2008.

MG-RAST. **MG-RAST metagenomics analysis server**. Disponível em: <<http://metagenomics.anl.gov/>>. Acesso em: 05 mar. 2010.

MOHANTY, S. R.; BHARATI, K.; DEEPA, N.; RAO, V. R.; ADHYA, T. K. Influence of heavy Metals on Methane Oxidation in tropical rice Soils. **Ecotoxicol. Environ. Saf.**, v. 47, p. 277-284, 2000.

MOTHUR. **Mothur**. Disponível em: <<http://www.mothur.org/>>. Acesso em: 05 mar. 2010.

NACIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. **GenBank**. Disponível em: <<http://www.ncbi.nlm.nih.gov/genbank/>>. Acesso em: 22 jun. 2010.

NAWROCKI, E.P.; EDDY, S.R. Query-Dependent Banding (QDB) for Faster RNA Similarity Searches. **PLoS Comput. Biol.** v.3, p. 540-554, 2007.

NIXON, S. W.; AMMERMAN, J. W.; ATKINSON, L. P.; BEROUNSKY, V. M.; BILLEN, G.; BOICOURT, W. C.; BOYNTON, W. R.; CHURCH, T. M.; DITORO D. M.; ELMGREN, R. The fate of nitrogen and phosphorus at the land–sea margin of the north Atlantic ocean. **Biogeochemistry**, v. 35, p.141-180, 1996.

OKAMURA, K.; KAWAI, A.; YAMADA, T.; HIRAISHI, A. *Acidipila rosea* gen.nov., sp.nov., anacidophilic chemoorganotrophic bacterium belonging to the phylum Acidobacteria. **FEMS Microbiol. Lett.**, v. 317, p. 138–142, 2011.

OPSAHL, S.; BENNER, R.; AMON, R. M. W. Major flux of terrigenous dissolved organic matter through the Arctic Ocean. **Limnology Oceanogr.**, v. 44, p. 2017–2023, 1999.

OTAWA, K.; ASANO, R.; OHBA, Y.; SASAKI, T.; KAWAMURA, E.; KOYAMA, F.; NAKAMURA, S.; NAKAI, Y. Molecular analysis of ammonia-oxidizing bacteria community in intermittent aeration sequencing batch reactors used for animal wastewater treatment. **Environ. Microbiol.**, v. 8, p. 1985–1996, 2006.

PANKRATOV, T. A.; SERKEBAEVA, Y. M.; KULICHEVSKAYA, I. S.; LIESACK, W.; DEDYSH, S. N. Substrate-induced growth and isolation of *Acidobacteria* from acidic *Sphagnum* peat. **ISME J.**, v. 2, p. 551–560, 2008.

PRUESSE, E.; QUAST, C.; KNITTEL, K.; FUCHS, B.; LUDWIG, W.; J. PEPLIES, AND F. O. GLÖCKNER. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. **Nucl. Acids Res.**, v. 35, p. 7188-7196, 2007.

QUAISER, A LÓPEZ-GARCÍA, P.; ZIVANOVIC, Y.; HENN, M. R.; RODRIGUEZ-VALERA, F.; MOREIRA, D. Comparative analysis of genome fragments of *Acidobacteria* from deep Mediterranean plankton. **Environ. Microbiol.**, v. 10, p. 2704–2717, 2008.

R. **The R Project for Statistical Computing**. Disponível em: < <http://www.r-project.org/> >. Acesso em: 06 jun. 2010.

RDP PYROSEQUENCING PIPELINE. **Pipeline Initial Process**. Disponível em: < <http://pyro.cme.msu.edu/init/form.spr> >. Acesso em: 09 mar. 2010.

RDP. **Ribosomal Database Project**. Disponível em: <<http://rdp.cme.msu.edu/>> Acesso em: 05 mar. 2010.

REBER H. H. Simultaneous estimates of the diversity and the degradative capability of heavy-metalaffected soil bacterial communities. **Biol Fert Soils**, v. 13, p.181-186, 1992.

RIKHVANOV, E. G.; VARAKINA, N. N.; SOZINOV. D. Y.; VOINIKOV, V. K. Association of bacteria and yeasts in hot springs. **Appl. Environ. Microbiol.**, v. 65, p. 4292-4293, 1999.

ROCHE. **454 Sequencing**. Disponível em:< <http://www.my454.com/>>. Acesso em: 23 fev. 2011.

