**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

**Mapeamento de agrupamentos gênicos envolvidos na fixação
biológica de nitrogênio em genomas de isolados brasileiros de
cianobactérias**

Bruno Costa Evangelista de Souza

Dissertação apresentada para obtenção do título de
Mestre em Ciências. Área de concentração:
Microbiologia Agrícola

**Piracicaba
2015**

Bruno Costa Evangelista de Souza
Bacharel em Ciências Biológicas

**Mapeamento de agrupamentos gênicos envolvidos na fixação biológica de
nitrogênio em genomas de isolados brasileiros de cianobactérias**

Orientadora:
Profa. Dra. **MARLI DE FÁTIMA FIORE**

Dissertação apresentada para obtenção do título de
Mestre em Ciências. Área de concentração:
Microbiologia Agrícola

Piracicaba
2015

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP

Souza, Bruno Costa Evangelista de
Mapeamento de agrupamentos gênicos envolvidos na fixação biológica de nitrogênio
em genomas de isolados brasileiros de cianobactérias / Bruno Costa Evangelista de
Souza. - - Piracicaba, 2015.
101 p. : il.

Dissertação (Mestrado) - - Escola Superior de Agricultura "Luiz de Queiroz".

1. Genômica 2. Nostocales 3. Nitrogenase 4. Heterócitos 5. *nifH* I. Título

CDD 589.46
S729m

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

DEDICATÓRIA

Aos

*Meus pais, Lêda e Gabriel, meus porto seguro, que sempre
estiveram ao meu lado me apoiando,*

À

*Minha querida avó Lindaura, sinto falta do
abraço de todo dia.*

Aos

*Meus irmãos, Machel, Flora e Elmer, os melhores
do mundo e sempre tão queridos.*

Ao

*Pequeno, traquino e querido Ian, que vejo crescer de longe,
mas que continua sendo
minha alegria mais sincera.*

A TODOS,

DEDICO

AGRADECIMENTOS

Inicialmente, gostaria de agradecer à minha família, a maior responsável pela minha formação. A distância é grande e dolorosa, mas tudo faz parte de um plano maior, nosso plano maior. Obrigado por todo o apoio e por partilhar comigo os melhores, e os piores, momentos.

À Prof^a. Dr^a. Marli de Fátima Fiore, pela orientação, confiança, ensinamentos e a oportunidade de desenvolver este projeto da melhor maneira possível.

À Universidade Estadual de Feira de Santana, pelo auxílio financeiro e concessão da bolsa de estudos durante o curso de Mestrado.

Ao Centro de Energia Nuclear na Agricultura pelo fornecimento de toda a infraestrutura necessária para a realização deste trabalho.

Ao Dr. Danillo Alvarenga e à Dr^a. Karina Heck pelos ensinamentos, lições, apoio e, principalmente, pela amizade.

Aos colegas do Laboratório de Biologia Celular e Molecular do CENA/USP, em especial aos amigos Bruno Weiss, Stella Thomaz de Lima, Ana Paula Dini Andreote, Gabriela Machineski, Janaina Rigonato, Paula Tamasauskas e Rafael Popin, pelas conversas, ajudas, ensinamentos e companheirismo. Vocês fizeram eu me sentir muito bem-vindo e tornaram mais leve a difícil tarefa que é ter que ficar longe de todos que te amam e apoiam.

Aos amigos de toda vida, Catharine Coelho, Carol Oliveira, Rafaela Oliveira, Danillo Matos, Sidney Pereira, Matheus Nolasco, Raíssa Sampaio, Ricardo Cedraz e Diogo Moura, que não permitiram que a distância enfraquecesse nosso amor e amizade.

Aos amigos conquistados em Piracicaba, Tânia Assis, Lu Soares, Tamires Cristina, Richtier Gonçalves, Marta Moitinho, Laura Bononi e Ana Paula Andrade, por ajudar a diminuir meus sentimentos de solidão.

Ao Programa de Pós-Graduação em Microbiologia Agrícola, Prof. Dr. Fernando Dini Andreote, Prof^a. Simone Lira e à secretária Maria Fernanda Prado toda a equipe, por todo o auxílio fornecido.

A todos que direta ou indiretamente contribuíram para a realização deste trabalho.

EPÍGRAFE

“Quem me pariu foi o ventre de um navio
Quem me ouviu foi o vento no vazio
Do ventre escuro de um porão
Vou baixar o seu terreiro
Epa raio, machado, trovão
Epa justiça de guerreiro
[...]
Vou aprender a ler
Pra ensinar os meu camaradas!.”.

Jose Carlos Capinam / Roberto Mendes

Sumário

RESUMO.....	11
ABSTRACT	13
1 INTRODUÇÃO	15
2 REVISÃO BIBLIOGRÁFICA	17
2.1 As cianobactérias	17
2.2 Ciclo do nitrogênio.....	18
2.3 Fixação biológica de nitrogênio	20
2.4 Cianobactérias fixadoras de Nitrogênio.....	23
2.5 Diferenciação de heterócitos em cianobactérias	26
2.6 Genética da Fixação Biológica de Nitrogênio.....	28
3 HIPÓTESE	31
4 OBJETIVO.....	31
4.1 Objetivos Específicos:	31
5 MATERIAL E MÉTODOS	33
5.1 Origem e cultivo das linhagens	33
5.2 Preparo do sequenciamento	34
5.3 Análises genômicas	34
5.3.1 Análise de qualidade das sequências e montagem do genoma.....	34
5.3.2 Linhagens com genomas depositados em bancos de dados	35
5.3.3 Busca do agrupamento gênico envolvido na fixação biológica de nitrogênio e diferenciação de heterócitos.....	36
5.3.4 Reconstrução filogenética dos genes.....	37
6 RESULTADOS E DISCUSSÃO.....	39
6.1 Sequenciamento e montagem dos genomas	39
6.2 Anotação de genes de Fixação Biológica de Nitrogênio	42
6.3 Anotação de genes ligados à diferenciação de heterócitos	47
6.4 Análises filogenéticas dos genes de fixação biológica de nitrogênio	49
7 CONCLUSÕES	61
REFERÊNCIAS.....	63
APÊNDICES.....	75
ANEXOS	97

RESUMO

Mapeamento de agrupamentos gênicos envolvidos na fixação biológica de nitrogênio em genomas de isolados brasileiros de cianobactérias

Cianobactérias são micro-organismos que realizam fotossíntese oxigênica e têm uma distribuição cosmopolita. Algumas cianobactérias são capazes de realizar fotossíntese e fixação biológica de nitrogênio (FBN), dois dos mais importantes processos na natureza, simultaneamente. As cianobactérias da ordem Nostocales são capazes realizar uma separação espacial destes dois processos por meio da formação de células especializadas, os heterócitos, onde acontece a fixação de nitrogênio, restringindo a fotossíntese às células vegetativas. Embora tenham importância ecológica, econômica e social, as cianobactérias foram muito pouco estudadas com enfoque genômico. Este trabalho teve como objetivo a caracterização dos agrupamentos gênicos envolvidos na fixação biológica de nitrogênio e na diferenciação de heterócitos de três isolados brasileiros de cianobactérias da ordem Nostocales. Para este fim, culturas das linhagens *Sphaerospermopsis torques-reginae* ITEP-024, *Nostoc* sp. CENA67 e *Fischerella* sp. CENA161 foram sequenciadas com a plataforma MiSeq e então foi realizada a montagem *ab initio* do genoma com as leituras obtidas. Além desses três isolados, os genomas de outras 31 linhagens disponíveis no banco de dados GenBank foram recuperados e utilizados para comparação. A anotação dos genes foi realizada por meio do alinhamento de sequências nucleotídicas já conhecidas de outras linhagens contra os genomas dessas linhagens, utilizando a ferramenta BLASTn, e por meio do servidor antiSMASH. Além disso, análises filogenéticas foram realizadas a partir dos genes anotados. O sequenciamento dos genomas das três linhagens apresentou altos valores de qualidade de bases e elevada cobertura genômica. A anotação dos agrupamentos envolvidos na FBN revelou a presença de um total de 22 genes envolvidos na síntese da Mo-nitrogenase, sendo 19 presentes em todas as linhagens. Os isolados brasileiros apresentaram sintenia com outras linhagens próximas filogeneticamente, apresentando variação apenas na presença de regiões de excisão e do gene *glbN*. A linhagem CENA161 apresentou um agrupamento gênico completo para síntese da V-nitrogenase, assim como apenas outras 5 linhagens em toda a ordem. Foram encontradas nas três linhagens sequenciadas os genes *devACB*, relacionados à formação da camada polissacarídica do heterócito. Os genes *hglEGDCA* e *hetM*, que estão relacionados à formação da camada glicolípida de heterócitos, foram encontrados completos nas linhagens ITEP-024 e CENA67, mas apenas parcialmente na CENA161. Genes reguladores do processo de diferenciação também foram acessados nas três linhagens brasileiras, entretanto apenas os reguladores globais do processo foram encontrados em CENA161. As análises filogenéticas mostraram que o gene *nifH* não reflete a filogenia do táxon e não é um bom marcador filogenético. Entretanto, a análise de todo o agrupamento refletiu o padrão filogenético de acordo com a taxonomia. Os resultados contribuem para a melhor compreensão dos aspectos genéticos e evolutivos do processo de FBN em cianobactérias.

Palavras-chave: Genômica; Nostocales; Nitrogenase; Heterócitos; *nifH*

ABSTRACT

Mapping of gene clusters involved in biological nitrogen fixation in genomes from Brazilian cyanobacterial isolates

Cyanobacteria are oxygenic photosynthetic microorganisms that have a worldwide distribution. Some cyanobacteria are capable of photosynthesis and biological nitrogen fixation (BNF), two of the most important processes in nature, simultaneously. Cyanobacteria from the Nostocales order perform a spatial separation of these processes through the formation of specialized cells, heterocytes, where nitrogen fixation is carried out, while photosynthesis is limited to vegetative cells. Although cyanobacteria have ecological, economic and social importance, they have been understudied with genomic approaches. This study aimed to characterize gene clusters involved in nitrogen fixation and heterocytes differentiation from three Brazilian strains of cyanobacteria from Nostocales order. For this purpose, non-axenic cultures of the strains *Sphaerospermopsis torque-reginae* ITEP-024, *Nostoc* sp. CENA67 and *Fischerella* sp. CENA161 were sequenced with the Illumina MiSeq platform and *ab initio* genome assembly was performed. In addition to these strains, genomes from 31 strains available in the GenBank database were retrieved and used for comparison. Gene annotation was performed through alignments between known genes present in other cyanobacteria and the strains genomes, using the BLASTn tool, and through the antiSMASH server. Furthermore, phylogenetic analyses were performed on the annotated genes. The genome sequencing showed high bases quality values and genomic coverage. The annotation of clusters involved in the BNF revealed the presence of a total of 22 genes involved in the synthesis of Mo-nitrogenase, among 19 were present in all strains. Brazilian isolates showed synteny with phylogenetically related strains, with variations only in the presence of excision regions and *glnB* gene. The CENA161 strain showed a complete gene cluster for synthesis of V-nitrogenase, present in only 5 other strains in this order. *devACB* genes, related to the heterocyte polysaccharide layer formation, were found in the three strains sequenced. The *hglEGDCA* and *hetM* genes, related to the formation of heterocyte glycolipid layer, were complete in ITEP-024 and CENA67 strains, but only partial in CENA161. Gene regulation of the heterocyte differentiation process have also been accessed in the three Brazilian strains, but only the general process regulatory genes were found in CENA161. Phylogenetic analysis showed that the gene *nifH* does not reflect the phylogeny of this group and should not be considered a good phylogenetic marker. However, the complete genes cluster analysis reflected the patterns of phylogenetic grouping according to taxonomy. The results contribute to a better understanding of the genetic and evolutionary aspects of the BNF process in cyanobacteria.

Keywords: Genomic; Nostocales; Nitrogenase; Heterocyte; *nifH*

1 INTRODUÇÃO

As cianobactérias têm sua origem estimada há aproximadamente 2,9 bilhões de anos e são reconhecidas pela sua fundamental participação na oxigenação da atmosfera terrestre. O ser humano, com sua necessidade de compreender a história evolutiva do mundo ao seu redor, busca nas cianobactérias uma forma de reconstruir a história da vida no planeta.

A plasticidade fenotípica e metabólica apresentada por estes organismos permite que este grupo esteja presente em ambientes diversos como solos, águas doces, águas marinhas e, ainda, em habitats extremos de temperatura e pH, sendo encontrados em vida livre ou associados a outros organismos, como plantas, algas, fungos e metazoários. Entre estas bactérias gram-negativas, encontram-se os únicos organismos capazes de realizar fotossíntese oxigênica e fixação biológica de nitrogênio, sendo, portanto, produtores primários e fornecedores de formas assimiláveis de nitrogênio para outros organismos presentes no ambiente. Essas capacidades metabólicas lhes conferem grande relevância ecológica, principalmente em ambientes oligotróficos, como oceanos.

Embora a importância das cianobactérias seja notória, as pesquisas com estes micro-organismos andam a passos lentos, quando comparadas com outros grupos microbianos. Os avanços tecnológicos da biologia molecular e a modernização e popularização das técnicas de sequenciamento de DNA permitiram o aumento de pesquisas nas áreas da genômica e metagenômica nas últimas décadas. Paralelamente, os avanços computacionais têm permitido a realização de investigações cada vez mais profundas na biologia dos organismos. Hoje é possível prever moléculas, rotas metabólicas, toxinas, taxonomia e uma série de outras informações biológicas a partir da investigação genômica.

Entretanto, devido a algumas dificuldades inerentes ao trabalho com cianobactérias, como as características genômicas e a dificuldade de obtenção de culturas axênicas, o número de genomas de cianobactérias disponibilizados ainda é muito pequeno e as pesquisas desenvolvidas são, geralmente, apenas com algumas linhagens-modelo, desconsiderando diversidade taxonômica e funcional deste grupo. Mesmo assim, o cenário é otimista e percebe-se nos últimos anos o aumento no número de pesquisas com genômica destes organismos.

A biodiversidade do Brasil é considerada uma das maiores do planeta e as cianobactérias dos seus diversos biomas podem apresentar um grande potencial de produção de metabólitos ainda desconhecidos. As singularidades dos ecossistemas brasileiros também podem favorecer pressões de seleção únicas, fazendo com que o metabolismo de carbono, nitrogênio, fósforo e outros elementos possam ser diferentes daqueles encontrados e estudados no hemisfério norte, onde estão concentradas as pesquisas em genômicas de cianobactérias. No Brasil, o sequenciamento genômico de cianobactérias ainda é incipiente, tendo o primeiro genoma de uma cianobactéria brasileira sido disponibilizado apenas no ano de 2013, por Fiore et al. Outros projetos de sequenciamento de linhagens diferentes e de ecossistemas pouco explorados estão em andamento, sendo explorado tanto o potencial de produção de substâncias bioativas quanto às interações ecológicas, taxonômicas e a caracterização do metabolismo de nitrogênio, foco do presente trabalho.

2 REVISÃO BIBLIOGRÁFICA

2.1 As cianobactérias

Cyanobacteria é um filo pertencente ao domínio Bacteria, composto por micro-organismos gram-negativos, com alta diversidade morfológica e ecológica, e constitui o único grupo de bactérias capazes de realizar fotossíntese oxigênica. Este grupo tem sua origem datada em 2,7-2,9 bilhões de anos (BLANK; SÁNCHEZ-BARACALDO, 2010) e provavelmente conteve os primeiros organismos a liberar oxigênio para a atmosfera, possuindo grande importância na oxigenação da atmosfera terrestre. Dessa forma, cianobactérias possibilitaram a formação de barreiras de proteção à radiação solar, como a camada de ozônio, e permitiram a evolução de seres mais complexos fora dos ambientes aquáticos, o que favoreceu a estruturação da vida tal qual a conhecemos hoje.

Cianobactérias, algas e plantas possuem aparatos fotossintéticos similares, utilizando a água como agente redutor na redução de gás carbônico em carboidratos, com liberação de oxigênio. A similaridade entre os aparatos fotossintéticos de cianobactérias, algas e plantas é o ponto de partida para a ideia de que os cloroplastos de organismos eucariotos fotossintetizantes surgiram a partir de um processo de endossimbiose. Esta teoria não é nova e foi proposta por Schimper e Mereschkowski há mais de um século, tendo ganhado força na década de 1970 a partir da análise de dados de ultra-estrutura e biologia molecular (MARGULIS, 1970; McFADDEN, 2001).

Embora a maioria das cianobactérias realize a fotossíntese oxigênica, com a água como doador de elétrons, algumas exceções têm sido encontradas. Alguns grupos de cianobactérias foram encontrados crescendo na ausência de luz, nutrindo-se por meio de fermentação de substratos orgânicos, sendo considerados, portanto heterotróficos facultativos (STAL; MOEZELAAR, 1997). A molécula doadora de elétrons para a fotossíntese também pode sofrer variações dentro do grupo, de modo que algumas cianobactérias são hábeis em utilizar sulfeto no lugar de água, em condições anaeróbicas (CASTENHOLZ; WATERBURY, 1989; COHEN et al., 1986). A presença de aparato fotossintético completo também não é unânime dentro das cianobactérias. Recentemente foi identificado um grupo de cianobactérias planctônicas em ambiente marinho, cuja análise genômica indica a falta de genes

codificadores para enzimas do ciclo de Calvin e para biossíntese de vários aminoácidos. Essas evidências indicam um novo modo de vida, sugerindo que este organismo depende de outro para sobreviver, em relações muito próximas ou simbióticas, visando à obtenção de nutrientes essenciais (BOTHE et al., 2010; TRIPP et al., 2010).

