GEORGES MIKHAEL NAMMOURA NETO

Desenvolvimento e validação de método para a identificação de micro-organismos metabolicamente ativos em biofilmes de amostras ambientais através da análise de rRNA 16S

Tese apresentada ao Programa de Pós-Graduação em Microbiologia do Instituto de Ciências Biomédicas da Universidade de São Paulo, para obtenção do Título de Doutor em Ciências.

Área de concentração: Microbiologia

Orientador: Prof. Dr. René Peter Schneider

Versão corrigida

São Paulo
2017

RESUMO

NAMMOURA NETO, G. M. **Desenvolvimento e validação de método para a identificação de micro-organismos metabolicamente ativos em biofilmes de amostras ambientais através da análise de rRNA 16S**. 2017. 192 f. Tese (Doutorado em Microbiologia) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, SP, 2017.

A otimização e padronização de métodos moleculares são fundamentais para evitar distorções nos resultados causadas por processamento de ácidos nucleicos da abundância relativa de micro-organismos de uma amostra. Nos estudos sobre identificação microbiana, a etapa de validação é, na maioria dos casos, negligenciada. As principais etapas em que podem ocorrer vieses capazes de alterar a informação sobre a composição da comunidade microbiana são a extração e a RT-PCR. O rRNA é a molécula ideal para identificar os micro-organismos metabolicamente ativos por estar em maior quantidade dentro da célula. A inexistência de um protocolo de extração de RNA “gold standard” e de kits comerciais específicos para cada tipo de amostra ambiental pode comprometer as etapas posteriores à extração. O protocolo de extração de RNA adaptado neste trabalho mostrou alta eficácia na lise celular e na remoção de contaminantes co-extraídos, garantindo a integridade e a qualidade do rRNA extraído de amostras contendo altas concentrações de contaminantes. Devido as diferentes características químicas das amostras ambientais, o uso deste protocolo adaptado pode preservar a informação sobre os indivíduos metabolicamente ativos de uma comunidade microbiana. Foram utilizados consórcios artificiais na validação da etapa de PCR. O efeito sinérgico do uso de aditivos, da Taq polimerase de alto rendimento, do programa de sub-ciclagem e da limitação do programa de amplificação em 10 ciclos, mantiveram a proporção dos *moldes* dos consórcios analisados próximo ao esperado. Após a padronização da técnica de ARDRA, foi definido que há necessidade de entre 500 e 550 UFC para uma cobertura completa da diversidade de lodo ativado. O uso de enzimas combinadas MspI/HaeIII e HhaI/RsaI em uma mesma reação “*double digest*” proporcionou a separação dos clones de lodo ativado utilizando menos etapas de ARDRA. E o critério de corte em $\geq 3\%$ do total agrupamentos, possibilitou selecionar os perfis mais representativos sem a perda de grupos importantes para a análise das bibliotecas. Foi concluído que, o uso de um protocolo inadequado de extração de RNA de amostras complexas pode gerar extratos contendo contaminantes co-extraídos que interferem na atividade da Taq polimerase e nas enzimas de restrição. Após o sequenciamento de nova geração, foi observado que o uso de diferentes iniciadores de PCR e do pré-processamento manual para os dados obtidos, foram essenciais para ampliar a cobertura observada para a amostra de lodo e melhorar a resolução dos resultados e a profundidade de identificação taxonômica, respectivamente.

Palavras-chave: Extração de RNA. Amostras ambientais. Bactérias metabolicamente ativas. MiSeq. RT-PCR. ARDRA.

ABSTRACT

NAMMOURA NETO, G. M. **Development and validation of method for the identification of metabolically active microorganisms in biofilms of environmental samples by 16S rRNA analysis.** 2017. 192 p. Ph. D. thesis (Microbiology) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, SP, 2017.

The optimization and standardization of molecular methods are fundamental to avoid distortions in the results caused by nucleic acid processing of the relative abundance of microorganisms in a sample. In the studies on microbial identification, the validation step is, in most cases, neglected. The main steps in which biases capable of altering the composition of the microbial community can occur are extraction and RT-PCR. The rRNA is the ideal molecule to identify metabolically active microorganisms because they are in the greatest amount within the cell. The lack of a gold standard RNA extraction protocol and specific commercial kits for each type of environmental sample may compromise the post-extraction steps. The RNA extraction protocol adapted in this work showed high efficacy in cell lysis and in the removal of co-extracted contaminants, guaranteeing the integrity and quality of the rRNA extracted from samples containing high concentrations of contaminants. Due to the different chemical characteristics of the environmental samples, the use of this adapted protocol can preserve the information about the metabolically active individuals of a microbial community. Artificial consortia were used to validate the PCR step. The synergistic effect of the use of additives, the high yield Taq polymerase, the sub-cycling program and the limitation of the amplification program in 10 cycles, kept the proportion of the molds of the consortia analyzed close to the expected. After standardization of the ARDRA technique, it was defined that there is a need for between 500 and 550 CFU for a complete coverage of the diversity of activated sludge. The use of MspI / HaeIII and HhaI / RsaI combined enzymes in the same double digest reaction provided the separation of activated sludge clones using fewer ARDRA steps. And the criterion of cut in $\geq 3\%$ of the total groupings, allowed to select the most representative profiles without the loss of important groups for the analysis of the libraries. It was concluded that the use of an inadequate RNA extraction protocol from complex samples can generate extracts containing co-extracted contaminants that interfere with Taq polymerase activity and restriction enzymes. After the sequencing of the new generation, it was observed that the use of different PCR primers and manual preprocessing for the obtained data were essential to increase the coverage observed for the sludge sample and to improve the resolution of the results and the depth of taxonomic identification, respectively.

Keywords: RNA extract. Environmental samples. Metabolically active bacteria. MiSeq. RT-PCR. ARDRA.

1 INTRODUÇÃO E REVISÃO DE LITERATURA

1.1 Estrutura e identificação da população de micro-organismos

Estudos envolvendo a prospecção da diversidade de procariotos em amostras ambientais utilizando métodos dependentes ou independentes de cultivo mostram que os micro-organismos raramente vivem sozinhos. Estes interagem uns com os outros para formar comunidades microbianas complexas, cuja estrutura é determinada por interações cooperativas e competitivas entre seus membros (ZHANG, 2013). Para a identificação de populações de bactérias em uma amostra complexa com uma diversidade grande de espécies de micro-organismos, é necessária a utilização de técnicas independentes de cultivo, também conhecidas como técnicas moleculares. As técnicas tradicionais de identificação bacteriana são baseadas no isolamento do organismo em meio de cultivo, porém, estes procedimentos são inadequados para identificação de bactérias de amostras ambientais, pois os meios de cultura artificiais não são capazes de mimetizar com precisão os diferentes nichos ecológicos (ALMEIDA, 2009; GIOVANNONI et al., 1990; TORSVIK et al., 1998). As reproduções de sistemas naturais em laboratório para pesquisas com biofilme utilizam um número controlado de micro-organismos, sendo uma grande simplificação dos processos naturais. Além disso, muitos micro-organismos aderidos em estruturas de biofilmes não são detectáveis por microscopia óptica convencional por permanecerem ligados a partículas de sedimentos (MUYZER et al., 1993). O biofilme é caracterizado pela síntese de substâncias poliméricas extracelulares (EPS), produzido por bactérias que podem ou não estar aderidas a uma superfície. O EPS pode ser composto de diversas substâncias orgânicas dependendo de onde o biofilme é formado. Altas concentrações de carboidratos são encontradas no EPS de muitas culturas bacterianas isoladas, enquanto proteínas são mais predominantes em biofilmes de lodo ativado de estações de tratamento de esgoto. Outros compostos como substâncias húmicas, ácido urônico e DNA também são encontrados no EPS (LIU, 2002). O EPS influencia na agregação e sedimentação dos flocos suspensos de biofilme em sistemas de tratamento de esgoto (ZENG et al., 2016). Essa matriz exopolimérica confere proteção aos micro-organismos do biofilme contra agentes antimicrobianos (COSTERTON et al., 1994), raios

UV, desidratação (a matriz de EPS contém moléculas altamente hidratadas) e predadores como protozoários (XAVIER et al., 2003). Existem diversos trabalhos caracterizando a diversidade microbiana de tanques de aeração dos sistemas de tratamento de esgoto (BOURREIN, 1999; GARBOR, 2003; GUOBIN, 2007; PUROHIT, 2002), águas marinhas (BAE, 2011; DANG, 2001), e sistemas de água potável (HONG, 2010) através do uso de técnicas moleculares, tendo como molécula alvo o 16S rDNA. Entretanto, com base na análise por rRNA 16S, ainda não se sabe quais são os micro-organismos biologicamente ativos presentes nessas amostras ambientais capazes de sintetizar componentes que formam a estrutura do biofilme.

O emprego de técnicas independentes de cultivo de comunidades microbianas teve início com a análise de perfis de ácidos graxos (KANEDA, 1967; NASSER, 1990) e recentemente os ácidos nucleicos têm se tornado as moléculas alvo preferidas para estudo de diversidade microbiana. O uso do gene rRNA como molécula padrão para classificação taxonômica foi inicialmente proposto por Woese; Fox (1977) e atualmente os genes das subunidades rRNA 16S são amplamente utilizados para estudos e caracterização de comunidades bacterianas em muitos nichos ecológicos (MIZRAHI-MAN et al., 2013). Qualquer gene que possua domínios conservados pode ser empregado para taxonomia, porém o 16S rDNA e rRNA são as moléculas mais utilizadas para este tipo de estudo. A vantagem adicional da abordagem genética em estudos de consórcios microbianos é a flexibilidade inerente ao método. A molécula de escolha depende da questão abordada, que pode ser: (1) quais são os tipos de micro-organismos presentes? (2) o que estes micro-organismos estão fazendo? e (3) como as atividades destes micro-organismos estão relacionados a funções do ecossistema? (RASTOGI; SANI, 2011). A extração de DNA permite a caracterização de comunidades microbianas com enfoque em biodiversidade, pois cada organismo com DNA intacto contribuirá para o resultado da análise. A extração de RNA ribossômico, por outro lado, direciona a análise para o conjunto de micro-organismos fisiologicamente ativos do ambiente estudado (QIN et al., 2008).

1.2 Moléculas alvo

1.2.1 Identificação de comunidades microbianas metabolicamente ativas por rRNA 16S

A identificação de comunidades microbianas a partir da extração do DNA total e posterior amplificação e sequenciamento do gene de rRNA 16S de amostras ambientais não permite discriminar bactérias metabolicamente ativas de latentes (KEER e BIRCH, 2003; NOVITSKY, 1986; ZAKRZEWSKI et al., 2012), devido a esses genes estarem presentes em um mesmo número de cópias dentro da célula (variando entre as espécies), independente da taxa de crescimento do micro-organismo. Entretanto, a abundância do RNA ribossômico é proporcional a atividade metabólica celular em uma amostra de forma que a quantidade de ribossomos da célula reflete a intensidade de seu metabolismo, isto é, quanto maior a taxa de crescimento de um micro-organismo maior é a quantidade de ribossomos produzido por este (BARBOSA, 1998; WAGNER et al., 1994). As análises feitas utilizando o rRNA mostram com maior sensibilidade a dinâmica de uma comunidade microbiana quando comparado aos resultados obtidos por rDNA. Isto é devido a persistência do DNA no ambiente mesmo após a morte celular (QIN et al., 2008). Estes fatores credenciam o RNA ribossômico como molécula-alvo mais adequada para a identificação das bactérias formadoras de biofilme metabolicamente ativas (ALMEIDA, 2009; FELSKÉ et al., 2000; MORGAN et al., 2002). No entanto o RNA é uma molécula menos estável do que o DNA, podendo ser degradado por diversos fatores químicos, bioquímicos e físicos do ambiente como calor, pH, íons metálicos e ribonucleases (BRISCO, 2012; EIGNER, 1961). Esta característica deve ser considerada na elaboração de protocolos de análise de populações microbianas a partir da extração do rRNA.

1.2.2 Determinação da abundância de espécies por rDNA 16S

Para identificação dos micro-organismos metabolicamente ativos pela análise do rRNA 16S, partimos da premissa de que, a quantidade para cada tipo de molde de rRNA (isto é, para cada espécie) reflete a proporção da abundância dos micro-organismos dentro de uma amostra. Sabe-se que a concentração de rRNA no interior da célula depende da taxa de crescimento, sendo que a concentração de rRNA depende da abundância intracelular de nucleosídeos trifosfato (TAO et al., 1999). Assim, a composição do meio onde o micro-organismo está presente interfere na taxa de crescimento. Entretanto, não

podemos estimar o número de células de cada espécie bacteriana de uma amostra adotando o rRNA 16S como molécula de escolha, porque a quantidade de rRNA não está relacionado com relacionado com a abundância. Ramagopal (1984) observou no cultivo da mesma cepa de *E. coli* em diferentes meios de cultura complexos concentrações distintas de ribossomos no mesmo ponto de análise da curva de crescimento, porém o número total de células (viáveis e não viáveis) permaneceu idêntica para ambos dos cultivos. Em contrapartida, para a espécie *Azotobacter agilis*, não houve alteração na concentração de ribossomos mesmo com o aumento de 80% de células não viáveis (SOBEK, 1966).