ROCHE. **Sistema Titanium**. Disponível em: <<http://www.454.com/>>. Acesso em: 30 ago. 2010.

ROESCH, L. F. W.; FULTHORPE, R. R.; RIVA, A.; CASELA, G.; HADWIN, A. K. M.; KENT A. D.; DAROUB S. H.; CAMARGO, F. A. O.; FARMERIE G. W. E TRIPLET, E. W. Pyrosequencing enumerates and contrasts soil microbial diversity. **ISME - Int Microbial Ecol.**, v. 1, p. 283–290, 2007.

RONAGHI, M. Pyrosequencing Sheds Light on DNA Sequencing. **Gen. Res.**, v. 11, p. 3-11, 2001.

RONAGHI, M., KARAMOHAMED, S., PETTERSSON, B., UHLEN, M. & NYREN, P. Real-time DNA sequencing using detection of pyrophosphate release. **An. Biochem.**, v. 242, p. 84–89, 1996.

RONAGHI, M.; UHLEN, M.; NYREN, P. A sequencing method based on real-time pyrophosphate, **Science**, v. 281, p. 363–365, 1998.

SANGWAN, P.; CHEN, X.; HUGENHOLTZ, P.; JANSSEN, P. H. *Chthoniobacter flavus* gen. nov., sp. nov., the first pure-culture representative of subdivision two, *Spartobacteria* classis nov., of the phylum Verrucomicrobia. **Appl. Environ. Microbiol.**, v. 70, p. 5875–5881, 2004.

SATHISH, S.; SWAMINATHAN, K. Genetic diversity among toxigenic clostridia isolated from soil, water, meat and associated polluted sites in South India. **Ind. J. Medical Microbiol.**, v. 27, n. 4, p. 311–320, 2009.

SCHLOSS P. D, HANDELSMAN J. Toward a census of bacteria in soil. **PLOS Comput. Biol.**, v. 2, p.786–793, 2006.

SCHLOSS, P. D. A High-Throughput DNA Sequence Aligner for Microbial Ecology Studies. **PLoS One**. v. 4, 2009.

SCHLOSS, P. D. The Effects of Alignment Quality, Distance Calculation Method, Sequence Filtering, and Region on the Analysis of 16S rRNA Gene-Based Studies. **PLoS Comput. Biol.**, v. 6, 2010.

SCHLOSS, P. D.; HANDELSMAN, J. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. **Appl. Environ. Microbiol.**, v. 71, p. 1501–1506, 2005.

SCHLOSS, P. D.; WESTCOTT, S. L.; RYABIN, T.; HALL, J. R. 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.; HORN, D. J. V.; WEBER C. F. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. **Appl. Environ. Microbiol.**, v. 75, p. 7537–7541, 2009.

SEVIOUR, R.J.; MINO, T.; ONUKI, M. The microbiology of biological phosphorus removal in activated sludge systems. **FEMS Microbiol. Rev.**, v. 27, p. 99–127, 2003.

SILVA. **Silva comprehensive ribosomal RNA databases**. Disponível em: <<http://www.arb-silva.de/>> Acesso em: 05 mar. 2010.

SIMO. **Sapelo Island Microbial Observatory**. Disponível em: <<http://simo.marsci.uga.edu/>>. Acesso em: 05 mar. 2010.

SIROVÁ, D.; ADAMEC, L.; VRBA J. Enzymatic activities in traps of four aquatic species of the carnivorous genus *Utricularia*. **New Phytol.**, v. 159, p 669–675, 2003.

SNOEYENBOS-WEST, O. L.; NEVIN, K. P.; ANDERSON, R. T. LOVLEY, D. R. Enrichment of *Geobacter* species in response to stimulation of Fe(III) reduction in sandy aquifer sediments. **Microbiol. Ecol.**, v. 39, p 153-167, 2000.

SOREK, R.; ZHU, Y.; CREEVEY, C. J.; FRANCINO, M. P.; BORK, P.; RUBIN, E. M. Genome-Wide Experimental Determination of Barriers to Horizontal Gene Transfer. **Science**, v. 318 p. 1390-1391, 2007.

STACKEBRANDT, E.; GOEBEL, B. M. Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology. **Int. J. Syst. Bacteriol.**, v. 44, p. 846-849, 1994.

STOLZ J. F.; BOTKIN D. B.; DASTOOR M. N. The integral biosphere. In: RAMBLER, M. B.; MARGULIS, L. E.; FESTER, R. (eds). **Global Ecology**. New York: Academic Press, p. 31-49, 1989.