Em condições apropriadas de luminosidade e disponibilidade de nutrientes, as cianobactérias podem formar florações, que podem ser definidas como o aumento massivo da biomassa cianobacteriana, com predomínio de uma ou duas espécies (OLIVER; GANF, 2000). Além da importância ecológica, as cianobactérias são produtoras de uma alta gama de substâncias biologicamente ativas, como inibidores de proteases, inibidores de fotossíntese, antimicrobianos, imunossupressores, toxinas, entre outras moléculas que podem ter importância social e valor econômico agregado (MOORE et al., 1984; NAMIKOSHI; RINEHART, 1996; BURJA et al., 2001; ETCHEGARAY et al., 2004, SILVA-STENICO et al., 2011, LEÃO et al., 2012, LIAIMER et al, 2015).

A grande diversidade metabólica e morfológica das cianobactérias é reflexo da longa história evolutiva deste grupo, o que permite que sejam encontrados em praticamente todos os ambientes do planeta, incluindo lugares considerados inóspitos como aqueles que apresentam extremos de pH, desertos, águas termais e lagos antárticos (CASTENHOLZ, 1973; DOR; DANIN, 1996; SKULBERG, 1995).

Essas características dotam as cianobactérias de fundamental importância como produtores primários na cadeia trófica e no fluxo de energia no ecossistema (RAMAKRISHNAN et al, 2010). Além de ocuparem a base da teia trófica, as cianobactérias têm importante atuação na disponibilização de elementos essenciais à vida, principalmente o nitrogênio.

2.2 Ciclo do nitrogênio

O nitrogênio faz parte da constituição de proteínas, ácidos nucleicos, peptidoglicanos, aminoácidos e outras moléculas essenciais, sendo, portanto um elemento fundamental para a manutenção da vida, cuja contribuição para a biomassa total dos seres vivos só é superada pelas contribuições de carbono, hidrogênio e oxigênio (YOUNG, 1992).

O nitrogênio é um elemento limitante à vida e sustentabilidade de toda uma cadeia trófica. Embora encontrado em grande quantidade na atmosfera, a entrada do nitrogênio nos ecossistemas está comumente associada à decomposição de matéria orgânica. No entanto, a fixação biológica de nitrogênio (FBN) é a principal fonte de entrada de nitrogênio em ambientes cuja taxa de metabolização de matéria orgânica é baixa, como manguezais, ou em ambientes oligotróficos, como os oceanos (HOLGUIN et al., 2001; ZEHR, 2011). O mecanismo de fixação de nitrogênio, realizado apenas por procariotos, principalmente cianobactérias, algumas bactérias de vida livre do solo e água e ainda bactérias simbiotes de raízes de plantas e arqueias, está ligado à presença e atividade do complexo enzimático conhecido como nitrogenase que consegue realizar a conversão de moléculas de nitrogênio elementar (N_2) em amônia (RAYMOND et al., 2004).

As transformações do nitrogênio em formas químicas passíveis de ser assimiladas por organismos vivos ocorrem em processos metabólicos interdependentes mediados por micro-organismos. Esse conjunto de processos compõe o vasto ciclo do nitrogênio (NELSON; COX, 2002).

A primeira etapa do ciclo de nitrogênio é a redução do nitrogênio atmosférico a amônia, por bactérias fixadoras de nitrogênio. O íon amônio, forma adquirida pela amônia quando em solução, é a forma prontamente absorvida pela maioria dos organismos tanto no solo, como na água (ESTEVES; AMADO, 2011). Várias bactérias obtêm energia por meio da oxidação da amônia em nitrito (NO^{2-}) e nitrato (NO^{3-}), de modo que praticamente toda a amônia do ambiente é transformada em matéria orgânica, ou em nitrato, em um processo conhecido por nitrificação. Após a morte dos organismos no ecossistema, a degradação da matéria orgânica pela ação dos micro-organismos decompositores devolve a amônia ao ambiente, no processo chamado de amonificação (MOREIRA; SIQUEIRA, 2006; GUERRERO et al., 1981).

A entrada de nitrito e nitrato em ambientes reduzidos pode ter causado grandes mudanças na biosfera. As cargas negativas dessas moléculas proporciona uma excelente solubilidade, tornando-os bastante biodisponíveis no ambiente. Por esse motivo, o nitrato tem o potencial de servir como fonte de nitrogênio para o crescimento dos seres vivos, de modo que esta capacidade generalizou-se inicialmente entre as bactérias e arqueias, e posteriormente entre os eucariotos (WRAGE et al, 2001; STOLZ; BASU, 2002). A disponibilidade de nitrogênio do solo é

regulada pelas vias dissimilatórias, que reduzem o nitrogênio a formas gasosas como N_2 e N_2O . Esse processo de saída de nitrogênio do ambiente, conhecido como desnitrificação, é mediado por bactérias anaeróbias facultativas que utilizam as formas oxidadas como aceptores finais de elétrons oriundos da oxidação de substratos reduzidos (ZUMFT, 1997; STOLZ; BASU, 2002). Ecologicamente, a saída do nitrogênio do ambiente é danosa principalmente para os ambientes terrestres; entretanto, em termos globais, a remoção do nitrogênio do solo antes que este seja lixiviado e atinja os oceanos é vital porque evita o enriquecimento deste ecossistema e, conseqüentemente, possíveis desequilíbrios ambientais (MOREIRA; SIQUEIRA, 2006).

Até recentemente, acreditava-se que a desnitrificação era o único processo de retirada de nitrogênio de um ambiente. Entretanto, em ambientes anaeróbios, porém ainda com disponibilidade de nitrito, pode ocorrer um processo conhecido por anammox (oxidação anaeróbica de amônio), na qual alguns organismos heterotróficos são capazes de oxidar o íon amônio, utilizando nitrito comoceptor de elétrons, liberando gás nitrogênio (REVSBECH et al., 2006; THAMDRUP; DALSGAARD, 2002; KUYERS et al., 2003). Em ambientes continentais, considerando os poucos estudos em rios, lagos e estuários, a representatividade deste processo não ultrapassa 25%, sendo a desnitrificação o processo preponderante (REVSBECH et al., 2006). Entretanto, a contribuição marinha do processo anammox vem sendo calculada entre 30 e 50% de todo nitrogênio existente (JETTEN et al., 2005; DALSGAARD et al., 2003).

2.3 Fixação biológica de nitrogênio

Carbono, hidrogênio, oxigênio e nitrogênio são os quatro elementos essenciais mais abundantes nos organismos vivos. Os três primeiros são obtidos facilmente por eucariotos e procariotos por meio da respiração e dos metabolismos fototróficos e quimiolitotróficos. O nitrogênio, entretanto, está disponível no ambiente em grande quantidade, porém em formas não combinadas, as quais não são prontamente assimiláveis por todos os eucariotos e pela maioria dos procariotos (MOREIRA; SIQUEIRA, 2006). O nitrogênio é o elemento químico quantitativamente mais abundante na atmosfera terrestre, compondo aproximadamente 78% dos seus gases (SIQUEIRA; FRANCO, 1988). Segundo Reis e Teixeira (2005), a abundância

de nitrogênio na atmosfera é explicada pela estabilidade que o átomo de nitrogênio adquire quando se liga a outro, formando uma forte ligação tripla entre os dois átomos de nitrogênio ($N\equiv N$), tornando a molécula de N_2 quase inerte e pouco reagente.

A reação que transforma o nitrogênio atmosférico em amônia exige temperatura e pressão muito elevadas, de modo a possibilitar o rompimento da ligação tripla covalente entre os dois átomos de nitrogênio. Entretanto, alguns micro-organismos são capazes de realizar a conversão de dinitrogênio atmosférico (N_2) a amônio em condições normais de temperatura e pressão, num processo enzimático conhecido como fixação biológica de nitrogênio (FBN), catalisado pela enzima nitrogenase (NEVES; RUMJANECK, 1992). A FBN é, portanto, uma fonte de entrada de formas assimiláveis de nitrogênio nos ecossistemas, podendo corresponder a 97% de todo o nitrogênio que entra em ambientes terrestres naturais (ZEHR et al., 2003; GALLOWAY et al., 2004).

O processo de fixação biológica de nitrogênio é mediado por uma parcela dos procariotos que, apesar de pequena, apresenta alta diversidade filogenética. Dentre os organismos fixadores há as cianobactérias, com espécies fixadoras distribuídas por todas as ordens, com exceção de Gloeobacterales, as Proteobacterias, com espécies fixadoras em todas as classes, o gênero *Frankia*, pertencente ao filo das actinobactérias e capaz de formar simbiose com raízes de plantas, algumas espécies de gêneros do grupo dos Firmicutes, como *Bacillus*, *Paenobacillus* e *Clostridium* e espécies de Archaea dos grupos dos halófilos e metanogênicos (EADY, 1991; MOREIRA; SIQUEIRA, 2006). Estudos recentes têm sugerido que a distribuição da capacidade fixação de nitrogênio nesses organismos é resultado de uma longa história evolutiva que envolve eventos sucessivos de transferências laterais e verticais entre estes grupos (RAYMOND et al., 2004; KECHRIS et al., 2006; BOLHUIS et al., 2009).

A nitrogenase é formada a partir da combinação de duas proteínas conhecidas como dinitrogenase (MoFe, dependente de molibdênio-ferro) e dinitrogenase redutase (Fe, dependente de ferro). Esses termos refletem a composição metálica dos grupos prostéticos contidos nas respectivas proteínas (PETER et al., 1995). A dinitrogenase redutase é um homodímero, codificado pelo gene *nifH* (MEVARECH et al., 1980), cuja função é hidrolisar o ATP, mediando a transferência de elétrons entre um doador externo e a dinitrogenase. Os genes *nifD*

e *nifK* (LAMMERS; HASELKORN, 1983; MAZUR; CHUI, 1982) codificam as subunidades alfa e beta, respectivamente, que formam a proteína heterotetrâmera dinitrogenase, correspondente ao sítio catalítico de ligação e redução do N_2 . Existem ainda outros dois sistemas alternativos de nitrogenase, que apresentam outros metais como cofatores. Um desses é o sistema dependente de vanádio no lugar de molibdênio, que é codificado pelos genes *vnfHGDK*.

A nitrogenase alternativa dependente de vanádio é composta por duas subunidades alfa (*VnfD*), duas subunidades beta (*VnfK*) e quatro subunidades delta (*VnfG*), formando uma enzima heterooctomérica com dois cofatores de ferro e vanádio (Fe-V) (LEE et al., 2009; EADY, 1988; 1995). Como a *NifH*, a *VnfH*, em conjunto com a ferredoxina, transfere elétrons para a dinitrogenase. A V-nitrogenase mostra diferente especificidade de substratos, sendo incapaz de reduzir acetileno em etileno, além de ser relativamente menos eficiente que a Mo-nitrogenase na redução de dinitrogênio, produzindo uma quantidade maior de hidrogênio (DILWORTH et al., 1988). Por causa dessa ineficiência da V-nitrogenase em reduzir N_2 e da elevada capacidade de produzir hidrogênio, linhagens de *Anabaena variabilis*, que possuem essa enzima, têm sido utilizadas para produção de hidrogênio em biorreatores (TSYGANKOV, et al., 2002).

O outro sistema possui apenas ferro e é codificado pelos genes *anfHDGK* (BISHOP; JOERGER, 1990; ZEHR et al., 2003). Essa terceira nitrogenase não contém molibdênio ou vanádio e o átomo heterometal no cofator é o ferro. Esta enzima é conhecida como nitrogenase-3 ou nitrogenase dependente de ferro (Fe-nitrogenase) (MULLER et al., 1992). Em *Azotobacter vinelandii*, os genes estruturais estão organizados como um único operon *anfHDGKOR* (JOERGER et al, 1989; MYLONA et al., 1996), sendo que as subunidades da dinitrogenase redutase são codificadas pelo gene *anfH*, enquanto que as subunidades alfa e beta da dinitrogenase são codificadas pelos genes *anfD* e *anfK*, respectivamente. Como na V-nitrogenase, uma subunidade delta adicional codificada pelo gene *anfG*, que é específica para nitrogenases alternativas, forma a estrutura hexamérica da Fe-nitrogenase (WAUGH et al., 1995) As funções dos genes *anfO* e *anfR* ainda são desconhecidas.

Segundo Postgate (1998), a reação de redução do nitrogênio atmosférico a amônia é extremamente custosa energeticamente, consumindo 16

moléculas de ATP e 8 elétrons para cada molécula de N₂ reduzido, podendo ser simplificada na seguinte reação estequiométrica:

Todo este processo é sensível ao oxigênio, uma vez que, por conta da atividade redox, a nitrogenase é pouco tolerante a este elemento. O oxigênio pode inibir a nitrogenase em três níveis: em nível genético, agindo na repressão da síntese da enzima; causando dano irreversível à proteína-Fe (dinitrogenase redutase) existente; e causando dano reversível à enzima. Neste último caso, a atividade retorna ao nível inicial quando a pressão parcial de oxigênio volta aos níveis adequados (REIS; TEIXEIRA, 2005). Essa sensibilidade pode ser explicada pela origem evolutiva do processo de FBN, uma vez que esta enzima originou-se em um período anterior à oxigenação da atmosfera terrestre (BERMAN-FRANK et al., 2003).

O surgimento da fixação biológica de nitrogênio e a evolução da nitrogenase ainda não apresentam uma forte sustentação empírica. Entretanto, as análises estruturais da nitrogenase e de complexos proteicos parálogos sugerem a possibilidade de uma origem comum entre a nitrogenase e enzimas associadas à biossíntese de clorofila e bacterioclorofila (BURKE et al., 1993; XIONG et al., 2000; BROCKER et al., 2008). Este complexo enzimático catalisa a ligação entre os carbonos C17 e C18 da protoclorofilida para formar a clorofilida (BOYD; PETERS, 2013). Essa reação é análoga ao mecanismo de transferência de elétrons requerido na fixação biológica de nitrogênio, envolvendo um análogo da dinitrogenase, BchL, que realiza a transferências de elétrons, e um análogo da nitrogenase redutase, BchNH, que apresenta o sítio ativo da redução da protoclorofilida.

2.4 Cianobactérias fixadoras de Nitrogênio

A disponibilidade de nitrogênio é um dos fatores que controlam a produtividade dos oceanos, portanto tem sido de grande interesse determinar a

magnitude e as vias de incremento de nitrogênio nos oceanos por meio da fixação biológica de nitrogênio (ZEHR, 2011). Sendo as cianobactérias os principais micro-organismos fixadores de nitrogênio dos oceanos abertos, com o gênero *Trichodesmium* responsável por aproximadamente 42% de nitrogênio global fixado (240 Tg de N₂ por ano), elas possuem significativa importância nos ciclos biogeoquímicos marinhos e no fluxo de gases entre os oceanos e a atmosfera, ajudando no controle das concentrações de gases causadores de efeito estufa na atmosfera, devido à retirada de CO₂ atmosférico pela realização da fotossíntese (BERMAN-FRANK et al., 2003; ZEHR, 2011).

O principal fator limitante para atividade da enzima nitrogenase é a presença de oxigênio (POSTGATE, 1998). Sendo as cianobactérias organismos que realizam fotossíntese oxigênica, foi necessário o desenvolvimento de adaptações para separar o N₂ do O₂. Estas estratégias incluem a separação apenas espacial da FBN e fotossíntese oxigênica, por meio da diferenciação de células especializadas chamadas heterócitos – que possuem camadas adicionais de glicolipídeos e polissacarídeos e ausência de aparato fotossintético; separação apenas temporal, com atividade fotossintética ocorrendo em período luminoso e a fixação de nitrogênio, durante o período escuro; separação espacial e temporal, com diferença no período de realização dos dois processos e a restrição da fixação de nitrogênio a células não diferenciadas que concentram a produção e a atividade da enzima nitrogenase; e a realização de fixação de nitrogênio em ambiente microaerofílicos (FIORE; HONDA, 2008; BERMAN-FRANK et al., 2001).

O gênero *Trichodesmium*, principal organismo diazotrófico dos oceanos, apresenta uma estratégia muito particular para separação do O₂ e N₂. Sendo uma cianobactéria filamentosa, ela confinou a realização da fixação de nitrogênio a algumas células, chamadas de diazócitos, que não se diferenciam morfológicamente das demais, como os heterócitos; entretanto, em períodos determinados, estas células apresentam uma taxa elevada de respiração, consumindo mais O₂ que o liberado pela fotossíntese, permitindo o funcionamento da nitrogenase (BERGMAN; CARPENTER, 1991). As cianobactérias podem também realizar o processo de fixação de nitrogênio associadas ou em simbiose com outros organismos, em interações com variações de dependência onde ocorre o fornecimento de substâncias como fonte de carbono e nitrogênio para o hospedeiro

em troca de proteção contra condições ambientais adversas (FIORE; HONDA, 2008).

A ocorrência de relações simbióticas de cianobactérias é relatada desde o século XIX, quando foram identificadas células de *Anabaena* nas raízes de gimnospermas e no caule de *Gunnera*. Posteriormente, foram identificadas associações entre *Nostoc* e briófitas e os aspectos fisiológicos da interação e disponibilização de nitrogênio entre cianobactérias e cicadáceas (FIORE; HONDA, 2008). A capacidade de fixar nitrogênio atmosférico faz das cianobactérias um excelente parceiro ecológico, de modo que as cianobactérias podem formar simbiose com diatomáceas, fungos, briófitas, pteridófitas e gimnosperma, até esponjas, ascídias e vermes da classe Echiura (ADAMS, 2000; WHITTON, 2000; RAI et al., 2000). As principais associações simbióticas de cianobactérias são encontradas com as formas heterocitadas; entretanto, cianobactérias não heterocitadas diazotróficas também são relatadas como cianobiontes, em relações nas quais as cianobactérias geralmente fornecem nitrogênio e carbono e em troca recebem proteção contra irradiância, dessecação e predação (ADAMS, 2000).