Alguns organismos, em condição de crescimento equilibrado, podem apresentar diferentes relações entre “taxa de crescimento/concentração de rRNA”. Como por exemplo: (1) taxa de crescimento baixa e concentrações de rRNA constantes; (2) taxa de crescimento intermediária, concentração de rRNA aumentando linearmente com a taxa de crescimento; (3) taxa de crescimento elevada, diminuição na concentração de rRNA com o aumento da taxa de crescimento e; (4) diminuição da taxa de crescimento com aumento da concentração de rRNA (BLAZEWICZ et al., 2013; BINDER; LIU, 1998; KERKHOF; KEMP, 1999; WORDEN; BINDER, 2003). O limite estequiométrico, isto é, a taxa de nutriente disponível (reagente) para a produção de moléculas (produto) estruturais celulares, é um fator importante na regulação do crescimento microbiano e na interação com o ambiente, onde em condições favoráveis, promove o aumento na biossíntese de proteínas ribossomais e de rRNA. De acordo com a hipótese da taxa de crescimento, a síntese de rRNA é mantida em função à taxa de crescimento celular, e o número de genes rRNA estão correlacionados a cada espécie bacteriana em responder a disponibilidade de recursos presentes no meio ambiente (ELSER et al., 2000; KARPINETS et al., 2006; STENER et al., 2002).

Outro fator que pode determinar a quantidade de rRNA e ribossomos é o número de cópias do gene rRNA (*rrn*), onde bactérias de crescimento rápido e com genoma relativamente grande contém maior número de operons *rrn*, essa característica está vinculada à necessidade de ajustes rápidos do metabolismo em que os organismos de crescimento rápido precisam ter perante as mudanças repentinas do suprimento nutricional. Em contrapartida, as bactérias que vivem em ambientes aquáticos estáveis e

com poucos nutrientes, tendem a exibir poucas cópias de operons *rrn* e possuem genomas menores (GYORFY et al., 2015). No domínio *Bacteria*, os genes que codificam o rRNA 16S estão organizados em um operon e embora a seleção natural favoreça a eliminação de genes redundantes em procariotos, é comumente visto a presença de múltiplas cópias de cada gene que codifica os rRNA 16S e 23S (KLAPPENBACH et al., 2000).

Os genes de rRNA (rDNA) são frequentemente usados para a identificação filogenética de micro-organismos, porém o rRNA é amplamente aplicado para caracterizar a atividade metabólica microbiana. Entretanto é necessário estar ciente que, todos os organismos em crescimento são ativos, mas nem todos os organismos ativos estão em crescimento (BLAZEWICZ et al., 2013). Para determinar se uma espécie metabolicamente ativa é a mais abundante dentro de uma amostra, não podemos ter como base somente os dados obtidos com análise do rRNA 16S devido as variações de concentração dessa molécula nas células. Campbell (2011) analisou a relação entre a abundância da comunidade bacteriana e a atividade metabólica pelo sequenciamento de rDNA e rRNA 16S respectivamente, cerca de 20% dos micro-organismos, quando na condição de raros (baixa abundância), apresentou maiores concentrações de rRNA 16S quando estes mesmos foram comparados na condição de predominante (alta abundância), sugerindo que as taxas de crescimento diminuem para algumas bactérias à medida que permanecem em alta abundância. Em amostras ambientais complexas cerca de 50% da comunidade microbiana alternam entre as condições de raros e abundantes, suportando a hipótese que algumas bactérias permanecem dormentes aguardando as mudanças de fornecimento nutricional (CAMPBELL et al., 2011; PEDRÓS-ALIÓ, 2006; SOGIN et al., 2006). Assim, o modo mais significativo de correlacionar à atividade metabólica com abundância da comunidade microbiana pela análise de rRNA e rDNA 16S é considerando também a quantidade do número de cópias do rDNA 16S por genoma de cada espécie bacteriana.

Algumas bactérias em fase estacionária podem diminuir concomitantemente o volume celular e o número de ribossomos por célula, entretanto outras espécies como as ultramicrobacterias, podem manter estáveis o tamanho celular e a concentração de ribossomos (FEGATELLA et al., 1998; SCHUT et al., 1997). Assim, em um determinado volume de amostra, duas espécies em fase de crescimento diferentes podem conter

concentrações de rRNA 16S similar, mas com concentrações distintas de rDNA, isto é, estar em abundâncias diferentes.

1.3 Validação de métodos de extração de RNA em amostras ambientais

Existe uma grande diversidade de protocolos de extração de DNA e RNA de amostras ambientais descritos na literatura. Porém pouco se sabe sobre a composição dos micro-organismos fisiologicamente ativos responsáveis pela formação dos biofilmes em amostras complexas, como por exemplo: lodo de esgoto ativado, membrana de osmose reversa, biocorrosão e marinho. A validação da extração de RNA é um ponto importante para a correta interpretação de resultados obtidos por análises moleculares subsequentes. Existe muitos estudos publicados nos últimos cinco anos sobre a caracterização de comunidades microbianas baseadas na extração de ácidos nucleicos, a maioria lastreada na extração de DNA das amostras (Scopus - 159 hits, agosto de 2017). Em contrapartida, estudos de comunidades microbianas a partir da extração de rRNA são mais raros (Scopus – 6 hits nos últimos cinco anos, agosto de 2017). Isto se deve as dificuldades técnicas na extração e no processamento de RNA de amostras microbianas complexas, que são bem maiores do que o processamento de DNA. Testes de aplicabilidade e otimizações em protocolos de extração de RNA ou na co-extração de DNA e RNA para amostras complexas foram descritos na literatura, tendo como alvo as amostras de: solo (FANG et al., 2014; KIM et al., 2013; NIKOLAUSZ et al., 2003; PAULIN et al., 2013; TÖWE et al., 2011), biorreator aerado (QIN et al., 2008), biorreator anaeróbio (IBRAHIM; AHRING et al., 2009), lodo ativado (JIN et al., 2010; McILROY et al., 2008; ROMANO; SPICA, 2001), água doce (McCARTHY et al., 2015), filtros marinhos (SCHNEIDER et al., 2017), sedimento (TATTI et al., 2016) e sistema de água potável (INKINEN et al., 2016). Na maior parte destes estudos, os resultados são interpretados quantitativamente com relação à proporção entre grupos taxonômicos de diversos níveis filogenéticos. As poucas publicações dedicadas à avaliação crítica dos procedimentos de

manipulação de ácidos nucleicos nos protocolos empregados nestes estudos descrevem grande número de interferências nas análises.

Tanto nas análises para a identificação parcial da comunidade microbiana (bibliotecas de clones, “fingerprinting”, Q-PCR, FISH, DNA microarrays, CARD-FISH) como para a identificação completa (DNA-DNA reassociação, sequenciamento genético inteiro, metagenômica, metatranscriptômica) são antecedidas pela etapa de extração de RNA ou DNA (RASTOGI; SANI, 2011). Apesar de muitos testes já terem sido realizados com kits comerciais e protocolos de extração de RNA, ainda não foi descoberto um método que funcione para todos os tipos de amostras ambientais, sendo necessária a otimização dos métodos já existentes (LIM et al., 2016). Alguns trabalhos descritos na literatura têm focado na validação de protocolos de extração de RNA de amostras complexas. McCarthy (2015) avaliaram a eficiência de rendimento entre diferentes protocolos de extração de RNA de amostras de água doce. Os autores observaram que a escolha de um protocolo de extração inadequado pode alterar a informação produzida na análise da composição da comunidade microbiana. Diferentes protocolos de extração de ácido nucleico podem recuperar seqüências de diferentes subsecções da comunidade devido à seletividade de lise e limpeza de extratos (LEITE et al., 2012). Xie (2013) e Lim (2016) focaram na otimização de protocolo de extração de RNA para amostras de biolixiviação e solo respectivamente. McIlroy (2009) avaliaram diferentes protocolos de extração de RNA para lodo ativado, os autores observaram que os melhores resultados de recuperação de RNA foram obtidos com protocolos que adotaram a lise física e sugere que diferentes métodos de extração sejam usados em paralelo para analisar com maior precisão a diversidade da amostra. Assim, a validação do método de extração é uma etapa fundamental para evitar a perda de informação sobre a comunidade microbiana em estudo. E este tipo de validação é possível somente através do emprego de consórcios artificiais de composição conhecida. Stark (2014) avaliaram diferentes métodos de lise na extração de RNA de consórcios bacterianos. Os autores observaram que as espécies Gram positivas apresentaram diferentes níveis de eficiência de extração de RNA, o que resultou na alteração da proporção entre componentes do consórcio nos extratos comparado com o consórcio original.

Um protocolo de extração pode ser dividido em três etapas essenciais (1) lise celular por métodos físicos, químicos ou enzimáticos (2) purificação do extrato através da remoção de material particulado, compostos orgânicos e inorgânicos da solução, finalizando com a (3) recuperação do ácido nucleico por precipitação ou filtração. A otimização do processo de extração é um ponto crítico, pois a aplicação de um protocolo com baixa eficiência de extração pode produzir informações erradas sobre a proporção entre membros da população microbiana (LEVER et al., 2015). Cada amostra ambiental possui uma característica particular, fazendo com que nenhum protocolo denominado “universal” de extração de RNA tenha ainda sido desenvolvido para todos os tipos de amostra. Feinstein (2009) observaram que, o uso de diferentes protocolos de extração de ácido nucleico gerou diferentes estimativas na composição da comunidade microbiana. Assim, muitos trabalhos em que o objetivo é a caracterização filogenética de comunidades microbianas em amostras ambientais, tendo como alvo o RNA, acabam por desconsiderar a etapa de validação do método empregado na pesquisa em questão.

1.4 Amplificação por PCR (*polymerase chain reaction*) em amostras ambientais

A técnica de amplificação de ácidos nucleicos por PCR tem sido muito utilizada em microbiologia ambiental. As estratégias de identificação de comunidades microbianas geralmente incluem o uso da PCR (RASTOGI; SANI, 2011). Entretanto, a amplificação de amostras contendo *moldes* heterogêneos de DNA e cDNA pode estar sujeita à introdução de artefatos durante o processamento que podem alterar significativamente a proporção de *moldes* na amostra. A etapa de amplificação por PCR é essencial para aumentar a quantidade de DNA para análise.

Erros introduzidos na reação de PCR podem ser causados por diversos mecanismos. Iniciadores denominados “universais” podem não reconhecer os sítios de anelamento de grupos importantes de micro-organismos presentes em uma determinada amostra. A alta temperatura de anelamento no programa de PCR pode impedir o iniciador de anelar no sítio do *molde* (SIPOS et al., 2010). A formação de estruturas secundárias no ácido nucleico pode dificultar o acesso e a ligação do iniciador ao seu respectivo sítio de ligação (LIESACK et al., 1991; MEYERHANS et al., 1990; PALLANSCH et al., 1990). Os

ácidos húmicos presentes nos extratos podem se ligarem covalentemente ao DNA impedindo a sua amplificação e inibir a atividade da enzima polimerase. O mecanismo desta inibição ainda é desconhecido devido à natureza amórfica das substâncias húmicas (MATHESON et al., 2010). E por fim a enzima polimerase, onde mesmo o uso de enzimas de alta fidelidade podem produzir mutações (substituições de bases) durante a síntese do DNA (ECKERT e KUNKEL, 1991; McINERNEY et al., 2014). Hong (2010) observaram a perda de informação da diversidade amostral em decorrência do conjunto de iniciadores utilizados na reação de PCR. Mistura de diferentes iniciadores em uma única reação de PCR também pode causar erros como por exemplo, a formação de quimeras durante amplificação (WANG, 1997). Esta alteração na proporção de diferentes membros da comunidade também é observada no emprego de PCR com muitos ciclos de amplificação (QIU, 2001; SUZUKI; GIOVANNONI, 1996). Brooks et al. (2015), analisou a alteração dos dados de composição de consórcios artificiais durante as etapas de extração, seguido de amplificação por PCR, sequenciamento e classificação taxonômica. O autor relatou que as etapas de extração e PCR foram as que mais introduziram alterações na proporção entre sinais dos diferentes membros da comunidade microbiana. Entretanto, os autores deste estudo submeteram as reações a um número extraordinariamente elevado de 30 e 35 ciclos de amplificação de PCR. Sipos (2007) avaliaram o efeito da variação do número de ciclos de PCR na diversidade microbiana em amostra ambiental. O autor concluiu que, tanto os programas de PCR com números reduzidos de ciclos (18) quanto os programas com muitos ciclos (48) de amplificação, não alteraram os resultados de composição microbiana. Entretanto, Suzuki e Giovannoni (1996) observaram que o aumento de ciclos pode de fato alterar a composição microbiana. Entretanto, um dos principais mecanismos de erros na reação de PCR de extratos com diversidade de sequências é a amplificação preferencial, onde ocorre a amplificação seletiva de certas sequências de DNAs de um extrato, alterando a proporção de cada sequência após a reação (KANAGAWA, 2003; POLZ e CAVANAUGH, 1998). A amplificação preferencial ocorre devido a diferenças de eficiência de amplificação entre *moldes* e é progressiva com o incremento do número de ciclos adotados.