STREICHAN, M.; GOLECKI, J. R.; SCHÖN, G. Polyphosphate-accumulating bacteria from sewage plants with different processes for biological phosphorus removal. **FEMS Microbiol. Lett.**, v. 73, p. 113–124, 1990.

SUNDQUIST, A.; BIGDELI, S.; JALILI, R.; DRUZIN, M. L.; WALLER, S.; PULLEN, K. M.; EL-SAYED, Y. Y.; TASLIMI, M. M. ; BATZOGLOU, S.; RONAGHI, M. Bacterial flora-typing with targeted, chip-based pyrosequencing. **BMC Microbiol.**, v. 7, p. 108, 2007.

SUZUKI, M. T.; GIOVANNONI, S. J. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. **Appl. Environ. Microbiol.**, v.62, p. 625-630, 1996.

TAMAKI, H.; SEKIGUCHI, Y.; HANADA, S.; NAKAMURA, K.; NOMURA, N.; MATSUMURA, M.; KAMAGATA, Y. Comparative Analysis of Bacterial Diversity in Freshwater Sediment of a Shallow Eutrophic Lake by Molecular and Improved Cultivation-Based Techniques. **Appl. Environ. Microbiol.**, v. 71, p. 2162-2169, 2005.

TAYLOR, P. **The genus *Utricularia*: a taxonomic monograph**. London, Royal Botanic Gardens, Kew. 1989.

TOPP, E. Bacteria in agricultural soils: Diversity, role and future perspectives. **Can. J. Soil Sci.**, v. 83, p. 303–309, 2003.

TORSVIK V.; DAAE F. L.; SANDAA R. A.; OVREAS, L. Novel techniques for analyzing microbial diversity in natural and perturbed environments. **J. Biotechnol.**, v. 64, p. 53-62, 1998.

TRUPER H. G. Prokaryotes: an overview with respect to biodiversity and environmental importance. **Biodivers. Conserv.**, v.1 p. 227-36, 1992.

VENTER, J. C., K. REMINGTON, J. F. HEIDELBERG, A. L. HALPERN, D. RUSCH, J. A. EISEN, D. WU, I. PAULSEN, K. E. NELSON, W. NELSON, D. E. FOUTS, S. LEVY, A. H. KNAP, M. W. LOMAS, K. NEALSON, O. WHITE, J.

PETERSON, J. HOFFMAN, R. PARSONS, H. BADEN-TILLSON, C. PFANNKUCH, Y. H. ROGERS, H. O. SMITH. Environmental genome shotgun sequencing of the Sargasso Sea. **Science**, v. 304, p. 66-74, 2004.

VON MERING, C.; HUGENHOLTZ, P.; RAES, J.; TRINGE, S. G.; DOERKS, D.; JENSEN, L. J.; WARD, N.; BORK, P. Quantitative phylogenetic assessment of microbial communities in diverse environments. **Science**, v. 315 p. 1126-1130, 2007.

VON WINTZINGERODE, F.; GOBEL, U. B.; STACKEBRANDT, E. Determination of microbial diversity in environmental samples: pitfalls of PCRbased rRNA analysis. **FEMS Microbiol. Rev.**, v. 21 p. 213-229, 1997.

WAGNER, M.; ERHART, R.; MANZ, W.; AMANN, R.; LEMMER, H.; WEDI, D.; SCHLEIFER, K-H. Development of an rRNA targeted oligonucleotide probe specific for the genus *Acinetobacter* and its application for in situ monitoring in activated sludge. **Appl. Environ. Microbiol.**, v. 60, p. 792–800, 1994.

WALTERS, K. J.; GASSNER, G. T.; LIPPARD, S. J.; WAGER, G. Structure of the soluble methane monooxygenase regulatory protein B. **Proc Natl Acad Sci. Biophys.**, v. 96, p 7877-7882, 1999.