A origem da fixação de nitrogênio em cianobactérias ainda não é um consenso no meio científico. As pesquisas nessa área tentam verificar se o primeiro ancestral cianobacteriano já seria capaz de fixar nitrogênio. Alguns estudos como o de Sánchez-Baracaldo et al. (2005) previam a pouca possibilidade de isso ocorrer. Entretanto, estudos recentes sugerem que a nitrogenase surgiu em uma cianobactéria ancestral e foi transferida verticalmente, acompanhando a evolução filogenética do grupo, ocorrendo, eventualmente, transferências laterais e redução genômica, com perda da capacidade fixadora (LATYSHEVA et al., 2012). Os eventos de transferências horizontais envolvendo cianobactérias são eventualmente identificados, como a aquisição do aparato fixador por *Microcoleus chthonoplastes* que adquiriu os genes de fixação de um organismo da classe Deltaproteobacteria (BOLHUIS et al., 2010). Outro exemplo foi a aquisição de genes de fixação de nitrogênio pela bactéria *Wolinella succinogenes*, isolada de rúmen bovino, cuja análise genômica identificou agrupamento gênico para fixação de nitrogênio mais parecido com o encontrado em cianobactérias que com os de micro-organismos fixadores de nitrogênio mais proximamente relacionados evolutivamente, indicando a provável transferência horizontal entre esses organismos (BAAR et al., 2003).

2.5 Diferenciação de heterócitos em cianobactérias

O avanço da biologia molecular e a abordagem polifásica na classificação de cianobactérias têm contribuído consideravelmente para o avanço na sistemática e taxonomia de cianobactérias, incluindo a descrição de numerosos novos gêneros e espécies, bem como a reclassificação e reorganização de táxons já existentes. (BOHUNICKA et al., 2015) A classificação dos táxons heterocitados têm sofrido revisões ao longo dos anos, devido às incongruências entre algumas características morfológicas e as polifilias reveladas pelas análises filogenéticas.

Até o ano de 2004, as cianobactérias heterocitadas eram separadas em duas ordens: Stigonematales, organismos com células organizadas em colônias filamentosas, heterócito em posição terminal, lateral ou intercalar e presença de ramificações verdadeiras; e Nostocales, que diferiria de Stigonematales apenas pela inexistência de ramificações verdadeiras (ANAGNOSTIDIS; KOMÁREK, 1990). A partir de 2005, entretanto, as análises moleculares mostraram que os gêneros pertencentes às duas ordens, Stigonematales e Nostocales, constituíam um grupo monofilético, mostrando que a capacidade de diferenciar células vegetativas em estruturas especializadas no processo de fixação biológica de nitrogênio, os heterócitos, representava a melhor característica diagnóstica deste grupo de cianobactérias e a capacidade de formar ramificações verdadeiras passou a ser característica de diagnose de algumas famílias dentro desta ordem (HOFFMANN et al., 2005).

A formação de heterócitos em cianobactérias sempre esteve supostamente relacionada à capacidade de fixação de N_2 e sua importância na história evolutiva da Terra, de acordo com os registros fósseis, mas a ligação entre a estrutura e o processo só foi comprovada em 1968, em trabalho publicado por Fay et al. A formação dos heterócitos é um processo mediado pela ativação de genes que são responsáveis por diferentes processos que culminam na diferenciação da célula vegetativa funcional em um centro de fixação de nitrogênio. O principal fator determinante para esta diferenciação celular é a carência de nitrogênio no ambiente. Essa escassez do nutriente provoca o acúmulo de 2-oxoglutarato no interior das células que, quando percebido pela enzima NtcA, aumenta sua atividade metabólica, provocando, dentre outras consequências, a transcrição do gene *nrrA*, que regula a transcrição de *hetR*, principal regulador da diferenciação do heterócito. Dessa forma,

os principais genes envolvidos na diferenciação dessas células especializadas são *hetR* e *hetP* (indução da diferenciação), *patL*, *patN* e *patA* (padrão de diferenciação), *abp3*, *hetF*, *hetZ* e *hglT* (diferenciação) (LAURENT et al., 2005; HERRERO et al., 2001; VALLADARES et al., 2008; EHIRA; OHMORI, 2006; PLOMISKY et al., 2013).

A principal característica do heterócito é o espessamento do envoltório celular, composto por duas camadas. A camada mais interna, sintetizada pelos genes *hgl*, é composta por glicosídeos de longas cadeias de trióis, tetróis e hidroxicetonas (HERRERO et al, 2013). Esta camada glicolípídica mais interna é a principal barreira para a difusão de gases para dentro do heterócito (XU et al., 2008). A camada mais externa é constituída por polissacarídeos, sintetizados pelos genes *dev*, e está envolvida na proteção da camada mais interna (ZHANG et al. 2006). Anteriormente, acreditava-se que a camada polissacarídica era fluida o suficiente para permitir a entrada de N₂ atmosférico requerido para a FBN; entretanto, hoje é sabido que a única via de entrada de gases nos heterócitos é através dos poros de comunicação existentes nas células vegetativas (WALSBY, 1985; 2007).

As células vegetativas e os heterócitos do filamento apresentam uma relação de interdependência, uma vez que cada uma realiza metabolismos essenciais para a sobrevivência de qualquer célula viva. Os heterócitos não apresentam o fotossistema II da fotossíntese, responsável pela hidrólise e liberação de O₂ e elétrons. Em compensação, a membrana citoplasmática está dominada por fotossistemas I e ATPases (CARDONA et al, 2009). Dessa forma, os elétrons na forma de NADPH e ferredoxinas são transferidos para a produção de energia nos heterócitos (HERRERO et al. 2013). Em contrapartida, o nitrogênio produzido nos heterócitos é rapidamente transformado em moléculas de glutamina utilizando a enzima glutamina sintase, que une a amônia resultante de FBN a uma molécula de glutamato. É por meio dessa molécula que o heterócito fornece nitrogênio para as células vegetativas, exportando moléculas de glutamina e importando glutamato das células vegetativas. Além disso, parte do nitrogênio fixado é acumulada nas regiões polares próximas às células vegetativas, formando grânulos de cianoficina que também atuam como um controle do processo de FBN (PICOSI et al., 2005).

2.6 Genética da Fixação Biológica de Nitrogênio

Os genes de fixação de nitrogênio têm sido identificados em numerosas cianobactérias, mas poucas tiveram seus agrupamentos gênicos caracterizados com maiores detalhes. Os agrupamentos gênicos envolvidos na FBN mais bem caracterizados são os das cianobactérias heterocitadas *Nostoc* sp. PCC 7120, *Anabaena variabilis* ATCC 29413, *Cylindrospermopsis raciborskii* CS-505 e *Nostoc punctiforme* PCC 73102, e de uma cepa unicelular, *Synechococcus* sp. RF-1. A nomenclatura dos genes *nif* de cianobactérias é idêntica à utilizada para qualquer outra bactéria diazotrófica (BÖHME, 1998). Quinze genes de fixação ou relacionados ao processo, incluindo os genes estruturais para nitrogenase, *nifHDK*, estão agrupados de acordo com a seguinte ordem: *nifB-fdxN-nifS-nifU-nifH-nifDnifK-nifE-nifN-nifX-orf2-nifW-hesA-hesB-fdxH* (Figura 1). Esses genes geralmente estão organizados em seis unidades de transcrição: *nifB-fdxN-nifS-nifU*, *nifHDK*, *nifEN*, *nifX-orf2*, *nifW-hesA-hesB*, e *fdxH* (FLORES et al., 1999; HASELKORN; BUIKEMA, 1992; HUANG et al., 1999; STUCKEN et al., 2010).

Figura 1 – Representação esquemática do agrupamento de genes *nif*. O esquema representa o agrupamento gênico de 15 kb contendo os genes *nifHDK* e outros 13 genes relacionados a FBN na linhagem *Cylindrospermopsis raciborskii* CS-505 e sua comparação com os agrupamentos gênicos (*nif 1* e *nif 2*) da linhagem *Anabaena variabilis* ATCC29413, bem como sua ausência na linhagem *Raphidiopsis brookii* D9. Retirado de Stucken et al. (2010)

Acessório a esses genes, já foi encontrado também o gene *nifP*, na cepa RF-1, o gene *nifJ*, agrupado com *nifV*, *nifZ* e *nifT*, encontrados em outros locais no cromossomo de *Anabaena* sp. PCC 7120, um gene chamado de *glnN*, que codifica uma proteína homóloga a um monômero de hemoglobina, encontrado justaposto ao *nifU* e *nifH* na cepa *Nostoc commune* UTEX 584 e em algumas outras cepas de *Nostoc* (BÖHME, 1998; HILL et al., 1996; POTTS et al., 1992).

Os genes *nif* e outros genes relacionados à fixação de nitrogênio não são expressos em culturas com suprimento de nitrogênio assimilável e seus produtos podem estar confinados em heterócitos, nas formas heterocíticas. Algumas cianobactérias heterocitadas, como a *Anabaena variabilis* ATCC 29413, carregam um segundo agrupamento que é expresso em células vegetativas sob condições de microaerobiose ou anaeobiose (THIEL et al., 1995). Estas cianobactérias também portam genes para nitrogenase dependente de vanádio: *vnfDG*, *vnfK*, *vnfE* e *vnfN* (THIEL, 1993).

Os demais genes associados ao agrupamento são responsáveis pela síntese de cofatores, formação de cisteína dessulfurase, produção de ferredoxinas que transferem elétrons para dinitrogenase redutase e proteínas de ligação ao NAD/FAD (HASELKORN; BUIKEMA, 1992; STUCKEN et al., 2010). Rubio e Ladden (2005) fizeram uma compilação dos estudos publicados elencando os genes *nif* já descritos e suas respectivas funções sugeridas (Tabela 1).

Tabela 1 – Produtos do gene *nif* e sua função (conhecida ou proposta) na fixação do nitrogênio. Adaptado de Rubio e Ladden (2005) (Continua)

Gene	Identificação ou produto do gene
<i>nifH</i>	Fe-proteína. Doador obrigatório de elétrons para a MoFe proteína durante o ciclo de catálise da nitrogenase. NifH também é requerido para a síntese do cofator FeMo e maturação da proteína apo-MoFe.
<i>nifD</i>	Subunidade da proteína MoFe. Forma um tetrâmero $\alpha_2\beta_2$ com a subunidade β da proteína MoFe. O sítio de redução do substrato, FeMo-co, está ligado à subunidade α da proteína MoFe.
<i>nifK</i>	Subunidade da proteína MoFe..

Tabela 1 – Produtos do gene *nif* e sua função (conhecida ou proposta) na fixação do nitrogênio. Adaptado de Rubio e Ludden (2005) (Conclusão)

Gene	Identificação ou produto do gene
<i>nifU</i>	Molde para a formação do cluster Fe-S para os componentes da nitrogenases.
<i>nifE</i>	Forma um tetrâmero $\alpha_2\beta_2$ com NifN. Necessário para a síntese do FeMo-co.
<i>nifN</i>	Requerido para a síntese do FeMo-co. Forma um tetrâmero com NifE.
<i>nifX</i>	Envolvido na síntese do FeMo-co. Acumula um precursor biossintético do cofator.
<i>nifT</i>	Desconhecida
<i>nifS</i>	Envolvido na mobilização de S para a síntese e reparo do cluster Fe-S.
<i>nifV</i>	Homocitrato sintase envolvida na síntese do FeMo-co.
<i>nifW</i>	Envolvido na estabilidade da proteína MoFe.
<i>nifZ</i>	Desconhecido
<i>nifM</i>	Requerido para a maturação de NifH.
<i>nifF</i>	Flavodoxina. Doador fisiológico de elétrons para NifH em <i>Klebsiella pneumoniae</i> .
<i>nifL</i>	Regulador negativo
<i>nifA</i>	Regulador negativo
<i>nifB</i>	Requerido para a síntese do FeMo-co. Seu produto metabólico, NifB-co, é um doador específico de Fe e S para o FeMo-co.
<i>fdxN</i>	Ferredoxina. Em <i>Rhodospirillum capsulatus</i> funciona como doador de elétrons para a nitrogenase.
<i>nifQ</i>	Envolvido na síntese do FeMo-co. Admite-se que seja responsável pelo processamento do MoO ₄
<i>nifJ</i>	Piruvato: flavodoxina (ferredoxina) oxidoredutase. Doador de elétrons para a proteína Fe em <i>K. pneumoniae</i> .

3 HIPÓTESE

Os agrupamentos gênicos da fixação biológica de nitrogênio nas cianobactérias são conservados e podem refletir a filogenia deste grupo de micro-organismos, sendo que o gene *nifH*, muito utilizado como marcador molecular de organismos diazotróficos, não fornece resolução suficiente para a inferência de padrões evolutivos quando utilizado isoladamente.

4 OBJETIVO

O objetivo geral deste estudo é analisar aspectos evolutivos relacionados aos agrupamentos gênicos envolvidos na fixação biológica de nitrogênio em cianobactérias da ordem Nostocales.

4.1 Objetivos Específicos:

- I) Identificar e caracterizar os genes relacionados à FBN no genoma de três isolados brasileiros de cianobactérias;
- II) Comparar os agrupamentos gênicos dos isolados com outros disponíveis em bancos de dados públicos;
- III) Analisar filogeneticamente os genes de FBN em cianobactérias para inferência de padrões evolutivos;
- IV) Caracterizar genes relacionados à diferenciação de heterócitos.

5 MATERIAL E MÉTODOS

5.1 Origem e cultivo das linhagens

No presente trabalho foram utilizadas três linhagens de gêneros diferentes de cianobactérias pertencentes à ordem Nostocales, sendo que elas são mantidas em condições de cultivo no Laboratório de Biologia Celular e Molecular do Centro de Energia Nuclear na Agricultura (CENA) da Universidade de São Paulo (USP).

A linhagem *Sphaerospermopsis torques-reginae* ITEP-024 (Figura 2A) foi isolada em 2002 (MOLICA et al., 2005) de uma floração ocorrida no reservatório de Tapacurá (8°02'14"S, 35°09'46"W). Essa linhagem foi cedida ao Laboratório de Biologia Celular e Molecular do CENA/USP pela Dra. Vera Regina Werner, do Museu de Ciências Naturais e Fundação Zoobotânica do Rio Grande do Sul, Porto Alegre, RS. Desde então, a linhagem está sendo mantida em cultivo em meio de cultura líquido ASM-1 (GORHAM et al., 1964), sob exposição à luz fluorescente (40-50 μmol fótons $\cdot\mu\text{m}^{-2}\cdot\text{s}^{-1}$), com fotoperíodo de 14 horas de luz e 10 horas de escuro, a 25 °C $\pm$ 1 °C.

A linhagem *Nostoc* sp. CENA67 (Figura 2B) foi isolada em 2004 de uma amostra de solo de terra preta antropogênica do sítio arqueológico Hatahara, (3°16'29"S e 60°12'16"O), localizado no município de Iranduba, a 30 km a sudoeste de Manaus-AM. Essa linhagem é mantida em condições de cultivo em meio de cultura líquido AA/4 (ALLEN; ARNON, 1955), sob exposição à luz fluorescente (40-50 μmol fótons $\cdot\mu\text{m}^{-2}\cdot\text{s}^{-1}$), com fotoperíodo 14 horas de luz e 10 horas de escuro, a 25 °C $\pm$ 1 °C.

A linhagem *Fischerella* sp. CENA161 (Figura 2C) foi isolada em 2004 de uma amostra de água coletada de um pequeno reservatório de concreto (4 m³) construído para represar água de uma nascente na Fazenda Capuava, Piracicaba, São Paulo (22°39'25"S e 47°37'24"O) (FIORE et al., 2009). Ela é mantida em câmara de crescimento com temperatura constante de aproximadamente 25 °C $\pm$ 1 °C com fotoperíodo de 14 horas de claro e 10 horas de escuro, com irradiância de 40 μmol fótons m⁻²s⁻¹, em meios de cultura BG-11 (ALLEN, 1968), sem adição de nitrogênio inorgânico.

Figura 2 – Fotomicrografias das linhagens brasileiras utilizadas neste trabalho. *Sphaerospermopsis torques-reginae* ITEP-024 (A), *Nostoc* sp. CENA67 (B) e *Fischerella* sp. CENA161 (C). Barras: A e C = 10 µm; B = 20 µm. Fonte: Laboratório de Biologia Celular e Molecular CENA/USP

5.2 Preparo do sequenciamento

As três linhagens cianobacterianas foram lavadas com solução de lavagem (NaCl 50 mM, Tris-HCl 10 mM pH 7,5, EDTA 2,5 mM pH 8,0 e etanol 50%) para retirada de contaminantes. O DNA foi extraído e preparado para sequenciamento com AxyPrep Bacterial Genomic DNA Miniprep Kit (Axygen Biosciences, Union City, CA, EUA). Cerca de 1 ng de DNA foi utilizado para o preparo de bibliotecas genômicas com leituras pareadas 2x300pb utilizando o kit Nextera XT DNA Sample Prep Kit (Illumina Inc., San Diego, CA, EUA) de acordo com as instruções do fornecedor. A biblioteca foi sequenciada com Miseq Reagent Kit (Illumina) no Centro de Genômica Funcional Aplicada à Agropecuária e à Agroenergia, Campus “Luiz de Queiroz” da Universidade de São Paulo, em Piracicaba-SP.