Erros na amplificação por PCR são frequentemente imputados à interferência das regiões ricas em G+C (KEBSCHULL; ZADOR, 2015; SEIFI et al., 2012). Essas regiões

produzem dobramentos inter e intracadeias em decorrência do aumento das ligações de hidrogênio com guaninas vizinhas, causando o término prematuro da amplificação pela enzima polimerase. Suzuki e Giovannoni (1996) demonstraram que a variação da porcentagem de guanidina e citosina (G+C) entre *moldes* é um dos fatores responsáveis pela diferença de cinética de amplificação. Quanto maior o teor de G+C maior será a energia e o tempo necessários para a desnaturação do DNA. Assim, a quantidade de *amplicons* dos *moldes* com maior teor de G+C é diminuída progressivamente a cada ciclo de PCR e estes estarão sub-representados no final do programa de amplificação (DUTTON et al., 1993; REYSENBACH et al., 1992). Atualmente o genoma bacteriano de todas as espécies identificadas, exibe uma variação no teor de G+C entre 13% e 75% (BROCCHIERI, 2014; LASSALLE et al., 2015). Este amplo espectro no teor de G+C pode influenciar negativamente na preservação da proporção de *moldes* durante a amplificação. Dentre algumas medidas para minimizar os vieses de PCR causados por *moldes* estruturalmente heterogêneos na mistura estão, a emulsificação da reação de PCR com óleo de silicone (Hori et al., 2007) e a redução do número de ciclos na reação de PCR (SUZUKI e GIOVANNONI, 1996; WILSON e BLITCHINGTON, 1996; WHITFORD et al., 1998).

O número mínimo de ciclos necessários para evitar a distorção entre a proporção de *moldes* do extrato de uma amostra com a respectiva proporção de *amplicons* gerados durante a PCR não está definido. Muitos estudos sobre diversidade microbiana de amostras complexas que utilizaram a técnica de PCR não consideram a introdução de possíveis vieses em etapas de amplificação por PCR com muitos ciclos de PCR. Entre esses estudos, foram utilizados programas de amplificação com 35 ciclos (BRAUN et al., 2016; POITELON et al., 2009; JANG et al., 2016; NAZINA et al., 2017; SMETS et al., 2016), 30 ciclos (BAUTISTA-DE LOS SANTOS et al., 2016; BELILA et al., 2016; DELAFONT et al., 2016; FISCHER et al., 2016), 28 ciclos (XIE et al., 2016), 27 ciclos (CHEN et al., 2016; CAI et al., 2017), 25 ciclos (CAI et al., 2016; JEONG et al., 2016; MCCARTHY et al., 2017; MATAR et al., 2017), 21 ciclos (BAKKE et al., 2015; GONZALES-SILVA et al., 2016), 20 ciclos (SAUNDERS et al., 2016) ou não informado (DOHERTY et al., 2017). Almeida (2009) observou que o número de 10 ciclos de amplificação por PCR diminuiu os vieses

ocasionados em bibliotecas contendo *moldes* heterogêneos de cDNA e permitiu a manutenção da proporção relativa de cada molde no produto da reação.

O uso de aditivos como a betaína (N,N,N-trimetilglicina), DMSO (dimetilsulfóxido), formamida, glicerina, glicerol, DTT (ditiotreitól), BSA (albumina de soro bovino), não somente previnem a formação de estruturas secundárias como também aumentam a taxa de anelamento dos iniciadores em temperaturas mais baixas (Tabela 1; JENSEN, 2010; SIMONOVIC, 2012; SARKAR, 1990; STRIEN, 2013). Seifi et al. (2012) e Jensen (2010) demonstraram que a adição de sulfato de amônio na reação de PCR com concentrações padronizadas de cloreto de magnésio, betaína e DMSO melhorou a amplificação. Lajin (2013) relatou que a adição de 1M betaína melhorou a amplificação por PCR e que devido ao efeito desnaturante do aditivo, a temperatura de anelamento teve de ser reduzida entre 2 a 3 °C em relação a temperatura ótima da reação de PCR sem a betaína. Farrell; Alexandre (2012) concluiu que o BSA, quando usado na combinação com DMSO, atua como um poderoso intensificador de amplificação na PCR. Outros vieses também podem interferir na PCR como por exemplo: componentes da reação e especialmente o DNA, podem adsorver a superfícies poliméricas dos tubos e ponteiros durante o processamento da amostra ou na PCR; nucleases podem degradar as moléculas de RNA e DNA; e o anelamento do iniciador pode ser perturbado por inibidores de PCR, como polissacarídeos (SCHRADER et al., 2012).

Tabela 1 - Modos de ação e efeitos na PCR de diferentes tipos de aditivos.

Tipo	Aditivo	Modo de ação	Efeitos na PCR
Amidas	Formamida	Quebra no pareamento de bases	Aumenta a especificidade e eficiência Diminuição da T_m Redução das estruturas secundárias
Sulfóxido	DMSO	Quebra no pareamento de bases	Aumenta a especificidade e eficiência, especialmente em amostras ricas em G+C Diminuição da T_m Redução das estruturas secundárias

Solutos Compatíveis	Betaina	Estabiliza e facilita a separação das fitas de DNA Diminuição da T _m Fraco efeito de termo- estabilização na Taq polimerase	Melhora a eficiência, especificidade e processividade da PCR Aumenta o intervalo ideal da concentração de MgCl ₂ Aumenta a eficiência de qPCR
	BSA	Elimina os contaminantes	Diminui as interferências na PCR

1.5 Técnicas moleculares de identificação microbiana

Como uma alternativa para a identificação dependente de cultivo, as ferramentas moleculares têm sido aplicadas extensivamente nas últimas décadas para identificação microbiana ambiental. As principais vantagens desta metodologia são o nível de precisão alcançado na determinação genotípica e o grande número de micro-organismos identificados em um curto tempo de análise. Inicialmente, o rRNA 5S era utilizado para as análises de diversidade microbiana, mas as baixas variações de nucleotídeos desta molécula que consequentemente limitava sua utilidade a ecossistemas menos complexos, levou a substituição pelo rRNA 16S (THERON; CLOETE, 2007). A princípio, qualquer gene bacteriano que esteja presente em todos os micro-organismos que possua domínios variáveis e conservados pode ser empregado para taxonomia, mas o rRNA 16S é o gene preferido para estes estudos por estar incorporado na estrutura do ribossomo sendo altamente estável (ALMEIDA, 2009). Dentre as técnicas moleculares de identificação estão as baseadas em “fingerprinting” destrutivo e não-destrutivo dos *amplicons* gerados na etapa de PCR e em sequenciamento de nova geração. Este último é capaz de analisar uma quantidade extremamente grande de *amplicons* em uma única análise.

1.5.1 Clonagem bacteriana

A clonagem é uma técnica molecular muito difundida em ecologia microbiana, que consiste na extração de DNA de uma comunidade microbiana, amplificação do fragmento de interesse, inserção deste em um vetor de clonagem, por fim a inserção (transformação) desse vetor em bactérias competentes. A utilização da clonagem permite maior resolução

na identificação de micro-organismos presentes em amostras ambientais (STOECK; EPSTEIN, 2003; LIANG et al., 2008). Isto é devido a capacidade de isolar cada fragmento clonado em UFC, permitindo o estudo preciso de taxonomia e a detecção de grupos minoritários, os quais são difíceis de detectar por técnicas de “fingerprinting” (ALMEIDA, 2009).

As desvantagens da técnica de clonagem são: o alto tempo consumido; o método é laborioso; a extração de ácido nucleico pode ser dificultada quando a amostra é de solo ou sedimento; muitos clones têm que ser sequenciados para atingir cobertura amostral; a técnica não é quantitativa, pois a etapa de PCR pode intensificar ou mascarar a quantidade dos *moldes* heterogêneos devido às diferenças de acessibilidade aos sítios das moléculas de DNA pelos componentes da reação (ALMEIDA, 2009; SANZ; KOCHLING, 2007). Análises de rarefação de bibliotecas de clones indicam que a diversidade microbiana ambiental não é completamente alcançada com esse tipo de técnica, mas apesar dessas limitações, a clonagem era considerada o “padrão ouro” na pesquisa preliminar de diversidade microbiana (RASTOGI; SANI, 2011).

1.5.2 Técnica de “Fingerprinting”

Métodos de “fingerprint” são amplamente empregados para a caracterização de comunidades microbianas e de suas dinâmicas populacionais (BOTEVA et al., 2013; HEYNDRIKX et al., 1996; PARK, 2015; WEIDNER et al., 1996; ZHANG et al., 2017; YAMEI et al., 2017). Estas análises se iniciam com a extração de DNA ou rRNA das amostras e amplificação dos genes-alvo, sendo geralmente o gene rRNA 16S em procariotos e 18S em eucariotos, por reações de PCR ou RT-PCR com iniciadores universais. Como todos os *amplicons* terão o mesmo comprimento e, portanto, sendo impossível diferenciar sequências de diferentes espécies em géis de agarose, os produtos das reações de PCR devem ser submetidos a técnicas complementares para recuperar a informação sobre a constituição fenotípica da amostra.

1.5.2.1 Análises de Restrição do DNA Ribossomal Amplificado (ARDRA)

As técnicas destrutivas de análise dos produtos de PCR de extratos de comunidades microbianas incluem todos os procedimentos que dependem da digestão dos *amplicons* com uma ou mais enzimas de restrição de alta especificidade, seguida da separação eletroforética em gel de agarose. Os *amplicons* com sítios de digestão em posições distintas produzem um conjunto de fragmentos com tamanhos diferentes, que são identificados em géis de agarose através do perfil característico das bandas. Esta técnica não permite distinguir *amplicons* sem sítios de restrição e a resolução pode ser comprometida quando os produtos da digestão apresentarem pouca diferença em pares de bases. A combinação de fingerprinting não-destrutivo através da construção de bibliotecas clonais com digestão enzimática otimiza a resolução do método. A grande vantagem de submeter clones de bibliotecas clonais a análise com enzimas de restrição é a rapidez na varredura da amostra com identificação rápida de sequências repetidas com consequente economia nos custos de sequenciamento de *amplicons* individuais de DNA. A técnica de ARDRA (*Amplified Ribosomal DNA Restriction Analysis*) é baseada no emprego de endonucleases de alta especificidade para sítios de restrição pouco frequentes em sequências de DNA (ALMEIDA, 2009). A execução de ARDRA com uma única enzima de restrição não é recomendável na análise de consórcios microbianos, onde existe a possibilidade de *amplicons* de diferentes origens não possuírem sítios de clivagem ou da digestão destes *amplicons* resultarem em fragmentos de restrição muito similares ou até idênticos entre si. O uso de enzimas de restrição em conjunto proporciona maior poder de discriminação nas análises (BRUNK et al., 1996; LIN et al., 2009). Digestões enzimáticas utilizando duas (*double digest*) ou mais enzimas de restrição em conjunto para determinação filogenéticas pode melhorar o poder de discriminação de fragmentos originários de organismos diferentes em uma amostra (TODOROVA et al., 2012). A combinação de diferentes enzimas de restrição em ARDRA produz perfis mais diferenciados devido o maior número de fragmentos gerados de tamanhos diferentes (BRUNK et al., 1996). Lin (2009) e Moyer (1996) avaliaram as diferentes combinações de enzimas de restrição na identificação da diversidade bacteriana, onde foi observado que

as combinações de enzimas *MspI/HaeIII*, *MspI/RsaI* possibilitaram o aumento da resolução na detecção de OTUs (unidade taxonômica operacional). As enzimas de restrição *MspI*, *HaeIII*, *MspI* e *RsaI* reconhecem sítios de corte de 4 bases, quando usadas em pares, permite ótimos níveis de fragmentação, isto é, são capazes de gerar um número maior de fragmentos permitindo a discriminação entre as espécies (MOYER, 2001). Além do uso de enzimas em conjunto, na análise de bibliotecas de clones é primordial o uso de consecutivas etapas de ARDRA, aumentando o limite de resolução para o agrupamento filogenético (INGRASSIA et al., 2001). Cada enzima de digestão reconhece diferentes sequências de pares de base no sítio de clivagem. Devido à grande diversidade de enzimas de restrição existentes há inúmeras possibilidades de combinação “*double digest*” em análises de ARDRA e a possibilidade de sequências diferentes serem classificadas como iguais no processo de análise. A aplicação de ARDRA na caracterização de comunidades microbianas complexas depende de uma estratégia para extrair o máximo de informação sobre a diversidade de sequências de uma amostra com um mínimo de esforço laboratorial. Assim, é necessário a padronização da técnica de ARDRA para determinação dos micro-organismos mais abundantes presentes em amostras ambientais complexas.