WARD, N. L.; CHALLACOMBE, J. F.; JANSSEN, P. H.; HENRISSAT, B.; COUTINHO, P. M.; WU, M.; XIE, G.; HAFT, D. H.; SAIT, M.; BADGER, J.; BARABOTE, R. D.; BRADLEY, B.; BRETTIN, T. S.; BRINKAC, L. M.; BRUCE, D.; CREASY, T.; DAUGHERTY, S. C.; DAVIDSEN, T. M.; DEBOY, R. T.; DETTER, J. C.; DODSON, R. J.; DURKIN, A. S.; GANAPATHY, A.; GWINN-GIGLIO, M.; HAN, C. S.; KHOURI, H.; KISS, H.; KOTHARI, S. P.; MADUPU, R.; NELSON, K. E.; NELSON, W. C.; PAULSEN, I.; PENN, K.; REN, Q.; ROSOVITZ, M. J.; SELENGUT, J. D.; SHRIVASTAVA, S.; SULLIVAN, S. A.; TAPIA, R.; THOMPSON, L. S.; WATKINS, K. L.; YANG, Q.; YU, C.; ZAFAR, N.; ZHOU, L.; KUSKE, C. R. Three genomes in the phylum Acidobacteria provide insight into their lifestyles in soils. **Appl. Environ. Microbiol.**, v. 74, p 2046–2056, 2009.

WARD, N.; LARSEN, O.; SAKWA, J.; BRUSETH, L.; KHOURI, H.; DURKIN, AS.; DIMITROV, G.; JIANG, J.; SCANLAN, D.; KANG, KH.; LEWIS, M.; NELSON, KE.; METHÉ, B.; WU, M.; HEIDELBERG, JF.; PAULSEN, IT.; FOUTS, D.; RAVEL, J.; TETTELIN, H.; REN, Q.; READ, T.; DEBOY, RT.; SESHADRI, R.; SALZBERG, SL.; JENSEN, HB.; BIRKELAND, NK.; NELSON, WC.; DODSON, RJ.; GRINDHAUG, SH.; HOLT, I.; EIDHAMMER, I.; JONASEN, I.; VANAKEN, S.; UTTERBACK, T.; FELDBLYUM, TV.; FRASER, CM.; LILLEHAUG, JR.; EISEN, JA. Genomic Insights into Methanotrophy: The Complete Genome Sequence of *Methylococcus capsulatus* (Bath). **PLoS Biol.**, v. 2, p 1616-1628, 2004.

WILKINS, M. J.; LIVENS, F. R.; VAUGHAN, D. J.; LOYD, J. R. The impact of Fe(III)-reducing bacteria on uranium mobility. **Biogeochemistry**, v 78, p. 125-150, 2006.

WILMOTTE, A.; HERDMAN, M. Phylogenetic relationships among the Cyanobacteria based on 16S rRNA sequences. In: GARRITY, G. M., BOONE, D. R.; CASTENHOLZ,

R. W. **Bergey's Manual of Systematic Bacteriology**. v. 1, 2ed. New York, Springer-Verlag. 2001. p. 487–493.

WUYTS, J.; PERRIÈRE, G.; PEER YV. The European ribosomal RNA database. **Nucl. Acids Res.**, v 32, p D101-D103, 2004.

XING, P.; HAHN, M. W.; WU, Q. L. Low Taxon Richness of Bacterioplankton in High-Altitude Lakes of the Eastern Tibetan Plateau, with a Predominance of *Bacteroidetes* and *Synechococcus* spp. **Appl. Environ. Microbiol.**, v. 75, p. 7017-7025, 2009.

YOUSSEF, N; SHEIK, C. S.; KRUMHOLZ, L. R.; NAJAR, F. Z.; ROE, B. A.; ELSHAHED; M. S. Comparison of Species Richness Estimates Obtained Using Nearly Complete Fragments and Simulated Pyrosequencing-Generated Fragments in 16S rRNA Gene-Based Environmental Surveys. **Appl. Environ. Microbiol.**, v. 75, p. 5227–5236, 2009.

ZHANG, S.; HAAS, B.; ESKIN, E.; BAFNA, V. Searching genomes for noncoding RNA using FastR. **IEEE/ACM Transactions Comput. Biol. Bioinformatics**, v. 2, p. 366–379, 2005.

ZHUANG, K.; IZALLALEN, M.; MOUSER, P.; RICHTER, H.; RISSO, C.; MAHADEVAN, R.; LOVLEY, D. R. Genome-scale dynamic modeling of the competition between *Rhodoferrax* and *Geobacter* in anoxic subsurface environments. **ISME J.** v. 5, p 305-316, 2011.

ZWART, G.; CRUMP, B. C.; AGTERVELD, M.; HAGEN, F.; HAN, S. K. Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. **Aquatic Microbiol. Ecol.**, v. 28, p. 141–155, 2002.