5.3 Análises genômicas

5.3.1 Análise de qualidade das sequências e montagem do genoma

A qualidade das leituras obtidas com os sequenciamentos genômicos da plataforma MiSeq foi verificada e gráficos de análise foram gerados com o programa FastQC 0.10.1 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Os pares das sequências com complementaridade foram unidos utilizando o programa PEAR 0.9.6

(ZHANG et al., 2014) e as bases com qualidade Phred inferior a 30 foram filtradas com Seqclean 1.3.12 (<http://cores.ibest.uidaho.edu/software/seqclean>). Em seguida, os dados de cada sequenciamento foram montados individualmente com o programa SPAdes 3.1.1 (BANKEVICH et al., 2014) e os resultados das montagens foram analisados com QUAST 2.3 (GUREVICH et al., 2013). Após a montagem, foi realizada nova filtragem de sequências com identidade com cianobactérias, utilizando o programa Kraken (WOOD; SALZBERG, 2014). As montagens que geraram o menor número total de sequências contíguas, a sequência contígua mais longa e o maior valor de N50 foram selecionadas para as buscas dos genes de interesse.

5.3.2 Linhagens com genomas depositados em bancos de dados

Trinta e um genomas de cianobactérias da ordem Nostocales depositados no banco de dados GenBank, do NCBI (Tabela 2) contendo os agrupamentos gênicos relacionados à FBN foram selecionados e utilizados na análise comparativa dos genes da FBN.

Tabela 2 - Relação de genomas de cianobactérias da ordem Nostocales com sequenciamento completo e incompleto disponíveis no banco de dados GenBank em março de 2014 (Continua)

Genomas completos	Ambiente isolado	Tamanho (Mpb)	(G+C)%	Genes	NCBI Projeto ID	Referência
<i>Anabaena/Nostoc</i> sp. PCC 7120	água doce	6,41	41,3	6.222	244	Kaneko et al. 2001
<i>Anabaena variabilis</i> ATCC 29413	água doce	7,11	41,4	5.813	10642	Thiel et al. 2014
<i>Anabaena</i> sp. 90	água doce	5,31	38,1	4.797	30803	Wang et al., 2012
<i>Anabaena cylindrica</i> PCC 7122	água doce	7,06	38,8	6.258	43355	Shih et al., 2013
<i>Calothrix</i> sp. PCC 6303	água doce	6,96	39,8	5.841	158041	Shih et al., 2013
<i>Calothrix</i> sp. PCC 7507	pântano	7,02	42,2	6.250	158683	Shih et al., 2013
<i>Cylindrospermum stagnale</i> PCC 7417	solo	7,61	42,2	6.738	158809	Shih et al., 2013
<i>Nostoc</i> sp. PCC 7107	água doce	6,33	40,4	5.538	158705	Shih et al., 2013
<i>Nostoc azollae</i> 0708	endossimbionte de <i>Azolla filiculoides</i>	5,35	38,4	5.379	30807	Ran et al. 2010
<i>Nostoc punctiforme</i> PCC 73102	simbionte de <i>Macrozamia riedlei</i>	8.23	41,3	7.164	216	Ekman et al. 2013
<i>Nostoc</i> sp. PCC 7524	fonte termal	6.64	41,5	5.687	158707	Shih et al., 2013
<i>Rivularia</i> sp. PCC 7116	marinha	8,73	37,5	6.946	63147	Shih et al., 2013

Tabela 2 - Relação de genomas de cianobactérias da ordem Nostocales com sequenciamento completo e incompleto disponíveis no banco de dados GenBank em março de 2014 (Conclusão)

Genomas incompletos	Ambiente isolado	Tamanho (Mpb)	(G+C) %	Genes	NCBI Projeto ID	Referência
<i>Anabaena</i> sp. PCC 7108	marinha	5,89	38,8	5.263	158737	Shih et al., 2013
<i>Calothrix</i> sp. PCC 7103	–	11.58	38,6	10.120	159495	Shih et al., 2013
<i>Cylindrospermopsis raciborskii</i> CS-505	água doce	3,88	40,2	3.501	40109	Stucken et al., 2010
<i>Chlorogloeopsis</i> sp. PCC 9212	água termal	7,65	41,5	6.749	104963	Dagan et al, 2012
<i>Chlorogloeopsis fritschii</i> PCC 6912	solo	7,75	41,5	6.914	104961	Dagan et al, 2012
<i>Fischerella muscicola</i> PCC 7414	água termal	6,9	41,2	6.138	104967	Dagan et al, 2012
<i>Fischerella muscicola</i> PCC 73103	solo	7,36	40,3	6.384	104965	Dagan et al, 2012
<i>Fischerella thermalis</i> PCC 7521	água termal	5,44	41	4.684	104969	Dagan et al, 2012
<i>Fischerella</i> sp. JSC-11	água termal	5,38	41,1	4.671	61093	Shih et al., 2013
<i>Fischerella</i> sp. PCC 9339	–	8,4	40,1	6.801	159505	Shih et al., 2013
<i>Fischerella</i> sp. PCC 9431	–	7,14	40,2	–	158821	Shih et al., 2013
<i>Fischerella</i> sp. PCC 9605	rocha calcárea	8,2	42,6	–	158819	Shih et al., 2013
<i>Mastigocladopsis repens</i> PCC 10914	–	6,47	43,5	5.757	158735	Shih et al., 2013
<i>Mastigocoleus testarum</i> BC008	marinha	15,87	48,2	–	163055	Não publicado
<i>Microchaete</i> sp. PCC 7126	água doce	5,74	42,2	5.339	158817	Shih et al., 2013
<i>Nodularia spumigena</i> CCY 9414	marinha	5,32	41,3	4.904	13447	Voß et al., 2013
<i>Raphidiopsis brookii</i> D9	água doce	3,19	40,1	3.057	40111	Stucken et al., 2010
<i>Scytonema hofmanni</i> PCC 7110	rocha calcárea	12,01	41,5	10.706	157363	Dagan et al, 2012
<i>Scytonema hofmanni</i> UTEX 2349	rocha	8,18	41,2	–	63425	Shih et al., 2013

– origem não encontrada

5.3.3 Busca do agrupamento gênico envolvido na fixação biológica de nitrogênio e diferenciação de heterócitos

Os contigs gerados nas montagens dos genomas foram submetidos à predição gênica automática utilizando o programa Prokka 1.8 (SEEMAN, 2014). Todas as informações obtidas pela predição automática foram verificadas pela anotação manual usando o programa Artemis 16.0.0 (<http://www.sanger.ac.uk/Software/Artemis/>). As regiões de interesse foram identificadas por meio de alinhamento local com um banco de dados curado contendo os genes de interesse gerado a partir dos genomas das linhagens

Cylindrospermopsis raciborskii CS-505 e *Nostoc punctiforme* PCC 73102, utilizando a ferramenta BLASTn (ALTSCHUL et al., 1990).

A identificação do agrupamento gênico envolvido na formação do envelope glicolipídico dos heterócitos, uma policetídeo sintase tipo 1, foi realizada utilizando o servidor antiSMASH 3.0.0 (WEBER et al., 2015), que fornece predição de agrupamentos gênicos de metabólitos secundários.

5.3.4 Reconstrução filogenética dos genes

Todos os genes da FBN anotados manualmente foram comparados filogeneticamente. Um alinhamento múltiplo foi construído para cada sequência gênica utilizando o programa ClustalW (THOMPSON et al., 1994). Após a conversão deste arquivo em um formato nexus com o programa ALTER (GLEZ-PEÑA et al., 2010), foi realizada a construção de árvores filogenéticas. O melhor modelo evolutivo foi escolhido pelo programa jModelTest (POSADA, 2008) e árvores concatenadas e individuais foram reconstruídas por inferência bayesiana por meio do programa MrBayes 3.2.5 (RONQUIST; HUELSENBECK, 2003), com 5 milhões de gerações, 2 corridas independentes e 4 cadeias. A visualização das árvores foi feita com o programa FigTree 1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree>) e a edição, com Inkscape 0.48.5 (<http://inkscape.org>).

As reconstruções foram comparadas com árvores filogenéticas elaboradas com o gene codificador para subunidade RNAr 16S e com o filogenoma reportando a mais nova proposta de classificação taxonômica de cianobactérias (KOMÁREK et al., 2014).

6 RESULTADOS E DISCUSSÃO

6.1 Sequenciamento e montagem dos genomas

O sequenciamento genômico das linhagens *Sphaerospermopsis torques-reginae* ITEP-024, *Nostoc* sp. CENA67 e *Fischerella* sp. CENA161, realizado com a plataforma MiSeq, da Illumina, apresentou altos valores de qualidade e cobertura, conferindo confiabilidade aos dados gerados e à montagem. A figura 2, obtida a partir do programa FastQC, mostra a qualidade do sequenciamento das linhagens, baseada nos índices de qualidade Phred das bases sequenciadas. O comprimento da maioria das sequências no conjunto de dados final ficou entre 251 e 300 pb. Para a linhagem ITEP-024 foi obtido um total de 28 GB de dados, contendo 52.789.680 leituras. Para CENA67, foram obtidos 11 GB de dados, com 39.254.914 leituras. Já para a linhagem CENA161, foi obtido um conjunto de dados de 16 GB, com 49.088.820 leituras.

O sequenciamento de cianobactérias é dificultado pela dificuldade de obter culturas axênicas, de modo que a montagem de genomas cianobactérias é realizada partindo de uma abordagem metagenômica, na qual as sequências de todo o metagenoma são montadas e depois apenas aquelas sequências com afiliação a sequências de cianobactérias são separadas e utilizadas para análises finais sobre o organismo. Segundo Luo et al. (2011), essa abordagem de montagem pode ser considerada segura quando a cobertura desse genoma é maior que 20 x dentro do metagenoma, caso contrário, as chances de montagens quiméricas são maiores.

As montagens dos genomas das cianobactérias não resultaram em genomas fechados, entretanto gerou sequências longas e de confiança para serem trabalhadas. O fator que geralmente mais contribui para a dificuldade de gerar genomas completos de cianobactérias é a grande abundância de regiões repetitivas no genoma destes organismos, a qual, mesmo com a alta cobertura do sequenciamento genômico, contribui para que a maior parte dos genomas cianobacterianos disponíveis esteja incompleta (WANG et al., 2012). A montagem da linhagem ITEP-024 gerou 78 sequências contíguas, totalizando 5,22 Mpb, com N50 de 171.963, maior contig com 554.390 pb, índice GC de 37,5%, cobertura média de 843x e 4962 genes únicos preditos. Já a montagem da linhagem CENA67 gerou 267

contigs, com um tamanho total de 8,2 Mpb, com N50 de 58.528, maior contig com 221.053 pb, índice GC de 41,4%, cobertura média de 224x e 7.190 genes preditos. A linhagem CENA161, ao fim da montagem, apresentou 429 contigs, com um total de 7,2 Mpb, com N50 de 26.856, maior contig com 169.947 pb, índice GC de 40,27%, cobertura média de 79x e 5.938 genes preditos.

Figura 2 - Qualidade de leituras geradas por sequenciamento com MiSeq das linhagens ITEP-024 A e B: Conjunto de pares R1 e R2, respectivamente; CENA67 C e D: Conjunto de pares R1 e R2, respectivamente; e CENA161: E e F Conjunto de pares R1 e R2, respectivamente

Comparando o tamanho do genoma e o percentual GC dos isolados brasileiros com outros genomas relacionados disponíveis nos bancos de dados públicos, como o GenBank, observou-se que a linhagem CENA161 (Tabela 3) apresentou valores dentro do esperado para o gênero *Fischerella*, apresentando maior similaridade com as linhagens originárias de ambiente terrestre, como a linhagem PCC 73103.

Tabela 3 – Características gerais do genoma da *Fischerella* sp. CENA161 comparado com sequências de nucleotídeos de linhagens relacionadas ao gênero *Fischerella*

Linhagem	Origem	Tamanho (Mpb)	(G+C)%
<i>Fischerella</i> sp. CENA161	Água de nascente	7,2	40,27
<i>Fischerella muscicola</i> PCC 7414	água termal	6,9	41,2
<i>Fischerella muscicola</i> PCC 73103	solo	7,36	40,3
<i>Fischerella thermalis</i> PCC 7521	água termal	5,44	41
<i>Fischerella</i> sp. JSC-11	água termal	5,38	41,1
<i>Fischerella</i> sp. PCC 9339	–	8,4	40,1
<i>Fischerella</i> sp. PCC 9431	–	7,14	40,2
<i>Fischerella</i> sp. PCC 9605	rocha calcárea	8,2	42,6

– origem não encontrada

O genoma da ITEP-024 apresentou, junto com a linhagem de *Rivularia* sp. PCC 7116, o menor percentual GC dentre os genomas atualmente disponíveis para a ordem Nostocales, porém como é a primeira linhagem do gênero *Sphaerospermopsis* a ser sequenciada, não há outra relacionada para comparação. Já a linhagem CENA67 (Tabela 4) apresentou índice GC dentro do esperado e o tamanho do genoma apresentou tamanho semelhante ao da linhagem *Nostoc punctiforme* PCC 73102.

Tabela 4 – Características gerais do genoma de *Nostoc* sp. CENA67 comparado com sequências de nucleotídeos de linhagens do gênero *Nostoc*

Linhagem	Origem	Tamanho (Mpb)	(G+C)%
<i>Nostoc</i> sp. CENA67	terra preta antropogênica	8,2	41,4
<i>Nostoc azollae</i> 0708	endossimbionte de <i>Azolla filiculoides</i>	5,35	38,4
<i>Nostoc punctiforme</i> PCC 73102	simbionte de <i>Macrozamia riedlei</i>	8,23	41,4
<i>Nostoc</i> sp. PCC 7524	fonte termal	6,64	41,5
<i>Anabaena/Nostoc</i> sp. PCC 7120	água doce	6,41	41,3
<i>Nostoc</i> sp. PCC 7107	água doce	6,33	40,4

6.2 Anotação de genes de Fixação Biológica de Nitrogênio

Dentre as linhagens analisadas, a *Calothrix* sp. PCC 6303, que possui genoma completo, apresentou possíveis erros de montagem na região do agrupamento gênico da FBN, como genes incompletos e regiões invertidas sem a presença de enzimas invertases, que prejudicaram a anotação manual. A linhagem *Mastigocoleus testarum* BC008 apresentou apenas alguns dos genes, em contigs muito curtos que não permitiram a anotação completa do agrupamento gênico. Já as linhagens *Fischerella* sp. PCC 9605, *Calothrix* sp. PCC 7103 e *Raphidiopsis brookii* D9 não apresentaram genes relacionados ao processo de fixação de nitrogênio. Embora tal resultado já fosse esperado para a linhagem D9 (STUCKEN et al., 2010), dada a incapacidade do gênero de produzir heterócitos, a falta desses genes para as linhagens PCC 9605 e PCC 7103, pertencentes a dois gêneros tipicamente fixadores de nitrogênio, foi inesperada. Entretanto, seriam necessários dados da fisiologia e ecologia dessas linhagens para proposição de explicação dessas ausências.

De modo geral, todas as outras 27 linhagens apresentaram um agrupamento gênico de 15 kpb codificante para uma nitrogenase dependente de cofator MoFe, ou seja, a nitrogenase dependente de Molibdênio e a mais

amplamente distribuída entre os organismos diazotróficos; entretanto, algumas particularidades foram encontradas para algumas linhagens.

No isolado brasileiro de *Sphaerospermopsis torques-reginae* ITEP-024, os genes da FNB foram encontrados distribuídos em dois contigs. O primeiro contig apresentou um agrupamento com os genes *nifT*, *nifZ*, *nifV* e outro agrupamento com os genes *nifB*, *fdxN*, *nifS*, *nifU* e parte do *nifH*; o segundo contig continha o que supõe-se ser a continuação do *nifH*, dois fragmentos do *nifD*, *nifK*, um gene codificador de uma proteína semelhante à Mo-nitrogenase, *nifE*, *nifN*, *nifX*, *duf269*, *duf683*, *nifW*, *hesA*, *hesB*, *fdxH* e *feoA*. Nesta linhagem, portanto, o gene *nifH* estava com sua sequência distribuída em dois contigs, não sendo possível determinar sua sequência de nucleotídeos completa.

Interrupções, ou regiões de excisão, no agrupamento gênico da FBN são conhecidas desde 1987 no genoma de *Nostoc* sp. PCC 7120 (GOLDEN et al., 1987), tendo seu funcionamento posteriormente descrito no genoma da mesma linhagem (BRUSCA et al., 1990). Esse tipo de evento é uma forma de recombinação mediada por enzimas da família das serinas e tirosinas, conhecidas como invertases e recombinases (NUNES-DÜBY et al., 1998). Essas enzimas são capazes de reconhecer regiões específicas e distantes no genoma e deletar ou inverter a sequência gênica entre estas regiões, controlando o processo de expressão gênica. O processo de rearranjo genômico é ativado durante o processo de diferenciação de heterócitos, de modo que os agrupamentos são apenas expressos após a retirada ou a inversão destas regiões, o que permite a leitura correta dos genes codificadores da nitrogenase ou transportadores (ADAMS; DUGGAN, 1999). Dessa forma, apenas após a retirada das regiões de interrupção o agrupamento se torna ativo e codificante. Não há dados na literatura que apontem a existência de regiões de excisão no agrupamento gênico de cianobactérias não heterocitadas, de modo que a interrupção gênica pode funcionar como mais um elemento regulatório exclusivo da fixação biológica de nitrogênio da ordem Nostocales.

O genoma da linhagem ITEP-024 apresentou pelo menos uma interrupção no gene *nifH* e uma interrupção de 15 kpb no gene *nifD*. Embora a linhagem mais próxima filogeneticamente, a *Cylindrospermopsis raciborskii* CS-505, não apresente nenhuma interrupção ou inversão no agrupamento gênico da FBN, outras linhagens também próximas, como a *Anabaena* sp. 90, apresentam duas regiões de excisão no gene *nifH*, uma de 5,8 kpb e outra de 80kpb, além de uma

interrupção de 15 kpb no *nifD*. Portanto, embora a sequência gênica da ITEP-024 não seja a mais próxima da *Anabaena* sp. 90, as duas apresentaram, dentre as Nostocales sequenciadas, a organização gênica mais semelhante (Figura 4).