1.5.2.2 DGGE e TGGE

O processamento não-destrutivo de *amplicons* inclui técnicas como DGGE (*Denaturing Gradient Gel Electrophoresis*) ou TGGE (*Temperature Gradient Gel Electrophoresis*), possibilitam a análise de cada *amplicon* individualmente através de gel de acrilamida com gradiente desnaturante químico ou físico, respectivamente. Em ambas das técnicas, um dos iniciadores utilizados na amplificação dos *moldes* de DNA por PCR contém grampo rico em bases guanina e citosina (G+C) com aproximadamente 40 pb (MYERS et al., 1985), com isso os *amplicons* são separados de acordo com o peso molecular determinado pela quantidade de G+C do DNA. Porém durante a corrida eletroforética, ocorre o aumento gradual do agente desnaturante, o qual faz com que o ácido nucleico de fita dupla se abra ficando presa na malha do gel, assim os *amplicons* com maior quantidade de G+C terão maior mobilidade no gel. Esse método permite

detectar diferenças no comportamento de desnaturação de pequenos *amplicons*, que teoricamente diferem em apenas uma base. Por ser uma técnica dependente de amplificação por PCR, outros fatores podem contribuir para a distorção nos resultados como: a presença de inibidores co-extraídos; *amplicons* de DNA de diferentes organismos com mesmo padrão de desnaturação; a presença do grampo G+C acoplado ao final de um dos iniciadores causando a baixa eficiência de amplificação por PCR aumentando o risco de geração de artefatos na etapa de anelamento, como a formação de heteroduplex (FERRIS et al., 1997; PITERINA; KOZDRÓJ; van ELSAS, 2000; LEE et al., 1996; PEMBROKE, 2013; RUANO e KIDD 1992; SEKIGUCHI, 2001); e a formação de dímeros no iniciador (MOESENEDER, 1999). O DGGE, assim como outras técnicas que empregam o uso de PCR, possui limitações de detecção e de comparações quantitativas das populações detectadas. O principal fator é a perda da correlação entre a abundância da população bacteriana com as intensidades de bandas em géis (PIERANGELA et al.; 2013). Araújo e Schneider (2008) visaram identificar numericamente os organismos dominantes de consórcios artificiais utilizando a técnica de DGGE. Os autores observaram que o uso dos iniciadores contendo o grampo G+C levou à amplificação preferencial de *amplicons* sub-representados nos consórcios. Os autores concluíram ser impossível obter com esta técnica uma representação exata das proporções relativas de um consórcio artificial ou dos organismos de uma comunidade. Outros fatores limitantes na técnica de PCR-DGGE e TGGE em culturas mistas incluem: a alta complexidade dos perfis em géis desnaturantes; informação de sequência limitada (< 500 pb) de bandas de DNA; *amplicons* com ponto de fusão semelhantes e o número de cópias do gene rDNA 16S, podendo distorcer a informação sobre a diversidade em análises de comunidades com composição desconhecida (RASTOGI; SANI, 2011). Assim pode ser concluído que os métodos de DGGE e TGGE não são indicados para a análise de comunidades complexas com grupos minoritários de bactérias.

1.5.3 Sequenciamento de Nova Geração

A identificação de comunidades microbianas por clonagem e técnicas de “fingerprinting” são frequentemente utilizados. Entretanto, uma vez que somente um

número limitado de clones pode ser analisado a complexidade da comunidade não é completamente representada (ZAKRZEWSKI et al., 2012). Os avanços no sequenciamento de nova geração permitiram descobrir as diversidades microbianas em uma ampla variedade de amostras, tendo como alvo o gene rRNA 16S para determinação das relações filogenéticas (REUTER, SPACEK; SNYDER, 2015). Os primeiros estudos de comunidades microbianas tipicamente sequenciavam o gene rRNA 16S inteiro, porém os dados da diversidade eram limitados devido pouca profundidade de sequenciamento. Com os sequenciadores de nova geração o foco mudou para sequenciamento de curtos fragmentos de genes (MIZRAHI-MAN et al., 2013). Atualmente os estudos com comunidades microbianas foram direcionados para as novas tecnologias de sequenciamento de nova geração devido ao barateamento e melhoramento das ferramentas de bioinformática (KOZICH et al., 2013). Os sequenciadores de nova geração estão sendo amplamente utilizados para caracterização microbiana utilizando o gene rRNA 16S como molécula alvo, mas com a atual limitação da técnica, ainda não é possível sequenciar o gene 16S inteiro (1500 pb).

O sequenciamento de nova geração não requer etapas preliminares de clonagem, levando a redução no custo, rapidez e sequenciamento paralelo de muitos fragmentos de DNA. Esta característica gera uma cobertura de sequenciamento suficiente para análises de comunidades microbianas complexas, como por exemplo: sistemas de tratamento de esgoto (JI et al.; 2015) e sistemas de distribuição de água potável (LIU et al.; 2014). Um dos métodos mais novos para sequenciamento é o pirosequenciamento, onde a leitura da sequência é realizada a partir de uma combinação de reações enzimáticas que se inicia com a liberação de um pirofosfato (PPi), oriundo da adição de um desoxinucleotídeo à cadeia. Em seguida, esse pirofosfato é convertido para ATP pela enzima ATP sulfurilase, sendo este, utilizado pela luciferase para oxidar a luciferina produzindo um sinal de luz que é captado por uma câmera CCD (CARVALHO; SILVA, 2010; RONAGHI, 2001). As plataformas mais utilizadas para estudos de comunidades microbianas são fornecidas pela Illumina (Califórnia, USA) e o Ion Torrent (New Hampshire, EUA), ambos utilizam ferramentas de bioinformática para remontagem das sequências identificadas (JEON et al., 2015).

O processo de sequenciamento dos equipamentos da Illumina envolve a ligação do DNA em uma superfície de vidro através de adaptadores específicos. A leitura das bases é feita em ciclos, iniciados com a incorporação da base, lavagem, captação da imagem e clivagem. Cada nucleotídeo incorporado emite uma fluorescência característica (Figura 1A). O sequenciador Ion Torrent desenvolvido pela empresa Life Technologies utiliza a tecnologia baseada na ligação do DNA em uma “bead” e amplificado por PCR-emulsão, o chip semi-condutor capta a mudança do pH através da liberação de íons hidrogênio durante a incorporação de dNTPs (Figura 1B; REUTER, SPACEK; SNYDER, 2015).

Figura 1 - Tecnologias utilizadas nos sequenciadores de nova geração Illumina (A) e Ion Torrent (B). Adaptado de: Reuter, Spacek; Snyder (2015). (A) anelamento da sequência no chip e amplificado, formando clusters, então tem o início do sequenciamento. (B) anelamento da sequência em beads e amplificado, cada bead é alocada em um poço no chip e então sequenciado.

Essa característica torna os sequenciadores de nova geração uma ferramenta de grande importância quando na caracterização microbiana de amostras ambientais, em particular os biofilmes, possibilitando realizar análises qualitativas e quantitativas. As

vantagens desse método sobre a clonagem bacteriana são: diminuição das etapas de PCRs e manipulação das amostras; análise de grande número de amostras em curto período de tempo; maior confiabilidade qualitativa e quantitativa; utilização de *barcodes* para diferenciação de várias amostras em uma única corrida de sequenciamento. Os sequenciadores de nova geração permitem caracterizar as complexas comunidades microbianas.

Considerando que o termo genoma é toda a informação genética de um organismo e a metagenômica é a coleção de toda a informação genética de um ecossistema inteiro, o transcriptoma pode ser descrito como a informação de todas as moléculas de RNA codificantes ou não expressados por uma célula sob uma determinada condição. Consequentemente, a metatranscriptoma pode ser descrito pela análise quantitativa de todos os genes expressos em uma amostra com diversos organismos ou por um ecossistema inteiro (WECHX, 2009). A análise por metagenômica se baseia na extração de DNA para a identificação dos micro-organismos e que consiste na análise de amostras contendo uma complexa diversidade microbiológica independentemente de cultivo em laboratório, tendo em mente que somente uma pequena fração de micro-organismos podem ser cultivados. Em amostras ambientais, a comunidade microbiológica tem melhor cobertura com as análises por metagenômica pelos sequenciadores de última geração devido à enorme quantidade de fragmentos sequenciados que esta tecnologia proporciona.

Os principais objetivos da transcriptômica são: a catalogação de todos os genes transcritos em uma determinada amostra, incluindo mRNA, RNAs não-codificantes e pequenos RNAs; determinação da estrutura transcricional dos genes, em termos de sítios de início, padrão de *splicing* e outras modificações pós-transcricionais; e quantificação da mudança nos níveis de expressão de cada transcriptoma (WANG, 2009). O estudo das moléculas de mRNA por metatranscriptômica em biofilmes se torna importante para identificar quais nutrientes estão sendo consumidos e para se averiguar quais proteínas estão sendo formados para a formação da matriz do biofilme ambiental.

1.5.3.1 Sequenciamento de Regiões Hipervariáveis para Classificação Filogenética

Dentre as nove regiões hipervariáveis algumas possuem maior capacidade de identificação filogenética para linhagens bacterianas em particular, entretanto não há uma região específica para estudo de comunidades complexas, sendo recomendado o uso de várias regiões alvo em conjunto (JEON et al., 2015). Uma simples região hipervariável não pode diferenciar todas as espécies bacterianas, porém o uso concatenado dessas regiões pode melhorar a distinção das espécies bacterianas (CHAKRAVORTY et al., 2007).

Muitos estudos sugerem o uso das regiões V1, V2 e V3 (BROOKS et al., 2015; BARTRAM et al., 2011; KUMAR et al., 2011; PITTA et al., 2014; WERNER et al., 2012), outros apontam que essas regiões superestimam a riqueza de espécies e promovem o uso da região V4, V5 e V6 como o mais apropriado (CAPORASO et al., 2011; GAIDOS, et al., 2011; JEON et al., 2015; KOZICH et al., 2013). No entanto, outros estudos demonstram que a região V7-V8 são mais representativos para caracterização de uma comunidade (KUMAR et al., 2011). Em estudo realizado por Chakravorty (2007) as regiões V2, V3 e V6 mostraram maior poder discriminatório para 110 espécies bacterianas analisadas. Zakrzewski (2012) e Mizrahi-Man (2013) recomendam as regiões V3 e V4 para estudos de comunidades bacterianas.

A escolha dos iniciadores para amplificação da região de interesse também é um fator determinante. A otimização de iniciadores é crucial para alcançar uma representação sem interferências da composição da comunidade microbiana (HUGERTH et al., 2014). Bakke (2011) e Prosdocimi (2013) observaram que a amplificação da região V3 com o iniciador 518R pode levar a erros de interpretação dos resultados em comunidade bacteriana de amostras ambientais, pois esse iniciador possui homologia para genes (SSU rRNA) de eucariotos. Os iniciadores 341F - 534R (região V3) e 341R – 907R (região V3–V5) podem amplificar genes de diferentes organismos, dificultando a análise de dados obtidos por biologia molecular de amostras heterogêneas (HUYS et al., 2008).