Figura 4 - Agrupamento gênico responsável pela Fixação Biológica de Nitrogênio na linhagem ITEP-024, em comparação com outras linhagens filogeneticamente próximas, indicando as interrupções nos genes *nifH* e *nifD*

No isolado CENA67, o agrupamento gênico da FBN foi encontrado em um único contig. Os genes estavam também organizados em dois agrupamentos distantes 7,3 kpb entre si e orientados em direções opostas. A ordem dos genes encontrada foi basicamente a mesma observada na ITEP-024, porém o *nifH* está íntegro e há uma interrupção de 9,8 kpb no *nifD*. Essa conformação é muito semelhante à encontrada na maior parte do gênero *Nostoc* (Figura 5). Com exceção da linhagem *Nostoc azollae* 0708, que não apresentou interrupções no agrupamento, as demais linhagens apresentaram pelo menos uma interrupção no gene *nifD* e apenas a linhagem *Nostoc* sp. PCC 7524 apresentou o gene *nifH* interrompido.

Figura 5 - Agrupamento gênico responsável pela FBN na linhagem CENA67, comparado com outras linhagens filogeneticamente próximas, indicando a interrupção no gene *nifD* e a presença do gene da cianoglobina

Além de regiões de excisões, o agrupamento gênico da FBN apresentou, entre os genes *nifU* e *nifH*, um gene codificador para uma cianoglobina (*glnN*). Esse gene foi relatado pela primeira vez para uma cianobactéria durante a investigação do metabolismo de nitrogênio em *Nostoc commune* (POTTS et al., 1992) e foi descrito posteriormente durante o sequenciamento do genoma da linhagem *Synechocystis* sp. PCC 6803 (KANEKO et al., 1996). Inicialmente, devido ao pequeno banco de dados, esse gene foi relacionado a globinas de protozoários e vertebrados. Apenas com o aumento no número de sequenciamentos na última década, outras globinas foram identificadas nos genomas de várias linhagens de cianobactérias e microalgas (JOHNSON; LECOMTE, 2013). Embora os estudos sejam recentes e escassos, sabe-se que o gene *glnN* é regulado pelo metabolismo de nitrogênio (FLORES; HERRERO, 1994) e supõe-se que essa proteína atue como um aprisionador de oxigênio nos heterócitos, entregando moléculas de oxigênio no complexo terminal de citocromo oxidase, ajudando na produção de moléculas de ATP durante a privação de oxigênio, da qual depende o metabolismo de fixação biológica de nitrogênio (HILL et al. 1996).

O gene de cianoglobina foi encontrado nas linhagens *Nostoc* sp. PCC 7107, *Nostoc punctiforme* PCC 73102, *Cylindrospermum stagnale* PCC 7417, *Calothrix* sp. PCC 7507, *Nodularia spumigena* CCY 9414, *Scytonema hofmanni* PCC 7110 e UTEX 2349, *Fischerella* sp. PCC 9339 e *Mastigocladopsis repens* PCC 10914. Com exceção do gênero *Nodularia*, todos os outros são descritos por Hoffmann et al. (2005) como pertencentes a ambientes terrestres, epilíticos ou epifíticos, ou seja, ambientes que apresentam uma concentração maior de oxigênio e que poderia prejudicar o processo de fixação de nitrogênio. Dessa forma, a presença de globinas em algumas cianobactérias pode ser uma adaptação evolutiva desses organismos a esse tipo de ambiente.

No isolado CENA161, os genes envolvidos na FBN foram encontrados em dois contigs, um contendo o agrupamento dos genes *nifV*, *nifT* e *nifZ* e outro contendo o outro agrupamento completo. Diferentemente dos outros dois isolados brasileiros, não havia nesse agrupamento nenhuma região de excisão nos genes *nif* e essa linhagem apresentou também o gene *glnN* codificador para cianoglobinas (Figura 6).

Figura 6 - Agrupamento gênico responsável pela Fixação Biológica de Nitrogênio na linhagem CENA161, indicando a presença de genes de cianoglobina e a ausência de interrupções, em comparação com outras linhagens filogeneticamente próximas

Dentre as cianobactérias mais próximas filogeneticamente, a linhagem *Fischerella* sp. PCC 9339 é a mais próxima da CENA161 em relação à sequência nucleotídica do agrupamento gênico; entretanto, com relação à sintenia, a linhagem *Fischerella muscicola* PCC 73103 apresenta uma organização mais semelhante ao isolado brasileiro, embora não possua o gene *glnN*. Além do agrupamento responsável pela codificação da nitrogenase dependente de Molibdênio, a linhagem CENA161 apresentou um agrupamento gênico codificador da nitrogenase alternativa que utiliza Vanádio como cofator (V-nitrogenase).

A atividade da V-nitrogenase foi inicialmente identificada em cianobactérias na linhagem *Anabaena variabilis* ATCC 29413 (KENTEMICH et al., 1988). Os genes desta nitrogenase alternativa e a organização do agrupamento gênico só foram descritas mais tarde, quando foi identificado um pequeno agrupamento de 4 genes, *vnfDG*, *vnfK*, *vnfE* e *vnfN*, e os genes *vupABC*, que codificam transportadores de vanadato, bem como os genes repressores da Mo-nitrogenase, *vnfR1* e *vnfR2* (THIEL; PRATTE, 2013). Diferentemente do que é encontrado em V-nitrogenases de bactérias heterotróficas, em cianobactérias os genes *vnfD* e *vnfG* encontram-se fusionados. Além das linhagens CENA161 e ATCC 29413, apenas outras três linhagens apresentam o agrupamento da V-nitrogenase: *Fischerella muscicola* PCC 7414, *Fischerella* sp. PCC 9339 e *Chlorogloeopsis* sp. PCC 7702 (Figura 7). Entretanto, enquanto CENA161, PCC 7414 e ATCC 29413 apresentam ortólogos dos genes *vnfDGKEN*, bem como dos transportadores de vanadato (*vupABC*), a linhagem PCC 9339 perdeu os genes dos transportadores, enquanto a linhagem PCC 7702 perdeu parte do gene *vnfDG* (THIEL; PRATTE, 2014).

Figura 7 - Agrupamento gênico associado a V-nitrogenase na linhagem CENA161, indicando os genes estruturais, reguladores e transportadores, comparando com outras linhagens de cianobactérias onde estes genes já foram também identificados

6.3 Anotação de genes ligados à diferenciação de heterócitos

Na linhagem ITEP-024, foram encontrados os genes *hetR* e *hetP*, principais responsáveis pela indução da diferenciação do heterócito (BUIKEMA; HASELKORN, 1991), os genes *patA*, *patN* e *patS*, responsáveis pelo padrão de diferenciação celular no filamento das cianobactérias (LIANG et al., 1992; GOLDEN; YOON, 1998; YOON; GOLDEN, 2001), e os genes *hetF*, *abp3* e *hetZ* (que participam da cascata de sinalização da diferenciação celular do heterócito (RISSER; CALLAHAN, 2008; ZHANG et al., 2007). Foram encontrados ainda os genes *ntcA* (MURO-PASTOR et al., 2002; CAMARGO et al., 2012), regulador global do metabolismo de nitrogênio, e o gene *nrrA*, que regula a transcrição do gene *hetR* (EHIRA; OHMORI, 2006).

Na linhagem CENA67, dentre os genes citados acima, apenas os genes *patS* e *abp3* não foram encontrados. Entretanto, o gene *patU3*, que atua na coordenação do padrão de diferenciação de heterócitos (ZHANG et al., 2007), foi encontrado apenas nesse isolado.

A anotação do isolado CENA161, contudo, resultou apenas na detecção dos genes reguladores principais *hetF*, *hetR* e *ntcA*. Tanto a falta de um banco de dados com sequências mais próximas do organismo estudado como a atual situação do genoma da linhagem, com um número maior de contigs curtos quando comparado aos genomas das linhagens ITEP-024 e CENA67, podem ter dificultado a identificação e anotação dos genes reguladores do processo de diferenciação de heterócitos deste isolado. Entretanto, a ausência desses genes em

outras cianobactérias próximas como as linhagens *Fischerella* sp. PCC 9339 e *Chlorogloeopsis* sp. PCC 7702, distantes filogeneticamente de *Sphaerospermopsis* e *Nostoc*, apresentando ramificações verdadeiras e talos mais complexos, pode ser indício de que esses táxons apresentam outros genes e formas de regulação da diferenciação de heterócitos.

O processo de anotação destes genes foi realizado com base na anotação manual por meio das ferramentas BLASTn e BLASTp, usando um banco de dados local curado. A construção deste banco de dados é muito importante, pois quanto mais próximo filogeneticamente o organismo-modelo utilizado, maiores são as chances de encontrar os genes de interesse. Com base nisso, utilizou-se o genoma da linhagem *Cylindrospermopsis raciborskii* CS-505 como referência para buscar os genes reguladores da diferenciação de heterócito na linhagem ITEP-024. Para a linhagem CENA67, utilizou-se como referência o genoma da *Nostoc punctiforme* PCC 73102. Já para a CENA161, não foi encontrado nenhum organismo de referência da família Fischerellaceae, de modo que tentou-se utilizar os genomas de cianobactérias da família Nostocaceae acima citadas, assim como tentou-se construir um banco de dados com os genes da linhagem *Chlorogloeopsis* sp. PCC7702, uma Chlorogleopsidaceae, mais próxima filogeneticamente da CENA161 e que possui um genoma mais completo e melhor montado e anotado, mas sem obter sucesso.

O heterócito é caracterizado morfológicamente pela presença de um envelope celular mais espesso para dificultar a entrada de oxigênio, que inibe o processo de fixação biológica de nitrogênio. Esse espessamento é proporcionado pela deposição interna de uma camada glicolipídica, que impede a entrada de O₂, e uma camada externa polissacarídica, que protege a camada interna de danos e diluição (ZHANG et al. 2006; XU et al., 2008). A camada interna é sintetizada por um agrupamento gênico contendo um gene que codifica uma poliketide sintase tipo 1 composto pelos genes *hglEGDCA* e *hetM* (STUCKEN et al, 2010). Esse agrupamento gênico (Figura 8) foi encontrado inteiro em um único contig na linhagem ITEP-024, enquanto que na linhagem CENA67 o gene *hglE* estava quebrado. A parte do primeiro domínio deste gene estava no final de um contig e a segunda parte, com o segundo domínio do gene, estava no início de outro contig. Entretanto, no genoma da linhagem CENA161 só foi encontrado parte do

agrupamento, com parte do gene *hglC* e os genes *hglA* e *hetM* no início de um contig.

Sphaerospermopsis torques-reginae ITEP-024

Nostoc sp. CENA67

Fischerella sp. CENA161

Figura 8 - Agrupamento gênico responsável pela codificação da camada glicolipídica nos três isolados brasileiros de cianobactérias

A formação da camada polissacarídica dos heterócitos é coordenada pelos genes *devACB* que codifica para um transportador de glicolipídios, o qual confere também estabilidade a esta camada (FIEDLER et al, 1998; STARON; MALDENER, 2012). Esse agrupamento foi encontrado nos genomas das três linhagens, CENA67, CENA161 e ITEP-014 (Figura 9).

Sphaerospermopsis torques-reginae ITEP-024

Nostoc sp. CENA67

Fischerella sp. CENA161

Figura 9 - Agrupamento gênico envolvido na formação da camada polissacarídica em heterócitos de três isolados brasileiros de cianobactérias

6.4 Análises filogenéticas dos genes de fixação biológica de nitrogênio

Uma árvore filogenética foi gerada por inferência bayesiana concatenando os 22 genes encontrados relacionados à FBN, utilizando linhagens de

cianobactérias da ordem Nostocales disponíveis no banco de dados do NCBI, que foram anotados, bem como as três linhagens de isolados brasileiros. Além das Nostocales, foi utilizada uma linhagem de *Cyanothece* sp. PCC 8802 e outra de *Trichodesmium erythraeum* IMS-101 para avaliar se alguma das cianobactérias estudadas apresentaria genes de fixação mais relacionados aos genes de cianobactérias de outras ordens, como Oscillatoriales ou Chroococcales, e como grupo externo para as análises foi utilizada a linhagem da bactéria *Wolinella succinogenes* DSMZ 1740, descrita por Baar et al. (2003) como possuidora de agrupamento gênico da FBN resultante de transferência horizontal a partir de uma cianobactéria. Dessa forma, a análise teria um grupo externo com genes próximos, mas de uma linhagem distante filogeneticamente.

A árvore filogenética resultante da análise desses genes concatenados (Figura 10) mostra que as linhagens se agrupam seguindo um padrão semelhante às análises do gene RNAr 16S e filogenômicas.

O isolado ITEP-024 ficou proximamente relacionado à linhagem *C. raciborskii* CS-505, o que corrobora a filogenia apresentada para o grupo, uma vez que esta é a linhagem mais próxima ao genoma sequenciado disponível (WERNER et al., 2012). As sequências do gênero *Fischerella* também se agruparam formando um clado bem definido, indicando que a linhagem PCC 9339 tem o agrupamento gênico de FBN mais próximo do isolado CENA161. A análise do RNAr 16S também mostrou que essas duas linhagens são próximas filogeneticamente. Já os agrupamentos das linhagens de *Nostoc* não se agruparam de forma bem definida, o que pode ser um reflexo da indefinição taxonômica de algumas linhagens do clado e da conhecida polifilia do gênero. O isolado CENA67, que na análise de RNAr 16S agrupou-se com a linhagem *N. punctiforme* PCC 73102 dentro de um clado com outras linhagens do gênero, aqui mostrou-se como um grupo externo ao grupo onde estão situadas a maioria das linhagens de *Nostoc*.

Comparando esta árvore concatenada com a árvore gerada em Komárek et al. (2014), da qual deriva a atual proposta de sistema de classificação cianobacteriano, criada com base na concatenação de sequências de 31 proteínas conservadas, é possível observar que a topologia das duas árvores de Nostocales é bastante semelhante, sendo que ambas divergem da organização dos agrupamentos das linhagens de Nostocaceae, *Nostoc* e *Anabaena*, na árvore por inferência bayesiana do gene RNAr 16S (Figura 11).

O resultado apresentado pela inferência bayesiana do gene RNAr 16S mostra valores de probabilidades posteriores mais baixos que a árvore dos genes de fixação concatenados. A análise do gene RNAr 16S separou bem aqueles grupos que apresentam ramificações verdadeiras – *Fischerella*, *Chlorogloeopsis* e *Mastigocladopsis* – e que já são reconhecidamente distantes dos que não formam ramificações.

Figura 10 – Árvore filogenética de 22 genes concatenados do agrupamento gênico de FBN gerada por inferência bayesiana. Nos nós dos ramos encontram-se as probabilidades posteriores bayesianas

Figura 11 – Árvore filogenética do gene de RNAr 16S das linhagens utilizadas neste estudo, gerada por meio de inferência bayesiana

O pouco número de linhagens sequenciadas dentro de um grupo tão diverso sistematicamente pode comprometer as inferências, de modo que a presença de sequências de mais grupos com ramificações verdadeiras e falsas e que formam outros tipos de organização do talo poderia melhorar a resolução desta análise. Mesmo assim, o RNAr 16S, embora seja considerado um bom marcador

para níveis taxonômicos acima de gênero, não reconstruiu bem a filogenia entre as linhagens da família Nostocaceae em comparação com análises filogenômicas.

A filogenia de cianobactérias é bastante conturbada e tem passado grandes mudanças nas últimas décadas, principalmente após a popularização da análise de genes marcadores como os codificadores para RNAr 16S, regiões intergênicas (ITS) e proteínas estruturais como RNA polimerase (*rpoC*), DNA girase (*gyrB*) e RecA. Entretanto, o gene do RNAr 16S tem sido o mais utilizado para a reconstrução filogenética de bactérias dadas as características do gene, como seu tamanho, de aproximadamente 1500 pb, a presença de regiões conservadas e variáveis e a existência de um grande banco de dados que permite a comparação de linhagens. As árvores filogenéticas baseadas no RNAr 16S, na maioria das vezes, corroboram a filogenia obtida a partir de genomas completos para categorias taxonômicas ao nível de gênero e acima (KONSTANTINIDIS; TIEDJE, 2007). Nos últimos anos, entretanto, as análises simultâneas de vários genes ou análises genômicas têm sido apontadas como o caminho mais promissor para a resolução de entraves na sistemática de vários grupos bacterianos, incluindo cianobactérias (CHUN; RAINEY, 2014).

Alguns autores têm relatado o gene *nifH* como um marcador filogenético, sugerindo que há uma correlação entre a topologia das árvores geradas com sequências de *nifH* e a de outras obtidas utilizando sequências do RNAr 16S (ZEHR; CAPONE, 1996; ZEHR et al., 2003; ZEHR et al., 1997; BEM-PORATH; ZEHR, 1994). Entretanto, trabalhos como o de Genuário (2010), com sequenciamento do gene *nifH* de cianobactérias, não mostraram congruência entre os agrupamentos formados nas árvores filogenéticas a partir deste gene e o posicionamento das linhagens de acordo com o sistema de classificação e do gene RNAr 16S. O mesmo ocorreu quando foi contruída uma árvore filogenética utilizando apenas o gene *nifH* das linhagens anotadas, ocorrendo o agrupamento de linhagens distantemente relacionadas filogeneticamente (Figura 12).