1.6 Processamento por bioinformática e banco de dados

Existem diferentes softwares para análises de dados obtidos por sequenciamento de nova geração, os mais utilizados são o QIIME (www.qiime.org), o RDPipeline (www.omictools.com), o MOTHUR (www.mothur.org), o VAMPS (<http://vamps.mbl.edu>), MG-

RAST (www.metagenomics.anl.gov), CloVR-16S (<http://clovr.org>), EzTaxon (<http://www.ezbiocloud.net>), METAGENassist (<http://www.metagenassist.ca>), SnoWMan (www.snowman.genome.tugraz.at) e o MEGAN (www.ab.inf.uni-tuebingen.de). Entretanto, o software QIIME possui algumas vantagens em relação aos demais como: maior precisão na atribuição taxonômica, implementação de testes estatísticos, rapidez de processamento de grande volume de dados, requisito computacional acessível, facilidade de uso e suporte documental (D'ARGENIO et al., 2014; GOODRICH et al., 2013; NILAKANTA et al., 2014; PLUMMER, et al., 2015).

Devido ao fenômeno de reação cruzada de iniciadores de rRNA 16S com genes de eucariotos, muitas sequências anotadas no GenBank para rRNA 16S são atualmente fragmentos de genes de rRNA 18S. Esse fator também pode levar a superestimação da diversidade procariótica, interferência no padrão de bandas em gel de agarose e interferência no monitoramento de populações (HUYS et al., 2008; PROSDOCIMI et al., 2013). A escolha do banco de dados de referência taxonômica pode impactar na fidelidade do resultado de classificação (SOERGEL et al., 2012). Para os resultados gerados por sequenciamento de nova geração podem ser utilizados bancos de dados distintos. O RDP (<http://rdp.cme.msu.edu>) é amplamente utilizado para análise de comunidade microbiana e na microbiologia ambiental, sendo que 85% das sequências bacterianas no RDP são provenientes de isolados diretos de amostras ambientais (COLE et al., 2013). O SILVA (<http://www.arb-silva.de>) é um banco de dados exclusivo para RNA microbiano que é atualizado frequentemente. Além disso, o novo conjunto de dados não-redundante do SILVA faz com que seja uma referência de classificação ideal para sequenciamentos de nova geração (QUAST et al., 2012). O Greengenes (<http://greengenes.lbl.gov>) é um banco de dados dedicado ao rRNA 16S como referência taxonômica, com aproximadamente 406.135 sequências filtradas (maio de 2013). O Greengenes é baseado em uma árvore filogenética *de novo*, isto é, gerado a partir de novas classificações taxonômicas ou transferência de classificações taxonômicas entre bancos de dados (McDONALD et al., 2012). A identificação filogenética é baseada na identidade da sequência (%ID), os limites de 97% e 99% são os mais utilizados para estudos com bactérias (GOODRICH et al., 2013).

Existem diversos trabalhos na literatura que apresentam comparações analíticas entre diferentes softwares (ALBANESE, et al., 2015; MAJANEVA, et al., 2015; PLUMMER, et al., 2015; PUENTE-SÁNCHEZ et al., 2015) e bancos de dados para amostras microbianas (GARDNER, 2012; KOZLOV, et al., 2016; NEWTON; ROESELERS, 2012; WERNER et al., 2012). Entretanto, ainda existem muitas divergências entre os resultados apresentados para cada banco de dados utilizado em diferentes amostras analisadas. Apesar das vantagens apresentadas no uso de sequenciadores de nova geração, é necessário adotar medidas corretivas durante a análise dos dados para eliminar erros, como o uso de limiares de pontuação de alta qualidade para remoção de sequências de baixa qualidade; correção de sequências através de fluxogramas de correção; e o uso de sequências pré-alocadas (PINTO; RASKIN, 2012).

2 CONCLUSÃO

O advento das técnicas moleculares possibilitou grande avanço na compreensão das comunidades microbianas ambientais, tanto na composição da microbiota como na identificação de organismos metabolicamente ativos. Os micro-organismos em amostras ambientais geralmente estão presentes na forma de biofilmes e as características químicas destas amostras são complexas. As numerosas manipulações de ácidos nucleicos entre a etapa de extração e a análise final podem resultar em alteração significativa da proporção de moldes que representam os diferentes organismos das comunidades e desta forma inviabilizar a obtenção de dados confiáveis sobre a composição da comunidade. Assim sendo, processos de validação das diferentes etapas moleculares são de extrema importância para a obtenção de dados fidedignos sobre a característica filogenética da amostra em estudo. A análise detalhada de todas as etapas moleculares obtidas neste trabalho possibilitou identificar os pontos críticos do processo, permitindo otimizações e padronizações que garantiram uma maior fidedignidade dos resultados obtidos.

Para auxiliar na validação de todas as etapas moleculares, foram utilizados consórcios artificiais de bactérias em concentrações e proporções conhecidas. Após testes de eficiência com diferentes protocolos de extração de RNA, o protocolo G (adaptado neste trabalho) foi o que ofereceu as melhores condições de lise e de recuperação do rRNA para as amostras de lodo ativado, tanto quantitativamente quanto qualitativamente. Foi comprovado também o aumento da eficiência na eliminação de contaminantes co-extraídos pelo protocolo G. Essa comprovação foi obtida após a comparação analítica com outro protocolo (F) similar ao adaptado neste trabalho. O protocolo G adaptado nesse trabalho mostrou ser eficaz para extração de RNA de diferentes amostras ambientais *in natura* (lodo ativado, tanques de piscicultura, filtro de areia e biocorrosão) e para amostras ambientais cultivadas em diferentes meios de cultura (bactérias redutoras de ferro e biolixiviação). O teor de matéria orgânica/inorgânica não interferiu na extração de RNA com o protocolo G, entretanto, é necessária a determinação da alíquota de amostra utilizada para a extração. Foi possível recuperar o RNA de lodo armazenado por longos períodos à -80 °C utilizando o protocolo G, mostrando que a conservação em freezer preserva a integridade da molécula de RNA. A otimização da lise celular e da eliminação

de contaminantes do extrato de RNA obtidos no protocolo G possibilitou a redução de vieses nas etapas moleculares “downstream” aplicadas neste trabalho.

Na otimização da RT-PCR, a redução no número de ciclos de amplificação durante a etapa de PCR para 10 ciclos junto com o programa de subciclagem e o uso de uma enzima polimerase de alta fidelidade, minimizou significativamente as variações entre as proporções obtidas e esperadas dos membros dos consórcios artificiais. Estas diferenças na eficiência de amplificação por PCR ocorreram devido ao amplo espectro no teor de G+C de cada espécie bacteriana. Após a análise dos extratos de RNA de lodo ativado obtidos pelos protocolos G e F, foi concluído que a presença de contaminantes co-extraídos interferiu na atividade da enzima polimerase. Na avaliação de aditivos de PCR, o uso das combinações A (Betaína 0.5M + DMSO 5%), C (Buffer Pfu + MgCl₂ 4mM), F (DMSO 7.5%) e H (Betaína 1M) junto da diminuição da rampa de temperatura para 1 °C/s, melhoraram a amplificação por PCR para a amostra de lodo ativado. Ao final da otimização da etapa de RT-PCR, foi concluído que a etapa de extração possui fundamental importância para a manutenção da proporção dos *moldes* de RNA de uma amostra. Mesmo utilizando programas de amplificação otimizados e aditivos, os contaminantes co-extraídos podem influenciar negativamente em todo o processo de análise molecular, até mesmo em análises que empregam o uso de digestão por enzimas de restrição.

A padronização de ARDRA revelou que são necessários entre 500 e 550 clones para uma cobertura completa da diversidade de micro-organismos. A utilização de enzimas misturadas na mesma reação de digestão “*double digest*”, proporcionou a separação dos clones utilizando 2 etapas de ARDRA porém, para uma separação mais fidedigna, uma terceira etapa de ARDRA é recomendada. O critério de corte em 3% possibilitou selecionar somente os grupos filogenéticos com maior representatividade. Após o sequenciamento dos clones da amostra de lodo ativado foi concluído que o agrupamento realizado na segunda etapa de ARDRA é efetivo para a separação das espécies da amostra, porém é recomendado uma terceira digestão para eliminação de eventuais clones não pertencentes ao agrupamento correspondente. O resultado obtido mostrou que este método é válido para a detecção e agrupamento de micro-organismos em amostras complexas, porém, o mesmo se mostrou laborioso. Mesmo com o surgimento de novas tecnologias moleculares, o uso da técnica de ARDRA ainda é utilizada para caracterização filogenética. Após as

padronizações da etapa de clonagem e ARDRA foi observado que os contaminantes co-extraídos pelo protocolo F interferiram na atividade das enzimas de restrição. Assim, foi concluído que o uso do protocolo G junto da padronização aplicada para o ARDRA possibilitou a detecção de todos os membros do consórcio artificial.

Foi utilizado o sequenciamento de nova geração para identificação dos micro-organismos metabolicamente ativos em amostras complexas. O uso da ferramenta QIIME no pré-tratamento dos dados permitiu o melhoramento dos resultados obtidos, aumentando a resolução e profundidade taxonômica. Na correlação entre os resultados do sequenciamento do rRNA 16S (metabolicamente ativos) e do DNA (abundância em número de células) para lodo ativado, foi observado que dentre os principais gêneros responsáveis por produzir EPS (*Acinetobacter*, *Acidovorax* e *Zoogloea*) somente o gênero *Acidovorax* apresentou grande abundância na amostra.

Após a análise detalhada de todas as técnicas moleculares empregadas neste trabalho, foi concluído que a validação é uma etapa essencial para preservar a verdadeira abundância filogenética e na identificação os organismos metabolicamente ativos de maior relevância em uma amostra. Trabalhos que negligenciam esta etapa podem introduzir vieses que comprometeram os resultados, gerando uma visão distorcida sobre a verdadeira composição microbiana estudada. No anexo 4 foi descrito um fluxograma de recomendação, contendo cada etapa validada e padronizada. No aspecto geral, para garantir a qualidade dos dados para uma amostra desconhecida, é recomendado adotar os seguintes controles: processar a amostra logo após a coleta; padronizar o volume da alíquota para cada tipo de amostra a ser extraída com o protocolo G; para PCR, utilizar diferentes combinações de iniciadores, enzima polimerase de alta fidelidade, combinação de aditivos, diminuição da rampa de temperatura em 1 °C/s, 10 ciclos de amplificação e programa de subciclagem; para ARDRA, utilizar entre 500 e 550 clones, combinação de enzimas de restrição, 3 ciclos de agrupamento e adotar o critério de corte em $\geq 3\%$; para sequenciamento de nova geração, utilizar kits de preparo da amostra que não utilizem etapas de PCR, adotar o pré-tratamento manual dos dados e utilizar mais de um banco de dados. Embora a validação de todo o processo molecular tenha melhorado significativamente os resultados sobre a caracterização filogenética de organismos metabolicamente ativos é necessário que seja feito em estudos futuros uma avaliação das

vias metabólicas ativas, utilizando o mRNA como molécula alvo, com o intuito de gerar uma visão detalhada da atividade em comunidades microbianas.

7 REFERÊNCIAS*

- ALBANESE, D.; FONTANA, P.; FILIPO, C.; CAVALIERI, D. & DONATI, C. MICCA: a complete and accurate software for taxonomic profiling of metagenomic data. **Scientific Reports**, v.5, 7 p., 2015.
- ALMEIDA, R.N.A. **Desenvolvimento, validação e aplicação de método molecular baseado na extração, amplificação e sequenciamento do rRNA para a identificação das bactérias formadoras de biofilme na superfície de membranas de osmose reversa**. 2009. 122 f. Dissertação (Mestrado em Biotecnologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2009.
- BAE, H., KIM, H., JEONG, S., LEE, S. Changes in the relative abundance of biofilm-forming bacteria by conventional sand-filtration and microfiltration as pretreatments for seawater reverse osmosis desalination. **Desalination**, v. 273, p. 258-266, 2011.
- BAKKE, I.; DE SCHRYVER, P.; BOON, N.; VADSTEIN, O. PCR-based community structure studies of Bacteria associated with eukaryotic organisms: A simple PCR strategy to avoid co-amplification of eukaryotic DNA. **Journal of Microbiology Methods**, v, 84, p. 349-351, 2011.
- BAKER, B. J.; BANFIELD, J. F. Microbial communities in acid mine drainage. **FEMS Microbiology Ecology**, v. 44, p. 139-152, 2003.
- BAKKE, I.; COWARD, E.; ANDERSEN, T. and VADSTEIN, O. Selection in the host structures the microbiota associated with developing cod larvae (*Gadus morhua*). **Environmental Microbiology**, v. 17, n. 10, p. 3914-3924, 2015.
- BARBOSA, H. R. **Microbiologia básica**. São Paulo: Editora Atheneu. 1998.
- BAUTISTA-DE LOS SANTOS, Q.; SCHROEDER, J. L.; BLAKEMORE, O.; MOSES, J.; HAFLEY, M.; SLOAN, W. and PINTO, A. J. The impact of sampling, PCR, and sequencing replication on discerning changes in drinking water bacterial community over diurnal time-scales. **Water Research**, v. 900, p. 216-224, 2016.
- BELILA, A.; EL-CHAKHTOURA, J.; OTAIBI, O.; MUYZER, G.; GONZALES-GIL, G.; SAIKALY, P. E.; van LOOSDRECHT, M. C. M. and VROUWENVELDER, J. S. Bacterial community structure and variation in a full-scale seawater desalination plant for drinking water production. **Water Research**, v. 94, p. 62-72, 2016.
- BINDER, B. J.; LIU, Y. C. Growth rate regulation of rRNA content of a marine *Synechococcus* (cyanobacterium) strain. **Applied Environmental Microbiology**, v. 64, p.3346-3351, 1988.