Figura 12 – Árvore filogenética do gene *nifH* gerada por inferência bayesiana, revelando agrupamento de linhagens unicelulares e homocitadas dentro do grupo de heterocitadas. Nos nós dos ramos encontram-se as probabilidades posteriores bayesianas

Três hipóteses podem explicar a ausência de correlação entre as análises de *nifH* e rRNA 16S. A primeira é que a capacidade de fixação teve origens múltiplas, o que não é mais bem aceito na comunidade científica (LATYSHEVA et al., 2012). A segunda seria o surgimento ancestral da fixação de nitrogênio seguida por diversos eventos de perda, dando origem aos diversos ramos filogenéticos (YOUNG, 2000). E a terceira hipótese é que o caráter da fixação teve uma origem única, se estendendo a outros ramos por transferência lateral. A hipótese hoje mais aceita é uma fusão das duas últimas hipóteses, de modo que os eventos de transferência lateral, junto com os eventos de convergência evolutiva causadas

pelas pressões de seleção, poderiam explicar as diversas incongruências na filogenia do gene *nifH*.

Dentre todas as reconstruções filogenéticas utilizando cada gene individualmente (ver Apêndice), apenas as árvores resultantes dos genes *nifE* (Figura 13) e *nifN* (Figura 14) refletiram a real filogenia do grupo estudado. Entretanto, o número de sequências destes genes depositadas em bancos de dados públicos é pequeno e pode comprometer este tipo de análise em larga escala. Os genes *nifEN* estão envolvidos na biossíntese do cofator MoFe e são genes parálogos aos genes estruturais *nifDK*. A comparação das sequências destes genes revela uma significativa homologia, comprovando que estes genes evoluíram de um ancestral comum, a partir de eventos de duplicação gênica (FANI et al. 2000).

Estas duplicações teriam ocorrido a partir de um gene *nifD* ancestral que teria se diversificado, formando o *nifK*. Uma outra duplicação dos genes *nifDK* teria dado origem aos genes *nifEN* (FANI et al., 2000; RAYMOND et al. 2004; BOYD et al., 2011). Estudo realizado por Boyd et al (2011) sugere que esses eventos de duplicação ocorreram em Archaea, nas classes Methanobacteriales e Methanococcales, e foram transferidos para outros organismos.

Por conta dessa origem e por funcionar como um molde para a construção da enzima que constrói o cofator essencial para a realização do processo de fixação de nitrogênio, os genes *nifEN* possivelmente estão suscetíveis a pressões de seleção diferentes daquelas sofridas pelos genes *nifDHK*, de modo que aparentam apresentar maior conservação ao longo de sua história evolutiva. Dessa forma, os genes *nifE* ou *nifN* poderiam ser melhores marcadores filogenéticos para cianobactérias da ordem Nostocales que o *nifH* e *nifD*, mais utilizados atualmente para estes tipos de estudo.

Figura 13 – Árvore filogenética do gene *nifE* gerada por inferência bayesiana que revela a congruência com as análises de filogenômicas e dos genes concatenados. Nos nós dos ramos encontram-se as probabilidades posteriores bayesianas

Figura 14 – Árvore filogenética do gene *nifN* gerada por inferência bayesiana, indicando topologia parecida com a árvore de genes FBN concatenados. Nos nós dos ramos encontram-se as probabilidades posteriores bayesianas

As análises filogenéticas utilizando o consenso do agrupamento gênico completo da FBN não mostraram divergências do padrão taxonômico, sugerindo que os genes são transferidos verticalmente dentro deste grupo de cianobactérias. Para tentar ter uma visão mais ampla da transferência dos genes encontrados neste agrupamento, realizou-se a reconstrução filogenética por meio de inferência bayesiana com genes estruturais concatenados dos diferentes tipos de nitrogenase – Mo-nitrogenase (*nif*), V-nitrogenase (*vnf*) e Fe-Nitrogenase (*anf*) – de diferentes

organismos procariotos (Anexo A), além das linhagens utilizadas neste trabalho, utilizando os genes de síntese de bacterioclorofila (Bch) como grupo externo, seguindo a metodologia utilizada por Boyd et al. (2011).

A árvore resultante desta análise (Apendice) não mostra evidências de transferência lateral seja dentro do grupo das cianobactérias, seja entre estes organismos e outras bactérias e Archaea, embora haja registro de transferência lateral entre a cianobactéria homocitada *Microcoleus chthonoplastes* e a deltaproteobactéria *Desulfovibrio* (BOLHUIS et al., 2010), linhagens não utilizadas no presente estudo. Esse dado sugere que a fixação de nitrogênio em Nostocales é menos propensa a eventos de transferência lateral de genes que em outras bactérias, uma vez que *Microcoleus chthonoplastes*, uma cianobactéria filamentosa pertencente à ordem Oscillatoriales, é suscetível à transferência lateral. Possivelmente, a complexidade morfológica e fisiológica das Nostocales dificulta a transferência de genes de forma não deletéria.

Entretanto, o agrupamento das diferentes nitrogenases mostra que as sequências de V-nitrogenase e Fe-nitrogenase estão distribuídas em um clado no meio de sequências da Mo-nitrogenase, sugerindo que as nitrogenases alternativas podem ser derivações da nitrogenase dependente de Molibdênio (Figura 15). Ao aumentar a robustez e incluir novas sequências de nitrogenases alternativas exclusivas de cianobactérias, grupo bastante heterogêneo e cosmopolita, esse resultado corrobora os estudos de Boyd e Peter (2013) sobre a origem e evolução da nitrogenase.

Figura 15 – Árvore filogenética gerada por inferência bayesiana a partir dos genes concatenados *nif/anf/vnfHDK*

7 CONCLUSÕES

Este estudo apresenta dados sobre a estrutura do agrupamento gênico da fixação biológica de nitrogênio e de diferenciação de heterócitos de linhagens de cianobactérias brasileiras da ordem Nostocales. Estes resultados evidenciam que a estrutura deste agrupamento gênico é bem conservada, sofrendo poucas variações entre as linhagens da ordem. Esta é a primeira descrição do agrupamento gênico de FBN para o gênero *Sphaerospermopsis* e a primeira descrição de um agrupamento de nitrogenase alternativa em cianobactérias brasileiras, e que é pouco frequente no grupo de cianobactérias.

As análises filogenéticas realizadas evidenciam que a transferência vertical é predominante dentro do grupo das Nostocales. Além disso, os resultados mostram que o uso do gene *nifH* como um marcador filogenético deve ser feito com cautela, uma vez que é possível a ocorrência de agrupamentos que não refletem os padrões filogenéticos de marcadores como rRNA 16S. A partir destes dados, os genes *nifE* e *nifN* poderiam ser avaliados para sua eficácia como marcadores moleculares, sendo necessário o incremento do banco de dados destes genes.

A utilização de mais de um gene para inferência filogenética já tem sido estimulada na comunidade científica e os resultados da análise dos genes concatenados de todo o agrupamento gênico da FBN corroboram esse dado, ao mostrar que ele reflete bem a filogenia, quando comparada com as análises filogenômicas.

O resultado obtido a partir da inferência filogenética dos genes estruturais das nitrogenases sugere que as nitrogenases alternativas (V-nitrogenase e Fe-nitrogenase) são derivadas da Mo-nitrogenase de Archaea, sendo ainda a Fe-nitrogenase derivada de uma modificação da V-nitrogenase.

Por fim, este trabalho contribui para o melhor entendimento dos agrupamentos gênicos relacionados à FBN e diferenciação de heterócitos, revelando ainda a necessidade do sequenciamento de novos grupos de cianobactérias para aumentar a robustez destes dados.

REFERÊNCIAS

- ADAMS, D.G. Symbiotic Interactions. In: WHITTON, B.A.; POTTS, M. **The ecology of Cyanobacteria**. Netherlands: Kluwer Academic, p. 523-561, 2000.
- ADAMS, D.G.; DUGGAN, P.S. Heterocyst and akinete differentiation in cyanobacteria. **New Phytologist**, Malden, v. 144, p. 3-33, 1999.
- ALLEN, M.B. Simple conditions for growth of unicellular blue-green algae on plates. **Journal of Phycology**, Baltimore, v. 4, p. 1-4, 1968.
- ALLEN, M.B.; ARNON, D.I. Studies on nitrogen-fixing blue-green algae. I. Growth and nitrogen fixation by *Anabaena cylindrica* Lemm. **Plant Physiology**, Rockville, v. 30, p. 366-372, 1955.
- ALTSCHÜL, S.F.; GISH, W.; MILLER, W.; MYERS, E.W.; LIPMAN, D.J. Basic local alignment search tool. **Journal of Microbiological Methods**, Amsterdam, v. 215, n. 3, p. 403-410, 1990.
- ANAGNOSTIDIS, K.; KOMÁREK, J. Modern approach to the classification system of cyanophytes: 5- Stigonematales. Archiv für Hydrobiologie, **Algological Studies**, Stuttgart, v. 59, p. 1-73, 1990.
- BAAR, C.; EPPINGER, M.; RADDATZ, G.; SIMON, J.; LANZ, C.; KLIMMEK, O.; NANDAKUMAR, R.; GROSS, R.; ROSINUS, A.; KELLER, H.; JAGTAP, P.; LINKE, B.; MEYER, F.; LEDERER, H.; SCHUSTER, S. Complete genome and analysis of *Wolinella succinogenes*. **Proceedings of the National Academic of Sciences USA**, Washington, v. 100, p. 11690-11695, 2003.
- BANKEVICH, A.; NURK, S.; ANTIPOV, D.; GUREVICH, A.A.; DVORKIN, M.; KULIKOV, A.S.; LESIN, V.M.; NIKOLENKO, S.I.; PHAM, S.; PRJIBELSKI, A.D.; PYSHKIN, A.V.; SIROTKIN, A.V.; VYAHHI, N.; TESLER, G.; ALEKSEYEV, M.A.; PEVZNER, P.A. SPAdes: a new genome assembly algorithm and its applications to singlecell sequencing. **Journal of Computational Biology**, New York, v. 19, p. 455-477, 2014.
- BEN-PORATH, J.; ZHER, J. P. Detection and characterization of cyanobacterial *nifH* genes. **Applied and Environmental Microbiology**, Baltimore, v. 60, p. 880-887, 1994.
- BERGMAN, B.; CARPENTER, E.J. Nitrogenase confined to randomly distributed trichomes in the marine cyanobacterium *Trichodesmium*. **Journal of Phycology**, Baltimore, v. 27, p. 158-165, 1991.
- BERMAN-FRANK, I.; LUNDGREN, P.; CHEN, Y.-B.; KÜPPER, H.; KOLBER, Z.; BERGMAN, B.; FALKOWSKI, P. Segregation of nitrogen fixation and oxygenic photosynthesis in the marine cyanobacterium *Trichodesmium*. **Science**, Washington, v. 294, n. 5546, p. 1534-1537, 2001.
- BERMAN-FRANK, I.; LUNDGREN, P.; FALKOWSKI, P. Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria. **Research in Microbiology**, Paris, v. 154, p. 157-164, 2003.

BLANK, C.E.; SÁNCHEZ-BARACALDO, P. Timing of morphological and ecological innovations in the cyanobacteria – a key to understanding the rise in atmospheric oxygen. **Geobiology**, Oxford, v. 8, p. 1–23, 2010.

BÖHME, H. Regulation of nitrogen fixation in heterocyst-forming cyanobacteria. **Trends in Plant Science**, Oxford, v. 3, n. 9, p. 346-351, 1998.

BOHUNICKÁ, M.; PIETRASIAK, N.; JOHANSEN, J.R.; GÓMEZ, E.B.; HAUER, T.; GAYSINA, L.A.; LUKEŠOVÁ, A. *Roholtiella*, gen. nov. (Nostocales, Cyanobacteria) — a tapering and branching cyanobacteria of the family Nostocaceae. **Phytotaxa**, Auckland, v. 197, n. 2, p. 83-103, 2015.

BOLHUIS, H.; SEVERIN, I.; CONFURIUS-GUNS, V.; WOLLENZIEN, U. I. A.; STAL, L. J. Horizontal transfer of the nitrogen fixation gene cluster in the cyanobacterium *Microcoleus chthonoplastes*. **The ISME Journal**, Hampshire, v. 4, p. 121-130, 2009.

BOTHE, H.; TRIPP, H.J.; ZEHR, J.P. Unicellular cyanobacteria with a new mode of life: the lack of photosynthetic oxygen evolution allows nitrogen fixation to proceed. **Archives of Microbiology**, Heidelberg, v. 192, p. 783-790, 2010.

BOYD, E.S.; HAMILTON, T.L.; PETERS, J.W. An alternative path for the evolution of biological nitrogen fixation. **Frontiers in Microbiology**, Lausanne, v.2, n. 205, 2011.

BOYD, E.S.; PETERS, J.W. New insights into the evolutionary history of biological nitrogen fixation. **Frontiers in Microbiology**, Lausanne, v.4, n. 201, 2013.

BRÖCKER, M. J.; VIRUS, S.; GANSKOW, S.; HEATHCOTE, P.; HEINZ, D. W.; SCHUBERT, W.-D.; JAHN, D.; MOSER, J. ATP-driven reduction by darkoperative protochlorophyllide oxidoreductase from *Chlorobium tepidum* mechanistically resembles nitrogenase catalysis. **Journal of Biological Chemistry**, Berlin, 283, 10559–10567, 2008.

BRUSCA, J.S.; CHASTAIN, C.J.; GOLDEN, J.W. Expression of the *Anabaena* sp. Strain PCC 7120 *xisA* Gene from a Heterologous Promoter Results in Excision of the *nifD* Element. **Journal Of Bacteriology**, Washington, v. 172, p. 3925-3931, 1990

BUIKEMA, W. J.; HASELKORN, R. Expression of the *Anabaena* *hetR* gene from a copper-regulated promoter leads to heterocyst differentiation under repressing conditions. **Proceedings of the National Academic of Sciences USA**, Washington, v. 98, p. 2729–2734, 2001.

BURJA, A.M.; BANAIGS, B.; ABOU-MANSOUR, E.; BURGESS, J.G.; WRIGHT, P.C. Marine cyanobacteria: a prolific source of natural products. **Tetrahedron**, Oxford, v. 57, p. 9347-9377, 2001.

BURKE, D. H.; HEARST, J. E.; SIDOW, A. Early evolution of photosynthesis: clues from nitrogenase and chlorophyll iron proteins **Proceedings of the National Academy of Sciences of the USA**, Washington, v. 90, p. 7134–7138, 1993.

CAMARGO, S.; VALLADARES, A.; FLORES, E.; HERRERO, A. Transcription activation by NtcA in the absence of consensus NtcA-binding sites in an *Anabaena* heterocyst differentiation gene promoter. **Journal of Bacteriology**, Washington, v. 194, p. 2939–2948, 2012.

CARDONA, T.; BATTCHIKIVA, N.; ZHANG, P.; STENSJÖ, K.; ARO, E.-M.; LINDBLAD, P.; MAGNUSON, A. Electron transfer protein complexes in the thylakoid

membranes of heterocysts from the cyanobacterium *Nostoc punctiforme*.

Biochimica et Biophysics Acta, Amsterdam, v. 1787, p. 252-263, 2009.

CASTENHOLZ, R. W.; WATERBURY, J. B. Oxygenic photosynthetic bacteria (sect. 19), group I. Cyanobacteria. In: STALEY, J. T.; BRYANT, M. P.; PFENNING, N.; HOLT, J. P. (Ed.). **Bergey's of systematic bacteriology**. Baltimore: Williams and Wilkins, 1989, p. 1710-1728.

CHUN, J.; RAINEY, F.A. Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea. **International Journal of Systematic and Evolutionary Microbiology**, Reading, v. 64, p. 316-324, 2014.

COHEN, Y.; JØRGENSEN, B. B.; REVSBECH, N. P.; POPLAWSKI, R. Adaptation to hydrogen sulfide of oxygenic and anoxygenic photosynthesis among cyanocabteria. **Applied and Environmental Microbiology**, Baltimore, v. 51, p. 398-407, 1986.

DAGAN, T.; ROETTGER, M.; STUCKEN, K.; LANDAN, G.; KOCH, R.; MAJOR, P.; GOULD, S.B.; GOREMYKIN, V.V.; RIPPKA, R.; TANDEAU DE MARSAC, N.; GUGGER, M.; LOCKHART, P.J.; ALLEN, J.F.; BRUNE, I.; MAUS, I.; PÜHLER, A.; MARTIN, W.F. Genomes of Stigonematalean Cyanobacteria (Subsection V) and the Evolution of Oxygenic Photosynthesis from Prokaryotes to Plastids. **Genome Biology and Evolution**, Oxford, v. 5, p. 31–44, 2012.

DALSGAARD, T.; CANFIELD, D.E.; PETERSEN, J.; THAMDRUP, B.; GONZÁLEZ, J.A. N₂ production by the anammox reaction in the anoxic water column of Golfo Dulce, Costa Rica. *Nature*, London, v. 422, p. 606-608, 2003.

DILWORTH, M.J.; EADY, R.R.; ELDRIDGE, M.E. The vanadium nitrogenase of *Azotobacter chroococcum*. Reduction of acetylene and ethylene to ethane. **Biochemical Journal**, London, v. 249, p. 745–751, 1988.

EADY, R.R. The dinitrogen-fixing bacteria. In: BALLOWS, A.; TRUPER, H.G.; DWORKIN, M.; HARDER, W.; CHLEIFER, K. (Ed). **The Prokaryotes**. 2nd ed. New York: Springe-velag, v. 1, chap. 22, 1991.

EADY, R.R. The vanadium-containing nitrogenase of *Azotobacter*. **Biofactors**, Oxford, v. 1, p. 111–116, 1988.

EADY, R.R. Vanadium nitrogenases of *Azotobacter*. **Metal ions in biological systems**, New York, v. 31, p. 363–405, 1995.

EHIRA, S.; OHMORI, M. NrrA directly regulates expression of *hetR* during heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. **Journal of Bacteriology**, Washington, v. 188, n. 24, p. 8520–8525, 2006.