BOTEVA, S.; KENAROVA, A.; RADEVA, G.; TRAYKOV, I.; BOGOEV, V. Community Dynamics of Pelagic Bacteria in the High Mountain Lake Bubreka, Rila Mountain, Bulgaria. **Tome**, v. 66, n. 11, 2013.

BOURREIN, M., ACHOUAK, W., URBAIN, V., HEULIN, T. DNA extraction from activated sludges. **Current Microbiology**, v. 38, p. 315-319, 1999.

BLAZEWICZ, S. J.; BARNARD, R. L.; DALY, R. A. and FIRESTONE, M. K; Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. **The ISME Journal**, v. 7, p. 2061-2068, 2013.

BRAUN, B.; SCHRÖDER, J.; KNECHT, H. and SZEWZYK, U. Unraveling the microbial community of a cold groundwater catchment system. **Water Research**, v. 107, p. 113-126, 2016.

BRISCO, M. J. and MORLEY, A. A. Quantification of RNA integrity and its use for measurement of transcript number. **Nucleic Acids Research**, v. 40, n. 18, 2012.

BROCCHIERI, L. The GC content of bacterial genomes. **Journal Phylogen. Evol. Biol.**, v.2, n. 1, 2014.

BROOKS, J. P.; EDWARDS, D.; HARWICH JR, M. D.; RIVERA, W. C.; FETTWEIS, J. M.; SERRANO, M. G.; RERIS, R. A.; SHETH, N. U.; HUANG, B.; GIRERD, P., Vaginal Microbiome Consortium (add members), DTRAUSS, J. F.; JEFFERSON, K. K. and BUCK, G. A. The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. **BMC Microbiology**, v. 15, n. 66, 14 p., 2015.

BRUNK, C. F., AVANISS-AGHAJANI, E. and BRUNK, C. A. A computer analysis of primer and probe hybridization potential with bacterial small-subunit rRNA sequences. **Applied and Environmental Microbiology**, v. 62, n. 3, p.872-879, 1996.

CAI, L.; CHEN, T-B.; GAO, D.; YU, J. Bacterial communities and their association with the bio-drying of sewage sludge. **Water Research**, v. 90, p. 44-51, 2016.

CAI, W.; LI, Y.; NIU, L.; ZHANG, W.; WANG, C.; WANG, P.; MENG, F. New insights into the spatial variability of biofilm communities and potentially negative bacterial groups in hydraulic concrete structures. **Water Research**, v. 123, n. 15, p. 495-504, 2017.

CAMPBELL, B. J., YU, L., HEIDELBERG, J. F. and KIRCHMAN, D. L. Activity of abundant and rare bacteria in a coastal ocean. **PNAS**, v. 108, n. 31, 2011.

CAPORASO, J. G.; LAUBER, C. L.; WALTERS, W. A.; BERG-LYONS, S.; POZUPONE, C. A.; TUMBAUGH, P. J.; FIERER, N. and KNIGHT, R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *PNAS*, v. 108, p. 4516-4522, 2011.

CHAKRAVORTY, S.; HELB, D.; BURDAY, M.; CONNELL, N. and ALLAND, D. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. **JOURNAL Microbiology Methods**. v.69, n.2, p.330-339, 2007.

CHEN, H.; WAN, J.; CHEN, K.; LOU, G.; FAN, J.; CLARK, J.; ZHANG, S. Biogas production from hydrothermal liquefaction wastewater (HTLWW): Focusing on the microbial communities as revealed by high-throughput sequencing of full-length 16S rRNA genes. **Water Research**, v. 106, p. 98-107, 2016.

COSTERTON, J. W., LEWANDOWSKI, Z., DEBEER, D., CALDWELL, D., KORBER, D., JAMES, G. Biofilms, the Customized Microniche. **Journal of Bacteriology**, p. 2137-2142, 1994.

COLE, J. R.; WANG, W.; FISH, J. A.; CHAI, B.; MCGARRELL, D. M.; SUN, Y.; BROWN, T.; PORRAS-ALFARO, A.; KUSKE, C. R. and TIEDJE, J. M. Ribosomal Database Project: data and tools for high throughput rRNA analysis. **Nucleic Acids Research**, v. 42, p. 633-642, 2013.

DANG, H., LOVELL, C. R., Numerical Dominance and Phylotype Diversity of Marine *Rhodobacter* Species during Early Colonization of Submerged Surfaces in Coastal Marine Waters as Determined by 16S Ribosomal DNA Sequence Analysis and Fluorescence In Situ Hybridization. **Applied and Environmental Microbiology**, v. 68, p. 496-504, 2001.

D'ARGENIO, V.; CASABURI, G.; PRECONE, V. and SALVATORE, F. Comparative Metagenomic Analysis of Human Gut Microbiome Composition Using Two Different Bioinformatic Pipelines. **BioMed Research Int.**, ID 325340, 10 p., 2014.

DELAFONT, V.; BOUCHON, D.; HÉCHARD, Y. and MOULIN, L. Environmental factors shaping cultured free-living amoebae and their associated bacterial community within drinking water network. **Water Research**, v. 100, p. 382-392, 2016.

DOHERTY, M.; YAGER, P. L.; MORAN, M. A.; COLES, V. J.; FORTUNATO, C. S.; KRUSCHE, A. V.; MEDEIROS, P. M.; PAYET, J. P.; RICHEY, J. E.; SATINSKY, B. M.; SAWAKUCHI, H. O.; WARD, N. D. and CRUMP, B. C. Bacterial Biogeography across the Amazon River-Ocean Continuum. **Frontiers in Microbiology**, v. 8, n. 882, 2017.

DUTTON, C. M.; PAYNTON, C. and SOMMER, S. S. General method for amplifying regions of very high G+C content. **Nucleic Acids Research**, v. 21, n. 12, p. 2953-2954, 1993.

ECKERT, K. A., KUNKEL, T. A. DNA polymerase fidelity and the polymerase chain reaction. **PCR Methods Applied**, v. 1, p. 17-24, 1991.

ELSER, J.J., FAGAN, W. F., DENNO, R. F., DOBBERFUHL, D.R., FOLARIN, A., HUBERTY, A., INTERLANDI, S., KILHAM, S. S., McCAULEY, E., SCHULZ, K. L., SIEMANN, E. H., STERNER, R. W. Nutritional constraints in terrestrial and freshwater food webs. **Nature**, v. 408, p. 578-580, 2000.

EIGNER, J., BOEDTKER, H., MICHAELS, G. The thermal degradation of nucleic acids. **Biochim. Biophys. Acta**, v. 51, p. 163 – 165, 1961.

FELSKE, A. A.; WOLTERINK, R.; VAN LIS, W. M. D. V.; AKKERMANS, A. D. L. Response of a soil bacterial community to grassland succession as monitored by 16S rRNA levels of the predominant ribotypes. **Applied Environment Microbiology**, v. 66, p. 3998-4003, 2000.

FEGATELLA, F., LIM, J., KJELLENBERG, S. and CAVICCHIOLI, R. Implications of rRNA operon copy number and ribosome content in the marine oligotrophic ultramicrobacterium. **Applied And Environmental Microbiology** v.64, n.11, p.4433-4438, 1998.

FERRIS, M. J.; WARD, D. M. Seasonal distributions of dominant 16S rRNA-defined populations in a hot spring microbial mat examined by denaturing gradient gel electrophoresis. **Applied Environment Microbiology**, v. 63, p. 1375–1381, 1997.

FISCHER, M. A.; GÜLLERT, S.; NEULINGER, S. C.; STREIT, W. R. and SCHMITZ, R. A. Evaluation of 16S rRNA Gene Primer Pairs for Monitoring Microbial Community Structures Showed High Reproducibility within and Low Comparability between Datasets Generated with Multiple Archaeal and Bacterial Primer Pairs. **Frontiers in Microbiology**, v. 7, n. 1297, 2016.

GARDNER, D. P.; XU, W.; MIRANKER, D. P.; OZER, S.; CANNONE, J. J. and GUTELL, R. R. An Accurate Scalable Template-based Alignment Algorithm. **IEEE Int. C. Bioinfo. Biomed**, p. 1-7, 2012.

GAIDOS, E.; RUSCH, A. and ILARDO, M. Ribosomal tag pyrosequencing of DNA and RNA from benthic coral reef microbiota: community spatial structure, rare members and nitrogen-cycling guilds. **Envirom. Microbiology**, v. 13, n. 5, p. 1138-1152, 2011.

GIOVANNONI S. J.; BRITSCHGI T. B.; MOYER C. L.; FIELD K. G. Genetic diversity in Sargasso Sea bacterioplankton. **Nature**, v. 345, p. 60-63, 1990.

GOODRICH, J. K.; DI RIENZI, S. C.; POOLE, A. C.; KOREN, O.; WALTERS, W. A.; CAPORASO, J. G.; KNIGHT, R. and LEY, R. E. Conducting a Microbiome Study. **Cell**, v. 158, p. 2500-2512, 2014.

GUO BIN, S.; WEN BIAO, J.; LAM, E. K. H.; XING, X. Purification of total DNA extracted from activated sludge. **Journal Environmental Sci** 20, 80–87, 2007.

GYORFY, Z., DRASKOVITS, G., VERNYIK, V., BLATTNER, F. F., GAAL, T. and POSFAI, G. Engineered ribosomal RNA operon copy-number variants of *E. coli* reveal the evolutionary trade-offs shaping rRNA operon number. **Nucl. Acids Resear.**, v.3, 2015.

HORI, M., FUKANO, H., SUZUKI, Y. Uniform amplification of multiple DNAs by emulsion PCR. **Biochemical Biophys. Research Comm.**, v. 352, n. 2, p. 323-328, 2007.

HONG, P.Y., HWANG, C., LING, F., ANDERSEN, G. L., LeCHEVALLIER, M. W., LIU, W. Pyrosequencing Analysis of Bacterial Biofilm Communities in Water Meters of a Drinking Water Distribution System. **Applied and Environmental Microb**, v. 76, p. 5631-5635, 2010.

HUYS, G.; VANHOUTTE, T.; JOOSSENS, M.; MAHIOUS, A. S.; DE BRANDT, E.; VERMEIRE, S. SWINGS, J. Coamplification of Eukaryotic DNA with 16S rRNA Gene-Based PCR Primers: Possible Consequences for Population Fingerprinting of Complex Microbial Communities. **Curr. Microbiology**, v. 56, p. 553-557, 2008.

INGRASSIA, I., ROQUES, C., PRÉVOTS, F. Preliminary experiments for ARDRA validation on flora associated with intestinal mucosa. **Lait**, v. 81, p. 263-280, 2001.

JANG, I-A.; KIM, J. & PARK, W. Endogenous hydrogen peroxide increases biofilm formation by inducing exopolysaccharide production in *Acinetobacter oleivorans* DR1. **Scientific Reports**, v. 6:21121, 2016.

JEON, Y. S.; PARK, S. C.; LIM, J.; CHUN, J and KIM, B. S. Improved pipeline for reducing erroneous identification by 16S rRNA sequences using the Illumina MiSeq platform. **Journal of Microbiology**, v. 53, n. 1, p. 60-69, 2015.

JEONG, S-Y.; YI, Y.; LEE, C-H.; KIM, T. G. Spatiotemporal dynamics and correlation networks of bacterial and fungal communities in a membrane bioreactor. **Water Research**, v. 105, p. 218-230, 2016.

JI, B.; WEI, L.; CHEN, D.; WANG, H.; LI, Z.; YANG, K. Domestic wastewater treatment in a novel sequencing batch biofilm filter. **Applied Environmental Biotechnology**, v. 99, p. 5731-5738, 2015.

JIN, M.; ZHAO, Z. G.; QUI, Z. G.; WANG, J. F.; CHEN, Z. L.; SHEN, Z. Q.; LI, C.; WANG, X. W.; DONG, X. W.; LI, J. W. Rapid method to extract high-quality RNA from activated sludge. **Huanjing Kexue/Environmental Science**, v. 31, p. 260-265, 2010.

JUNIOR, F. B. R., MENDES, I. C., TEIXEIRA, K. R. S., REIS, V. M. **Uso de ferramentas moleculares em estudos da diversidade de micro-organismos do solo**. Planatina, DF: Embrapa Cerrados, 2002, 33 p.

KANAGAWA, T. Bias and Artifacts in Multitemplate Polymerase Chain Reactions (PCR). **Journal of Bioscience and Bioengineering**, v. 96, n. 4, p. 317-323, 2003.

KARPINETS, T. V., GREENWOOD, D. J., SAMS, C. E. and AMMONS, J. T. RNA: protein ratio of the unicellular organism as a characteristic of phosphorous and nitrogen stoichiometry and of the cellular requirement of ribosomes for protein synthesis. **BMC Biology**, v.4, n.30, 2006.

KEER J. T.; BIRCH L. Molecular methods for the assessment of bacterial viability. **Jornal of Microbiology Methods**, v. 53, p. 175-183, 2003.

KERKHOF, L., KEMP, P. Small ribosomal RNA content in marine Proteobacteria during non-steady-state growth. **FESM Microbiology Ecology**, v. 30, p. 253-260, 1999.

KIM, M. H.; HAO, O. J., WANG, N. S. Acinetobacter isolates from different activated sludge processes: characteristics and neutral network identification. **FEMS Microbiology Ecology**, v, 23, p. 217-227, 1997.

KLAPPENBACH, J. A., DUNBAR, J. M. and SCHMIDT, T. M. rRNA operon copy number reflects ecological strategies of bacteria. **Applied and Enviro. Microbiology**, v. 66, n.4, p. 1328-1333, 2000.

KEBSCHULL, J. M. and ZADOR A. M. Sources of PCR-induced distortions in high-throughput sequencing data sets. **Nucleic Acids Research**, v. 43, n. 21, 2015.

KOZDRÓJ, J., van ELSAS, J. D. Pllication of polymerase chain reaction-denaturing gradient gel electrophoresis for comparison of direct and indirect extraction methods of soil DNA used for microbial community fingerprinting. **Biol. Fertil Soils**, v. 31, p. 372-378, 2000.

KOZICH, J. J.; WESTCOTT, S. L.; BAXTER, N. T.; HOGHLANDER, S. K.; SCHLOSS, P. D. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the

MiSeq Illumina Sequencing Platform. **Applied and Environmental Microbiology**, v. 79, n. 17, p. 5112-5120, 2013.

KOZLOV, A. M.; ZHANG, J.; YILMAZ, P.; GLÖCKNER, F. O. and STAMATAKIS, A. Phylogeny-aware identification and correction of taxonomically mislabeled sequences. **Nucleic Acids Research**, p. 12, 2016.

LAJIN, B.; ALACHKAR, A.; SAKUR, A. A. Betaine significantly improves multiplex tetra-primer ARMS-PCR methods. **Mol. Biotechnol.**, v. 54, p. 977-982, 2013.

LANE, A. N. Solution conformation and dynamics of the octadeoxy-nucleotide d(CACTAGTG)₂: A multinuclear n.m.r. relaxation study. **Carbohydrate Research**, v. 221, p. 123-144, 1991.

LEE, D. H.; ZO, Y. G.; KIM, S. J. Nonradioactive method to study genetic profiles of natural bacterial communities by PCR single- strand-conformation polymorphism. **Environment of Microbiology**, v. 62, p. 3112–3120, 1996.

LEVER, M. A.; TORTI, A.; EICKENBUSCH, P.; MICHAUD, A. B.; SAINT-TEMKIV, T. and JORGENSEN, B. B. A modular method for the extraction of DNA and RNA, and the separation of DNA pools from diverse environmental sample types. **Frontiers in Microbiology**, v. 6, Art. 476, 2015.

LIU, Q. and SOMMER, S. S. Amplification of Regions with High and Low GC Content: Application to the Inversion Hotspot in the Factor VIII Gene. **BioTechniques**, v. 25, n. 6, 1998.

LIU, H., FANG, H. H. P., Extration of extracellular polymeric substances (EPS) of sludges. **Journal of Biotech**, v. 95, p. 249 – 246, 2002.

LIU, G.; BAKKER, G. L.; LI, S.; VREEBURG, J. H. G.; VERBERK, J. Q. J. C.; MEDEMA, G. J.; LIU, W. T. and VAN DIJK, J. C. Pyrosequencing Reveals Bacterial Communities in Unchlorinated Drinking Water Distribution System: An Integral Study of Bulk Water, Suspended Solids, Loose Deposits, and Pipe Wall Biofilm. **Environmental Sci. & Tech.**, v. 48, p. 5467-5476, 2014.

LIN, W., LI, J., SCHÜLER, D., JOGLER, C., PAN, Y. Diversoty analisys of magnetotactic bacteria in Lake Liyun, northern China, by restriction fragment length polymorphism. **Syst. and Applied Microb**, v. 32, n. 2, p. 342-350, 2009.

LIANG, Z., DRIJBER, R. A., LEE, D. J., DWIEKAT, I. M., HARRIS, S. D. and WEDIN, D. A. A DGGE-cloning method to characterize arbuscular mycorrhizal community structure in soil. **Soil Biol. and Biochemical**, v. 40, n. 4, p. 956-966, 2008.

LIESACK, W.; WEYLAND, H.; STACKEBRANDT, E. Potential risks of gene amplification by PCR as determined by 16S rDNA analysis of a mixed-culture of strict barophilic bacteria. **Microbial Ecology**, v. 21, p. 191-198, 1991.

MAJANEVA, M.; HYYTIÄINEN, K.; VARVIO, S. L.; NAGAI, S.; BLOMSTER, J. Bioinformatic Amplicon Read Processing Strategies Strongly Affect Eukaryotic Diversity and the Taxonomic Composition of Communities. **PLOS ONE**, v. 10, n. 6, 2015.

MATAR, G. K.; BAGCHI, S.; ZHANG, K.; OERTHER, D. B.; SAIKALY, P. E. Membrane biofilm communities in full-scale membrane bioreactors are not randomly assembled and consist of a core microbiome. *Water Research*, v. 123, p. 124-133, 2017.

MATHESON, C. D.; GURNEY, C.; ESAU, N. and LEHTO, R. Assessing PCR inhibition from humic substances. **The Open Enzyme Inhib. Journal**, v. 3, p. 38-45, 2010.

McCARTHY, D. T.; JOVANOVIĆ, D.; LINTERN, A.; TEAKLE, I.; BARNES, M.; DELETIC, A.; COLEMAN R.; ROONEY, G.; PROSSER, T.; COUTTS, S.; HIPSEY, M. R.; BRUCE, L. C.; HENRY, R. Source tracking using microbial community fingerprints: Method comparison with hydrodynamic modelling. **Water Research**, v. 109, p. 253-265, 2017.

McDONALD, D.; PRICE, M.; GOODRICH, J.; NAWROCKI, E. P.; DeSANTIS, T.; PROBST, A.; ANDERSEN, G. L.; KNIGHT, R. and HUGENHOLTZ. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. **The ISME Journal**, v. 6, p. 610-618, 2012.

McINERNEY, P.; ADAMS, P. and HADI, M. Z. Error Rate Comparison during Polymerase Chain Reaction by DNA Polymerase. **Hindawi**, 8 p., ID 287430, 2014.

MEYERHANS, A.; VARTANIAN, J. P.; WAIN-HOBSON, S. DNA recombination during PCR. **Nucleic Acids Research**, v. 18, p. 1687-1691, 1990.

MIZRAHI-MAN, O.; DAVENPORT, E. R.; GILAD, Y. Taxonomic Classification of Bacterial 16S rRNA Genes Using Short Sequencing Reads: Evaluation of Effective Study Designs. **PLOS One**, v. 8, n. 1, 2013.

MORGAN C. A.; HUDSON A.; KONOPKA, A.; NAKATSU C. H. Analyses of microbial activity in biomass-recycle reactors using denaturing gradient gel electrophoresis of 16S rDNA and 16S rRNA PCR products. **Canadian Journal of Microbiology**, v. 48, p. 333-341, 2002.

MOESENEDER, M. M.; ARRIETA, J. M.; MUYZER, G.; WINTER, C.; HERNDL, G. J. Optimization of terminal-restriction fragment length polymorphism analysis for complex marine bacterioplankton communities and comparison with denaturing gradient gel electrophoresis. **Environment of Microbiology**, v. 65, p. 3518-3525, 1999.

MOYER, C. L.; TIEDJE, J. M.; DOBBS, F. C.; KARL, D. M. A Computer-Simulated Restriction Fragment Length Polymorphism Analysis of Bacterial Small-Subunit rRNA Genes: Efficacy of Selected Tetrameric Restriction Enzymes for Studies of Microbial Diversity in Nature. **Applied and Environmental Microbiology**, v. 62, n. 7, p. 2501-2507, 1996.

MOYER, C. L. Molecular Phylogeny: Applications and Implications for Marine Microbiology. **Methods in Microbiology**, v. 30, p. 375-394, 2001.

MUYZER, G.; WALL, E. C.; UITTERLINDEN, A. G. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. **Applied and Environmental Microbiology**, v. 59, n. 3, p. 695-700, 1993.

MUYZER, G.; SMALLA, K. Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. **International Journal of General and Molecular Microbiology**, v. 73, p. 127-141, 1998.

MYERS, R. M.; FISCHER, S. G.; LERMAN, L. S.; MANIATIS, T. Nearly all single base substitutions in DNA fragments joined to a GCclamp can be detected by denaturing gradient gel-electrophoresis. **Nucleic Acids Research**, v. 13, p. 3131-3145, 1985.

NAZINA, T. N.; SHESTAKOVA, N. M.; SEMENOVA, E. M.; KORSHUNOVA, A. V.; KOSTRUKOVA, N. K.; TOUROVA, T. P.; MIN, L.; FENG, Q. and POUTARAUS, A. B. Diversity of Metabolically Active Bacteria in Water-Flooded High-Temperature Heavy Oil Reservoir. **Frontiers in Microbiology**, v. 8, n. 707, 2017.

NEWTON, I. L. G. and ROESELERS, G. The effect of training set on the classification of honey bee gut microbiota using the Naïve Bayesian Classifier. **BMC Microbiology**, v. 12, n. 221, 2012.

NIKOLAUSZ, M.; MÁRIALIGETI, K.; KOVÁCS, G. Comparison of RNA and DNA-based species diversity investigations in rhizoplane bacteriology with respect to chloroplast sequence exclusion. **Journal of Microbiology Methods**, v. 56, p. 365-373, 2004.

NILAKANTA, H.; DREWS, K. L.; FIRRELL, S.; FOULKES, M. A.; JABLONSKI, K. A. A review of software for analyzing molecular sequences. **BMC Research Notes**, v. 7, p. 9, 2014.

NOVITSKY, J.A. Degradation of dead microbial biomass in a marine sediment. **Appl Environ Microbiol**, v. 52, p. 504–509, 1986.

PALLANSCH, L., BESWICK, H., TALIAN, J., ZELENKA, P. Use of an RNA folding algorithm to choose regions for amplification by the polymerase chain reaction. **Analitic Biochemistry**, v. 185, p. 57-62, 1990.

PARK, J-S. Bacterial diversity of the Marine Sponge, *alichondria panicea* by ARDRA and DGGE. **Korean J. of Microbiology**, v. 51, n. 4, p. 398-406, 2015.

PIERANGELA, C., FRANZETTI, A., GANDOLFI, I., GUERRINI, E., BESTETTI, G. Bacterial DGGE fingerprints of biofilmes on electrodes of membranesless microbial fuel cells. **Internat. Biodeterio. & Biodegrad.**, v. 84, p. 211-219, 2013.

PINTO, A. J., RASKIN, L. PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. **PLoS ONE**. v.7, n. 8, 2012.

PITERINA, A. V. and PEMBROKE, J. T. Use of PCR-DGGE based molecular methods to analyse microbial community diversity and stability during the thermophilic stages of an ATAD wastewater sludge treatment process as an aid to performance monitoring. **ISRN Biotechnology**, p. 13, 2013.

PITTA, D. W.; PARMAR, N.; PATEL, A. K.; INDUGU, N.; KUMAR, S.; PRAJAPATHI, K. B.; PATEL, A. B.; REDDY, B.; JASHI, C. Bacterial Diversity Dynamics Associated with Different Diets and Different Primer Pairs in the Rumen of Kankrej Cattle. **PLoS ONE**, v. 9, n. 11, 2014.