EKMAN, M.; PICOSI, S.; CAMPBELL, E.L.; MEEKS, J.C.; FLORES, E.A. *Nostoc punctiforme* sugar transporter necessary to establish a cyanobacterium-plant symbiosis. **Plant Physiology**, Rockville v. 161, p. 1984-92, 2013.

ESTEVEES, F.A.; AMADO, A.M. Nitrogênio. In: ESTEVES, F.A. (Org.) **Fundamentos de Limnologia**. 3th ed. Rio de Janeiro: Interciência, v.1, p. 239-258, 2011.

ETCHEGARAY, A.; RABELLO, E.; DIECKMANN, R.; MOON, D.H.; FIORE, M.F.; VON DÖHREN, H.; TSAI, S.M.; NEILAN, B.A. Algicide production by the filamentous cyanobacterium *Fischerella* sp. CENA19. **Journal of Applied Phycology**, Dordrecht, v. 16, p. 237-243, 2004.

FAY, P. ; STEWART, W.D.P.; WALSBY, A.E.; FOGG, G.E. Is the heterocyst the site of nitrogen fixation in blue-green algae? **Nature**, London, v. 200,p. 810-812, 1968.

FIEDLER, G.; ARNOLD, M.; HANNUS, S.; MALDENER, I. The DevBCA exporter is essential for envelope formation in heterocysts of the cyanobacterium *Anabaena* sp. strain PCC 7120. **Molecular Microbiology**, Oxford, v. 27, p. 1193–1202, 1998.

FIORE, M.F.; ALVARENGA, D.O.; VARANI, A.M.; HOFF-RISSETI, C.; CRESPI, E.; RAMOS, R.T.J.; SILVA, A.; SCHAKER, P.D.C.; HECK, K.; RIGONATO, J.; SCHNEIDER, M.P.C. Draft genome sequence of the Brazilian toxic bloom-forming cyanobacterium *Microcystis aeruginosa* SPC777. **Genome Announcements**, Washington, v.1, ed.4, e00547-13, 2013.

FIORE, M.F.; GENUÁRIO, D.B.; SILVA, C.S.P.; SHISHIDO, T.K.; MORAES, L.A.B.; NETO, R.C.; SILVA-STENICO, M.E. Microcystin production by a freshwater spring cyanobacterium of the genus *Fischerella*. **Toxicon**, Oxford, v. 53,p. 754–761, 2009

FIORE, M.F.; HONDA, R.Y. Fixação do N₂ por cianobactérias. In: FIGUEIREDO, M.V.B.; BURITY, H.A.; STAMFORD, N.P.; SANTOS, C.E.R.S. (Ed.).

Microrganismos e agrobiodiversidade: o novo desafio para a agricultura. Guaíba: Agrolivros, 2008. p. 69-95.

FLORES, E.; HERRERO, A. Assimilatory nitrogen metabolism and its regulation. In: Bryant, D.A. (Ed.), **The molecular biology of cyanobacteria**. Netherlands: Kluwer Academic Publishers, 1994. p. 487-517.

FLORES, E.; MURO-PASTOR, A. M.; HERRERO, A. Cyanobacterial nitrogen assimilation genes and NtcA-dependent control of gene expression. In: PESCHEK, L. W.; SCHMETTERER, G. (Ed.). **The phototrophic prokaryotes**. New York: Plenum Press, 1999. p. 487-517.

GALLOWAY, J.N.; DENTENER, F.J.; CAPONE, D.G.; BOYER, E.W.; HOWARTH, R.W.; SEITZINGER, S.P.; ASNER, G.P.; CLEVELAND, C.C.; GREEN, P.A.; HOLLAND, E.A.; KARL, D.M.; MICHAELS, A.F.; PORTER, J.H.; TOWNSEND, A.R.; VOROSMARTY, C.J. Nitrogen cycles: Past, present, and future. **Biogeochemistry**, Dordrecht, v. 70, p. 153– 226, 2004.

GENUÁRIO, D. B. **Cianobactérias em ecossistemas de manguezais: isolamento, morfologia e diversidade genética**. 2010. 97 P. Dissertação (Mestrado em Ciências) – Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Piracicaba, 2010.

GLEZ-PEÑA, D.; GÓMEX-BLANCO, D.; REBOIRO-JATO, M.; FDEZ-RIVEROLA, F.; POSADA, D. ALTER: program-oriented format conversion of DNA and protein alignmentss. **Nucleic Acids Research**, London, v. 38, p. 14-18, 2010.

GOLDEN, J.W.; YOON, H.S. Heterocyst formation in *Anabaena*. **Current Opinion in Microbiology**, Kidlington, v. 1, p. 623–629, 1998.

GOLDEN, J.W.; MULLIGAN, M.; HASELKORN, R. (1987) DiVerent recombination site specificity of two developmentally regulated genome rearrangements. **Nature**, London, v. 327, p. 526–529, 1987.

GORHAM, P.R.; MCLACHLAN, J. R.; HAMMER, V.T.; KIM, W.K. Isolation and culture of toxic strains of *Anabaena flos-aquae* (Lyngb.) **Breb. Verhandlungen der**

Internationalen Vereinigung fuer Theoretische und Angewandte. Limnologie, Toulouse, v. 15, p. 796-804, 1964.

GUERRERO, M.G.; VEGA, J.M.; LOSADA, M. The assimilatory nitrate-reducing system and its regulation. **Annual Review of Plant Physiology**, Palo Alto v. 32, p. 169-204, 1981.

GUREVICH, A.; SAVELIEV, V.; VYAHHI, N.; TESLER, G. QUASt: quality assessment tool for genome assemblies. **Bioinformatics**, Oxford, v. 29, p. 1072-1075, 2013.

HASELKORN, R.; BUIKEMA, W. J. Nitrogen fixation in cyanobacteria. In: STACEY, G.; BURRIS, R. H.; EVANS, H. J. (Ed.). **Biological nitrogen fixation**. New York: Chapman & Hall, p. 166-190, 1992.

HERRERO, A.; MURO-PASTOR, A.M.; FLORES, E. Nitrogen control in cyanobacteria. **Journal of Bacteriology**, Washington, v. 183, n. 2, p. 411-425, 2001.

HERRERO, A.; PICOSI, S.; FLORES, E. Gene expression during heterocyst differentiation. **Advances in Botanical Research**, London, v. 65, p. 281-329, 2001.

HILL, D. R.; BELBIN, T. J.; THORSTEINSSON, M. V.; BASSAM, D.; BRASS, S.; ERNST, A.; BÖGER, P.; PAERL, H.; MULIGAN, M.E.; POTTS, M. Gln (cyanoglobins) is a peripheral membrane protein that is restricted to certain *Nostoc* spp. **Journal of Bacteriology**, Washington, v. 178, p. 6587-6598, 1996.

HOFFMANN, L.; KOMÁREK, J.; KAŠTOVSKÝ, J. System of cyanoprokaryotes (Cyanobacteria): state in 2004. **Algological Studies**, Stuttgart, v. 117, p. 95-115, 2005.

HOLGUIN, G.; VAZQUEZ, P.; BASHAN, P. The role of sediment microorganisms in the productivity, conservation, and rehabilitation of mangrove ecosystems: an overview. **Biology and Fertility of Soils**, Berlin, v. 33, p. 265-278, 2001.

HUANG, T-C.; LIN, R-F.; CHU, M-K.; CHEN, H-M. Organization and expression of nitrogen-fixation genes in the aerobic nitrogen-fixing unicellular cyanobacterium *Synechococcus* sp. strain RF-1. **Microbiology**, Reading, v. 145, p. 743-753, 1999.

JETTEN, M.; SCHMID, M.; VAN DE PAS-SCHNOONEN, K.; SINNINGHE DAMSTÉ, J. STROUS, M. Anammox organisms: enrichment, cultivation, and environmental analysis. **Methods in Enzymology**, Berkely, v. 397, p. 34-57, 2005.

JOERGER, R.D.; JACOBSON, M.R.; PREMAKUMAR, R.; WOLFINGER, E.D.; BISHOP, P.E. Nucleotide sequence and mutational analysis of the structural genes (anfH, anfG, anfK) for the second alternative nitrogenase from *Azotobacter vinelandii*. **Journal of Bacteriology**, Washington, v. 171, p. 1075-1086, 1989.

JOHNSON, E.A.; LECOMTE, J.T.J. The Globins of Cyanobacteria and Algae. **Advances in Microbial Physiology**, Oxford, v. 63, p. 196-272, 2013.

KANEKO, T.; NAKAMURA, Y.; WOLK, C.P.; KURITZ, T.; SASAMOTO, S.; WATANABE, A.; IRIGUCHI, M.; ISHIKAWA, A.; KAWASHIMA, K.; KIMURA, T.; KISHIDA, Y.; KOHARA, M.; MATSUMOTO, M.; MATSUNO, A.; MURAKI, A.; NAKAZAKI, N.; SHIMPO, S.; SUGIMOTO, M.; TAKAZAWA, M.; YAMADA, M.; YASUDA, M.; TABATA, S. Complete genomic sequence of the filamentous

nitrogenfixing cyanobacterium *Anabaena* sp. strain PCC 7120. **DNA Research**, Tokyo, v. 8, p. 205-213, 2001.

KANEKO, T.; SATO, S.; KOTANI, H.; TANAKA, A.; ASAMIZU, E.; NAKAMURA, Y.; MIYAJIMA, N.; HIROSAWA, M.; SUGIURA, M.; SASAMOTO, S.; KIMURA, T.; HOSOUCHI, T.; MATSUNO, A.; MURAKI, A.; NAKAZAKI, N.; NARUO, K.; OKUMURA, S.; SHIMPO, S.; TAKEUCHI, C.; WADA, T.; WATANABE, A.; YAMADA, M.; YASUDA, M.; TABATA, S. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. **DNA Research**, Tokyo, v. 3, n. 3, p. 109–136, 1996.

KECHRIS, K. J.; LIN, J. C.; BICKEL, P. J.; GLAZER, A. N. Quantitative exploration of the occurrence of lateral gene transfer by using nitrogen fixation genes as a case study. **Proceedings of the National Academy of Sciences of the USA**, Washington, v. 103, p. 9584–9589, 2006.

KENTEMICH, T.; DANNEBERG, G.; HUNDESHAGEN, B.; BOTHE. Evidence for the occurrence of the alternative, vanadium-containing nitrogenase in the cyanobacterium *Anabaena variabilis*. **FEMS Microbiology Letters**, Amsterdam, v.51, p. 19-24, 1988.

KOMÁREK, J.; KAŠTOVSKÝ, J.; MAREŠ, J.; JOHANSEN, J.R. Taxonomic classification of cyanoprokaryotes (cyanobacterial genera) 2014, using a polyphasic approach. **Preslia**, Praha, v. 86, p. 295-335, 2014.

KONSTANTINIDIS, K.T; TIEDJE, J.M. Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead. **Current Opinion in Microbiology**, Kidlington, v. 10, p. 504-509, 2007.

KUYPERS, M.M.M.; SLIEKERS, A.O.; LAVIK, G.; SCHMID, M.; JORGENSEN, B.B.; KUENEN, J.G.; SINNINGHE DAMSTÉ, J.S.; STROUS, M.; JETTEN, M.S.M. Anaerobic ammonium oxidation by anammox bacteria in the Black Sea. **Nature**, London, v. 422, p. 608–611, 2003.

LAMMERS, P.J.; HASELKORN, R. Sequence of the *nifD* gene coding for the alpha subunit of dinitrogenase from the cyanobacterium *Anabaena*. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v. 80, p. 4723-4727, 1983.

LATYSHEVA, N.; JUNKER, V.L.; PALMER, W.J.; CODD, G.A.; BARKER, D. The evolution of nitrogen fixation in cyanobacteria. **Bioinformatics**, Oxford, v. 28, n. 5, p. 603-606, 2012.

LAURENT, S.; CHEN, H.; BÉDU, S.; ZIARELLI, F.; PENG, L.; ZHANG, C.C. Nonmetabolizable analogue of 2-oxoglutarate elicits heterocyst differentiation under repressive conditions in *Anabaena* sp. PCC 7120. **Proceeding of the National Academy of Sciences of the USA**, Washington, v. 102, n. 28, p. 9907-9912, 2005.

LEÃO, P.N.; ENGEL, N.; ANTUNES, A.; GERWICK, W.H.; VASCONCELOS, V. The chemical ecology of cyanobacteria. **Natural Product Reports**, Cambridge, v. 29, p. 372-391, 2012.

LEE, C.C.; HU, Y.; RIBBE, M.W. Unique Features of the Nitrogenase VFe Protein from *Azotobacter vinelandii*, **Proceedings of the National Academy of Sciences of the USA**, Washington, v. 106, n. 23, p. 9209-9214, 2009.

LIAIMER, A.; HELFRICH, E.J.N.; HINRICHS, K.; GULJAMOW, A.; ISHIDA, K.; HERTWECK, C.; DITTMANN, E. Nostopeptolide plays a governing role during cellular differentiation of the symbiotic cyanobacterium *Nostoc punctiforme*. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v. 112, p. 1862-1867, 2015.

LIANG, J.; SCARPINO, L.; HASELKORN, R. The *patA* gene product, which contains a region similar to CheY of *Escherichia coli* controls heterocyst pattern formation in the cyanobacterium *Anabaena* 7120. **Proceedings of the National Academic of Sciences USA**, Washington, v. 89, p. 5655–5659, 1992.

LUO, C.; TSEMENTZI, D.; KYRPIDES, N.C.; KONSTANTINIDIS, K.T. Individual genome assembly from complex community short-read metagenomic datasets. **The ISME Journal**, Oxon, v. 6, p. 898-901, 2011.

MAZUR, B.J.; CHUI, C.F. Sequence of the gene coding for the beta-subunit of dinitrogenase from the blue-green alga *Anabaena*. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v. 79, p. 6782-6786, 1982.

McFADDEN, G.I. Primary and secondary endosymbiosis and origin of plastids. **Journal of Phycology**, Baltimore, v. 36, p. 951-959, 2001.

MARGULIS, L. **Origin of eukaryotic cells**. New Haven: Yale University Press, 1970.

MEVARECK, M.; RICE, D.; HASELKORN, R. Nucleotide sequence of a cyanobacterial *NifH* gene coding for nitrogenase reductase. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v. 77, p. 6476-6480, 1980.

Molica, R.J.R.; Oliveira, E.J.A.; Carvalho, P.V.V.C.; Costa, A.N.S.F.; Cunha, M.C.C.; Melo, G.L.; Azevedo, S.M.F.O. Occurrence of saxitoxins and an anatoxin-a(s)-like anticholinesterase in a Brazilian drinking water supply. **Harmful Algae**, Amsterdam, v. 4, p. 743–753, 2005.

MOORE, R.E.; CHEUK, C.; PATTERSON, G.M.L. Hapalindoles: New alkaloids from the blue-green algae *Hapalosiphon fontinalis*. **Journal of American Chemical Society**, Easton, v. 106, p. 6456-6457, 1984.

MOREIRA, F.M.M.; SIQUEIRA, J.O. **Microbiologia e Bioquímica do Solo**. 2 ed., Lavras: Editora UFLA, 2006. p. 449-542.

MÜLLER, A.; SCHNEIDER, K.; KNÜTTEL, K.; HAGEN, W.R. EPR spectroscopic characterization of an 'iron only' nitrogenase. S = 3/2 spectrum of component 1 isolated from *Rhodobacter capsulatus*. **FEBS Letter**, Amsterdam, v. 303, p. 36–40, 1992.

MURO-PASTOR, A. M.; VALLADARES, A.; FLORES, E.; HERRERO, A. Mutual dependence of the expression of the cell differentiation regulatory protein HetR and the global nitrogen regulator NtcA during heterocyst development. **Molecular Microbiology**, Oxford, v. 44, p. 1377–1385, 2002.

MYLONA, P.V.; PREMAKUMAR, R.; PAU, R.N.; BISHOP, P.E. Characteristics of *orf1* and *orf2* in the *anfH*HDGK genomic region encoding nitrogenase 3 of *Azotobacter vinelandii*. **Journal of Bacteriology**, Washington, v. 178, p. 204–208, 1996.

NAMIKOSHI, M.; RINEHART, K.L. Bioactive compounds produced by cyanobacteria. **Journal of Industrial Microbiology and Biotechnology**, Hampshire, v. 17, p. 373–384, 1996.

NELSON, D.L.; COX, M.M. **Lehninger princípios de bioquímica**. Tradução: A.A. Simões e W.R.N. Lori. 3. ed. São Paulo: Sarvier, 2002. p. 639-640.

NEVES, M.C.P.; RUMJANECK, N.G. Bioquímica e fisiologia da fixação de nitrogênio. In: CARDOSO, E.J.B.N.; TSAI, S.M.; NEVES, M.C.P. **Microbiologia do Solo**. Campinas: Sociedade Brasileira de Ciências do Solo, 1992. p.141-156.

NUNES-DÜBY, S.E.; KWON, H.J.; TIRUMALAI, R.S.; ELLENBERGER, T.; LANDY, A. Similarities and differences among 105 members of the Int family of site-specific recombinases. **Nucleic Acids Research**, London, v. 26, p. 391–406, 1998,

OLIVER, R. L.; GANF, G. G. Freshwater blooms. In: WHITTON, B. A.; POTTS, M. (Ed.). **The ecology of cyanobacteria, their diversity in time and space**. Dordrecht: Kluwer Academic Publishers, 2000. p. 149-194.

PETERS, J.W.; FISHER, K.; DEAN, D. Nitrogenase structure and function: a biochemical-genetic perspective. **Annual Review of Microbiology**, Palo Alto, v. 45, p. 335-366, 1995

PICOSSI, S.; MONTESINOS, M.L.; PERNIL, R.; LICHTLÉ, C.; HERRERO, A.; FLORES, E. ABC-type neutral amino acid permease N-I is required for optimal diazotrophic and is repressed in the heterocysts of *Anabaena* sp. Strain PCC 7120. **Molecular Microbiology**, Oxford, v. 57, p. 1582-1592, 2005.