PLUMMER, E.; TWIN, J.; BULACH, D. M.; GARLAND, S. M. and TABRIZI, S. N. A Comparison of Three Bioinformatics Pipelines for the Analysis of Preterm Gut Microbiota using 16S rRNA Gene Sequencing Data. **Journal of Proteo. & Bioinfo.**, v. 8, p. 2830291, 2015.

POITELON, J-B.; JOYEUX, M.; WELTÉ, B.; DUGUET, J-P.; PRESTEL, E.; LESPINET, O. and DUBOW, M. S. Assessment of phylogenetic diversity of bacterial microflora in drinking water using serial analysis of ribosomal sequence tags. **Walter Research**, v. 43, p. 4197-4206, 2009.

POLZ, M. F.; CAVANAUGH, C. M. Bias in template-to-product ratios in multitemplate PCR. **Applied and Environmental Microbiology**, v. 64, n. 10, p. 3724-3730, 1998.

PROSDOCIMI, E. M., NOVATI, S., BRUNO, R., BANDI, C., MULATTO, P., GIANNICO, R., CASIRAGHI, M., FERRI, E. Errors in ribosomal sequence datasets generated using PCR-coupled 'panbacterial'

pyrosequencing, and the establishment of an improved approach. **Mol. And Cell. Probes.**, v. 27, p. 65-67, 2013.

PUROHIT, H. J., KAPLEY, A., MOHARIKAR, A. A., NARDE, G. A novel approach for extraction of PCR-compatible DNA from activated sludge samples collected from different biological effluent treatment plants. **Journal of Microbiology Methods**, v. 52, p. 315-323, 2003.

PUENTE-SÁNCHEZ, F.; AGUIRRE, J. and PARRO, V. A novel conceptual approach to read-filtering in high-throughput amplicon sequencing studies. **Nucleic Acids Research**, p. 11, 2015.

QUAST, C.; PRUESSE, E.; YILMAZ, P.; GERKEN, J.; SCHWEER, T.; YARZA, P.; PEPLIES, J. and GLÖCKNER. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. **Nucleic Acid Research**, v. 28, 2012.

QIN Y., et. al. Population dynamics of ammonia-oxidizing bacteria in an aerated submerged biofilm reactor for micropolluted raw water pretreatment. **Springer Science**, v. 79, p. 135-145, 2008.

QIU, X., WU, L., HUANG, H., McDONEL, E. E., PALUMBO, A. V., TIEDJE, J. M., ZHOU, J. Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning.. **Applied and Environmental Microbiology**, v. 67, p.880-887, 2001.

RAMAGOPAL, S. Metabolic changes in ribosomes of *Escherichia coli* during prolonged culture in different media. **Eur. J. Biochemical**, v. 140, p. 353-361, 1984.

RASTOGI, G. and SANI, R. K. Molecular techniques to assess microbial community structure, function, and dynamics in the environment. **Microbes and Microbiology Tech.**, p. 29-57, 2011.

REYSENBACH, A. L.; GIVER, L. J.; WICKHAM G. S.; PACE, N. R. Differential amplification of rRNA genes by polymerase chain reaction. **Applied and Environmental Microbiology**, v. 58, p. 3417-3418, 1992.

REUTER, J. A.; SPACEK, D. V. and SNYDER, M. P. High-throughput sequencing technologies. **Molecular Cell**, v. 58, p. 587-597, 2015.

RUANO, G.; KIDD, K. K. Modeling of heteroduplex formation during PCR from mixtures of DNA templates. **PCR Methods Applied**, v. 2, p. 112–116, 1992.

SALIPANTE, S. J.; KAWASHIMA, T.; ROSENTHAL, C.; HOOGESTRAAT, D. R.; CUMMINGS, L. A.;

SANZ, J. L.; KOCHLING, T. Molecular biology techniques used in wastewater treatment: An overview. **Process Biochemistry** 42, p. 119–133, 2007.

SARKAR, G.; KAPELNER, S.; SOMMER, S. S. Formamide can dramatically improve the specificity of PCR. **Nucleic Acids Reserch**, v. 18, p. 7465, 1990.

SCHNEIDER, D.; WEMHEUER, F.; PFEIFFER, B.; WEMHEUER, B. Extraction of Total DNA and RNA from Marine Filter Samples and Generation of a cDNA as Universal Template for Marker Gene Studies. **Methods in Molecular Biology**, v. 1539, 2017.

SCHUT, F., GOTTSCHAL, J. C., PRINS, R. A. Isolation and characterization of the marine ultramicrobacterium *Sphingomonas* sp. strain RB2256. **FEMS Microbiology Rev.**, v. 20, p. 363-369, 1997.

SEKIGUCHI, H., TOMIOKA, N., NAKAHARA, T. & UCHIYAMA, H. A single band does not always represent single bacterial strains in denaturing gradient gel electrophoresis analisys. **Biotechnology Letter**, v. 23, p. 1205-1208, 2001.

SMETS, W.; LEFF, J. W.; BRADFORD, M. A.; McCULLEY, R. L.; LEBEER, S.; FIERER, N. A method for simultaneous measurement of soil bacterial abundances and community composition via 16S rRNA gene sequencing. **Soil Bio. & Biochemical**, v. 96, p. 145-151, 2016.

SIMONOVIC, A.; TRIFUNOVIC, M.; RASPOR, M.; CINGEL, A.; BOGDANOVIC, M.; DRAGICEVIC, M. Dimethyl sulfoxide improves sensitivity and specificity of RT-PCR and QRT-PCR amplification of low-expressed transgenes. **Arch. Biol. Science**, v. 63, n. 3, p.865-976, 2012.

SILVA, L. J. **Biocorrosion on metal surfaces in hydroelectric power plant**. 2015. 229 p. Ph.D. Thesis (Microbiology) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2015.

SILVA, C. L. V.; SILVA, A. L. S.; MEDEIROS, S. R. B.; LIMA, L. F. A. L.; BLAHA, C. A. G. Detecção de bactérias redutoras de Fe (III), em mangue da bacia petrolífera Potiguar com potencial biodegradador de petróleo. **4° PDPETRO**, p. 21-24, 2007.

SIPOS, R., SZÉKELY, A., RÉVÉSZ, S., MÁRIALIGETI, K. Addressing PCR biases in envivonmental microbiology studies. Bioremediation. **Methods in Molecular Biology**, v. 599, p. 37, 2010.

SOBEK, J. M., CHARBA, J. F. and FOUST, W. N. Endogenous metabolism of *Azotobacter agilis*. **Journal of Bact.**, v. 92, n. 3, p.687-695, 1966.

SOERGEL, D. A. W.; DEY, N.; KNIGHT, R. and BRENNER, S. E. Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. **The ISME Journal**, v. 6, p. 1440-1444, 2012.

SOGIN, M. L. Microbial diversity in deep sea and the underexplored “rare biosphere”. **Proc. Natl. Acad. Sci. USA**, v. 103, p. 12115-12120, 2006.

SONG, J., CHOO, Y. J. and CHO, J. C. *Perlucidibaca piscinae* gen. nov., sp. nov., a freshwater bacterium belonging to the family Moraxellaceae. **Intern. Journal of Syst. And Evol. Microbiology**, v. 58, p. 97-102, 2008

STOECK, T. and EPSTEIN, S. Novel eukaryotic lineages inferred from small-subunit rRNA analyses of oxygen-depleted marine environments. **Applied and Environmental Microbiology**, v. 69, n. 5, p. 2657-2663, 2003.

SUZUKI, M. T.; GIOVANNONI, S. J. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. **Applied and Environmental Microbiology**, v. 62, p. 625–630, 1996.

TAO, H., BAUSCH, C., RISHMOND, C., BLATTNER, F. R. and CONWAY, T. Functional genomics: Expression analysis of *Escherichia coli* growing on minimal and rich media. **Journal of Bacteriology**, v. 181, n. 20, p. 6425-6440, 1999.

THERON, J.; CLOETE, T. E. Molecular Techniques for Determining Microbial Diversity and Community Structure in Natural Environments. **Critical Reviews in Microbiology**, v. 26, n. 1, p. 37-57, 2007.

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

TODOROVA, N. H., RADEVA, G. and KARAMFILOV, V.K. Microbial diversity in *Zostera* sp. beds in south-western Black sea region analyzed by amplified ribosomal DNA restriction enzyme analysis (ARDRA). **Biotech & Biotech Equip.**, v. 26, n. 1, p. 5-11, 2012.

XAVIER, J. B., PICIOREANU, C., ALMEIDA, J. S., van LOOSDRECHT, M. C. M., **Monitorização e modelação da estrutura de biofilmes**. Departamento de Biotecnologia Universidade Técnica de Delft. Boletim de Biotecnologia. p. 13, 2003.

XIE, Y.; WANG, J.; WU, Y.; REN, C.; SONG, C.; YANG, J.; YU, H.; GEISY, J. P.; ZHANG, X. Using in situ bacterial communities to monitor contaminants in river sediments. **Envirom. Pollution**, v. 212, p. 348-357, 2016.

WANG, G. C. Y.; WANG, Y. Frequency of formation of chimeric molecules as a consequence of PCR coamplification of 16S rRNA genes from mixed bacterial genomes. **Applied and Environmental Microbiology**, v. 63, p. 4645-4650, 1997.

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

WERNER, J. J.; KOREN, O.; HUGENHOLTZ, P.; DeSANTIS, T.; WALTERS, W. A.; CAPORASO, J. G.; ANGEMENT, L. T.; KNIGHT, R. and LEY, R. E. Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. **The ISME Journal**, v. 6, p. 94-103, 2012.

WEIDNER, S.; ARNOLD, W. and PÜHLER, A. Diversity of Uncultured Microorganisms Associated with the Seagrass *Halophila stipulacea* Estimated by Restriction Fragment Length Polymorphism Analysis of PCR-Amplified 16S rRNA Genes. **Applied and Environmental Microbiology**, v. 62, n. 3, p. 766-771, 1996.

WHITFORD, M. F.; FORSTER, R. J.; BEARD, C. E.; GONG, J.; TEATHER, R. M. Phylogenetic analysis of rumen bacteria by comparative sequence analysis of cloned 16S rRNA genes. **Anaerobe**, v. 4, n. 3, p. 153-163, 1998.

WILSON, K. H.; BLITCHINGTON, R. B. Human colonic biota studied by ribosomal DNA sequence analysis. **Applied and Environmental Microbiology**, v. 62, n. 7, p. 2273-2278, 1996.

WORDEN, A. Z., BINDER, B. J. Growth regulation of rRNA content in *Prochlorococcus* and *Synechococcus* (marine cyanobacteria_ measured by whole-cell hybridization of rRNA-targeted peptide nucleic acids. **J. Phycol.**, v. 39, p. 527-534, 2003.

WOESE, C. R.; FOX, G. E. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. **Proc. Nalt. Acad. Sci. USA**, v. 74, p. 5088-5090, 1977.

YAMEI, G.; ANYI, Y.; JUN, B.; RUXIA, M.; LEI, Y.; YANJIE, W.; WEIDONG, W. Bioreactor performance and microbial community dynamics in a production-scale biogas plant in northeastern China. **Int. J. Agric. & Biol. Eng.**, v. 10, n. 1, p. 191-201, 2017.

ZAKRZEWSKI, M.; GOESMANN, A.; JAENICKE, S.; JÜNEMANN, S.; EIKMEYER, F.; SZCZEPANOWSKI, R.; AL-SOUD, W, A.; SORENSEN, S.; PÜHLER, A.; SCHLÜTER, A. Profiling of the metabolically active community from a production-scale biogas plant by means of high-throughput metatranscriptome sequencing. **Journal of Biotechnology**, v. 128, p. 248-258, 2012.

ZENG, J.; GAO, J-M.; CHEN, Y-P.; YAN, P.; DONG, Y.; SHEN, Y.; GUO, J-S.; ZENG, N. & ZHANG, P. Composition and aggregation of extracellular polymeric substances (EPS) in hyperhaline and municipal wastewater treatment plants. **Scientific Reports**, v. 6, n. 26721, 2016.

ZHANG, W., SILEIKA, T. and PACKMAN, A. I. Effects of fluid flow conditions on interactions between species in biofilms. **FEMS Microbiology Ecology** v. 84, n. 2, p. 344-354, 2013.

ZHANG, Q.; KANG, O.; JABEEN, S.; DICK, W. A. Alterations in soil microbial communities caused by treatments with penicillin or neomycin. **Environmental Sci Pollut Research**, v. 24, p. 18651-18662, 2017.