PLOMINSKY, Á.M.; LARSSON, J.; BERGMAN, B.; DELHERBE, N.; OSSES, I.; VÁSQUEZ, M. Dinitrogen Fixation Is Restricted to the Terminal Heterocysts in the Invasive Cyanobacterium *Cylindrospermopsis raciborskii* CS-505. **PLoS ONE**, San Francisco, v. 8, n. 2, e51682, 2013.

POSADA, D. jModelTest: Phylogenetic Model Averaging. **Molecular Biology and Evolution**, Oxford, v. 25, n. 7, p. 1253-1256, 2008.

POSTGATE, R. J. **Nitrogen fixation**. 3rd. ed. Cambridge: Cambridge University Press, 1998. 112 p.

POTTS, M.; ANGELONI, S. V.; EBEL, R. E.; BASSAM, D. Myoglobin in a cyanobacterium. **Science**, Washington, v. 256, p. 1690–1692, 1992.

RAI, A.N.; SÖDERBÄCK, E.; BERGMAN, B. Cyanobacterium-plant symbioses. **New Phytologist**, Malden, V, 147, p. 449-48, 2000.

RAMAKRISHNAN, B.; MEGHARAJ, M.; VENKATESWARLU, K.; NAIDU, R.; SETHUNATHAN, N. The impacts of environmental pollutants on microalgae and cyanobacteria. **Critical Reviews in Environmental Science and Technology**, Boca Raton, v. 40, n. 8, p. 699-821, 2010.

RAN, L.; LARSSON, J.; VIGIL-STENMAN, T.; NYLANDER, J.A.; ININBERGS, K.; ZHENG, W.W.; LAPIDUS, A.; LOWRY, S.; HASELKORN, R.; BERGMAN, B.

Genome erosion in a nitrogen-fixing vertically transmitted endosymbiotic multicellular cyanobacterium. **PLoS One**, San Francisco, v. 8, n. 5, e11486, 2010.

RAYMOND, J.; SIEFERT, J.L.; STAPLES, C.R.; BLANKENSHIP, R.E. The natural history of nitrogen fixation. **Molecular Biology and Evolution**, Chicago, v. 21, p. 541-554, 2004.

REIS, V.M.; TEIXEIRA, K.R.S. Fixação Biológica de Nitrogênio – Estado da Arte. In: AQUINO, A.M.; ASSIS, R.L. EMBRAPA INFORMAÇÃO TECNOLÓGICA.

Processos biológicos no sistema solo-planta: ferramentas para uma agricultura sustentável. Brasília: EMBRAPA Informação Tecnológica, 2005. p. 161-180.

REVSBECH, N.P.; RISGAARD-PETERSEN, N.; SCHRAMM, A.; NIELSEN, L.P. Nitrogen transformations in stratified aquatic microbial ecosystems. **Antonie van Leeuwenhoek**, Wageningen, v. 90, p. 361-275, 2006.

RISSER, D. D.; CALLAHAN, S. M. HetF and PatA control levels of HetR in *Anabaena* sp. strain PCC 7120. **Journal of Bacteriology**, Washington, v. 190, p. 7645–7654, 2008.

RONQUIST, F.; HUELSENBECK, J.P. MrBayes 3: Bayesian phylogenetic inference under mixed models. **Bioinformatics**, Oxford, v. 19, p. 1572-1574, 2003.

RUBIO, L. M.; LUDDEN, P. W. Maturation of Nitrogenase: a Biochemical Puzzle. **Journal of Bacteriology**, Washington, v. 187, n.2, p. 405–414, 2005.

SÁNCHEZ-BARACALDO, P.; HAYES, P. K.; BLANK, C. E. Morphological and habitat evolution in the Cyanobacteria using a compartmentalization approach. **Geobiology**, Oxford, v. 3, p. 145–165, 2005.

SEEMANN, T. Prokka: rapid prokaryotic genome annotation. **Bioinformatics**, Oxford, v. 30, p. 2068-2069, 2014.

SHIH, P.M.; WU, D.; LATIFI, A.; AXEN, S.D.; FEWER, D.P.; TALLA, E., CALTEAU, A.; CAI, F.; TANDEAU DE MARSAC, N.; RIPPKA, R.; HERDMAN, M.; SIVONEN, K.; COURSIIN, T.; LAURENT, T.; GOODWIN, L.; NOLAN, M.; DAVENPORT, K.W.; HAN, C.S.; RUBIN, E.M.; EISEN, J.A.; WOYKE, T.; GUGGER, M.; KERFELD, C.A. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. **Proceedings of the National Academic of Sciences USA**, Washington, v. 110, p. 1053-1058, 2013.

SHISHIDO, T.K. **Análise fenotípica, genética e de bioatividade de isolados brasileiros de cianobactérias dos gêneros Fischerella e Hapalosiphon.** 2009. 134p. Dissertação (Mestrado em Biologia na Agricultura e no Ambiente) - Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Piracicaba, 2009.

SILVA-STENICO, M.E.; SILVA, C.S.P.; LORENZI, A.S.; SHISHIDO, T.K.; ETCHEGARAY, A.; LIRA, S.P.; MORAES, L.A.B.; FIORE, M.F. Non-ribosomal peptides produced by Brazilian cyanobacterial isolates with antimicrobial activity. **Microbiological Research**, Amsterdam, v. 166, p. 161-175, 2011

SIQUEIRA, J.O.; FRANCO, A.A. **Biotecnologia do Solo – Fundamentos e Perspectivas.** Brasília, MEC-ESAL-FAEPE-ABEAS, 1988. p. 235.

STAL, L. J.; MOEZELAAR, R. Fermentation in cyanobacteria. **FEMS Microbiology Reviews**, Amsterdam, v. 21, p. 179-211, 1997.

STARON, P; MALDENER, I. All0809/8/7 is a DevDCA-like ABC-type efflux pump required for diazotrophic growth in *Anabaena* sp. PCC7120. **Microbiology**, Reading, v. 158, p. 2537-2545, 2012.

STOLZ, J.F.; BASU, P. Evolution of nitrate reductase: molecular and structural variations on a common function. **Chemical Biochemistry**, London, v. 3, p.198–206, 2002.

STUCKEN, K.; JOHN, U.; CEMBELLA, A.; MURILLO, A.A.; SOTO-LIEBE, K.; FUENTES-VALDÉS, J.J.; FRIEDEL, M.; PLOMINSKY, A.M.; VÁSQUEZ, M.; GLÖCKNER, G. The smallest known genomes of multicellular and toxic cyanobacteria: comparison, minimal gene sets for linked traits and the evolutionary implications. **PLoS ONE**, San Francisco, v. 5, n. 2, e. 9235, 2010.

THAMDRUP, B., DALSGAARD, T. Production of N₂ through anaerobic ammonium oxidation coupled to nitrate reduction in marine sediments. **Applied and Environmental Microbiology**, Washington, v. 68, p. 1312–1318, 2002.

THIEL, T. Characterization of Genes for an Alternative Nitrogenase in the Cyanobacterium *Anabaena variabilis*. **Journal of Bacteriology**, Washington, v. 175, n 19, p. 6276-6286, 1993.

THIEL, T.; LYONS, E. M.; ERKER, J. C.; ERNST, A. A second nitrogenase in vegetative cells of a heterocyst-forming cyanobacterium. **Proceedings of the National Academic of Sciences USA**, Washington, v. 92, p. 9358–9362, 1995.

THIEL, T.; PRATTE, B.S. Alternative Nitrogenases in *Anabaena variabilis*: The Role of Molybdate and Vanadate in Nitrogenase Gene Expression and Activity. **Advances in Microbiology**, Irvine, v. 3, p. 87-95, 2013

THIEL, T.; PRATTE, B.S. Regulation of Three Nitrogenase Gene Clusters in the Cyanobacterium *Anabaena variabilis* ATCC 29413. **Life**, Basel, v. 4, p. 944-967, 2014.

THIEL, T.; PRATTE, B. S.; ZHONG, J.; GOODWIN, L.; COPELAND, A.; LUCAS, S.; HAN, C.; PITLUCK, S.; LAND, M.L.; KYRPIDES, N.C.; WOYKE, T. Complete genome sequence of *Anabaena variabilis* ATCC 29413. **Standards in Genomic Sciences**, v. 9, n. 3, p. 562–573, 2014.

THOMPSON, J.D.; HIGGINS, D.G.; GIBSON, T.J. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weigh matrix choice. **Nucleic Acids Research**, London, v. 22, p. 4673-4680, 1994.

TRIPP, H.J; BENCH, S.R; TURK, K.A. FOSTER, R.A.; DESANY, B.A.; NIAZI, F.; AFFOURTIT, J.P.; ZEHR, J.P. Metabolic streamlining in a open-ocean nitrogen-fixing cyanobacterium. **Nature**, v. 464, p. 90-94, 2010.

TSYGANKOV, A.A.; FEDOROV, A.S.; KOSOUROV, S.N.; RAO, K.K. Hydrogen production by cyanobacteria in an automated outdoor photobioreactor under aerobic conditions. **Biotechnology and Bioengineering**, New York, v. 80, p. 777–783, 2002.

- VALLADARES, A.; FLORES, E.; HERRERO, A. Transcription activation by NtcA and 2-oxoglutarate of three genes involved in heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. **Journal of Bacteriology**, Washington, v. 190, n.18, p. 6126– 6133, 2008.
- VOß, B.; BOLHUIS, H.; FEWER, D.P. KOPF, M.; MÖKE, F.; HAAS, F.; EL-SHEHAWY, R.; HAYES, P.; BERGMAN, B.; SIVONEN, K.; DITTMANN, E.; SCANLAN, D.J.; HAGEMANN, M.; STAL, L.J.; HESS, W.R. Insights into the physiology and ecology of the brackish-water-adapted Cyanobacterium *Nodularia spumigena* CCY9414 based on a genome-transcriptome analysis. **PLoS ONE**, San Francisco, v. 8, n. 3, e60224, 2013.
- WALSBY, A.E. Cyanobacterial heterocysts: terminal pores proposed as sites of gas exchange. **Trends in Microbiology**, Cambridge, v. 15, p. 340-349, 2007.
- WALSBY, A.E. The permeability of heterocyst to the gases nitrogen and oxygen. **Proceedings of the Royal Society of London**, London, v. 226, p. 345-367, 1985.
- WANG, H.; SIVONEN, K.; ROUHIAINEN, L.; FEWER, D.P.; LYRA, C.; RANTALAYLINEN, A.; VESTOLA, J.; JOKELA, J.; RANTASÄRKÄ, K.; LI, Z.; LIU, B. Genome derived insights into the biology of the hepatotoxic bloom-forming cyanobacterium *Anabaena* sp. strain 90. **BMC Genomics**, London, v. 13, p. 613, 2012.
- WAUGH, S.I.; PAULSEN, D.M.; MYLONA, P.V.; MAYNARD, R.H.; PREMAKUMAR, R.; BISHOP, P.E. The genes encoding the delta subunits of dinitrogenases 2 and 3 are required for Mo-independent diazotrophic growth by *Azotobacter vinelandii*. **Journal of Bacteriology**, Washington, v. 177, p. 1505–1510, 1995.
- WEBER, T.; BLIN, K.; DUDELA, S.; KRUG, D.; KIM, H.U.; BRUCCOLERI, R.; LEE, S.Y.; FISCHBACH, M.A.; MULLER, R.; WOHLLEBEN, W.; BREITLING, R.; TAKANO, E.; MEDEMA, M.H. antiSMASH 3.0: a comprehensive resource for the genome mining of biosynthetic gene clusters. **Nucleic Acids Research**, London, v. 43, p. W237-243, 2015.
- WERNER, V.R.; LAUGHINGHOUSE IV, H.D.; SANT'ANNA, C.L.; FIORE, M.F.; HOFF, C.; SANTOS, K.R.S.; NEUHAUS, E.B.; MOLICA, R.J.R.; HONDA, R.Y.; ECHENIQUE, R.O. Morphological and molecular studies of *Sphaerospermopsis torques-reginae* (Cyanobacteria, Nostocales) from South American water blooms. **Phycologia**, Odense, v.51, p. 228-238, 2012.
- WHITTON, B.A. 2000 Soils and Rice-Fields In: WHITTON, B.A.; POTTS, M. **The ecology of Cyanobacteria**. Netherlands: Kluwer Academic, p. 523-561, 2000.
- WOOD, D.; SALZBERG, S.L. Kraken: ultrafast metagenomic sequence classification using exact alignments. **Genome Biology**, London, v. 15, n. 2, p. 1-12, 2014.
- WRAGE, S.F.; VELTHOF, G.L.; van BEUSICHEM, M.L.; OENEMA, O. Role of nitrifier denitrification in the production of nitrous oxide. **Soil Biology and Biochemistry**, Oxford, v. 33, p. 1723-1732, 2001.
- XIONG, J.; FISCHER, W. M.; INOUE, K.; NAKAHARA, M.; BAUER, C. E. Molecular evidence for the early evolution of photosynthesis. **Science**, Washington, v. 289, p. 1724–1730, 2000.

XU, X.; ELHAI, J.; WOLK, C.P. Transcriptional and developmental responses by *Anabaena* to deprivation of fixed nitrogen. In: HERRERO, A.; FLORES, E. **The cyanobacteria. Molecular biology, genomics and evolution**. Norfolk UK: Castier Academic Press, 2008. p. 383-422.

YOON, H.S.; GOLDEN, J. W. PatS and products of nitrogen fixation control heterocyst pattern. **Journal of Bacteriology**, Washington, v. 183, p. 2605–2613, 2001.

YOUNG, J. P. W. Molecular evolution in diazotrophs: do the genes agree? In: PEDROSA, F. O.; HUNGRIA, M.; YATES, G.; NEWTON, W. E. (Ed.). **Nitrogen Fixation: from molecules to crop productivity**. Dordrecht: Kluwer Academic Publishers, 2000. p. 161-164.

YOUNG, J.P.W. Phylogenetic classification of nitrogen-fixing organisms. In: STACEY, G.; EVANS, H.J.; BURRIS, R.H. (Ed.). **Biological nitrogen fixation**. London: Chapman & Hall, 1992. p. 43-86.

ZEHR, J. P. Nitrogen fixation by marine cyanobacteria. **Trends in Microbiology**, Cambridge, v. 19, No. 4, 2011.

ZEHR, J. P.; CAPONE, D. G. Problems and promises of assaying the genetic potential for nitrogen fixation in the marine environment. **Microbial Ecology**, New York, v. 32, p. 263-281, 1996.

ZEHR, J. P.; JENKINS, B. D.; SHORT, S. M.; STEWARD, G. F. Nitrogenase gene diversity and microbial community structure: A cross-system comparison. **Environmental Microbiology**, London, v. 5, p. 539-554, 2003.

ZHANG, C.C.; LAURENT, S.; SARK, S.; PENG, L.; BÉDU, S. Heterocyst differentiation and pattern formation in cyanobacteria: a chorus of signals. **Molecular Microbiology**, Oxford, v. 59, n. 2, p. 367-375, 2006.

ZHANG, J.; KOBERT, K.; FLOURI, T.; STAMATAKIS, A. PEAR: a fast and accurate Illumina paired-end read merger. **Bioinformatics**, Oxford, v. 30, p. 614-620, 2014.

ZHANG, W.; DU, Y.; KHUDYAKOV, I.; FAN, Q.; GAO, H.; NING, D.; WOLK, C.P.; XU, X. et al. A cluster that regulates both heterocyst differentiation and pattern formation in *Anabaena* sp. strain PCC 7120. **Molecular Microbiology**, Oxford, v. 66, p. 1429–1443, 2007.

ZUMFT, W.G. Cell biology and molecular basis of denitrification. **Microbiology and Molecular Biology Reviews**, Washington, v. 61, p. 533-616, 1997.

APÊNDICES

APÊNDICE A

Árvore gerada por inferência bayesiana do gene *nifT*

Árvore gerada por inferência bayesiana do gene *nifZ*

Árvore gerada por inferência bayesiana do gene *nifV*

Árvore gerada por inferência bayesiana do gene *nifB*

Árvore gerada por inferência bayesiana do gene *fdxN*

Árvore gerada por inferência bayesiana do gene *nifS*

Árvore gerada por inferência bayesiana do gene *nifU*

Árvore gerada por inferência bayesiana do gene *nifD*

Árvore gerada por inferência bayesiana do gene da enzima similar a Mo-nitrogenase

Árvore gerada por inferência bayesiana do gene *nifX*

Árvore gerada por inferência bayesiana do gene *duf269*

Árvore gerada por inferência bayesiana do gene *duf683*

Árvore gerada por inferência bayesiana do gene *nifW*

Árvore gerada por inferência bayesiana do gene *hesA*

Árvore gerada por inferência bayesiana do gene *hesB*

Árvore gerada por inferência bayesiana do gene *fdxH*

Árvore gerada por inferência bayesiana do gene *feoA*

ANEXOS

ANEXO A – Lista de espécies utilizadas para construção da árvore concatenada dos genes *nifHDK* e homólogos, com número de acesso dos genes (continua)

LINHAGENS	Número de acesso GenBank		
	<i>nifH</i>	<i>nifD</i>	<i>nifK</i>
<i>Acidithiobacillus ferrooxidans</i> ATCC 53993	WP_00956749 4.1	WP_01253658 0.1	WP_01253657 9.1
<i>Alkaliphilus metalliredigens</i> QYMF	WP_01206461 4.1	WP_01206461 1.1	WP_01206461 0.1
<i>Anabaena variabilis</i> ATCC 29413 vnf	WP_01132061 0.1	WP_01132071 5.1	WP_01132071 6.1
<i>Anabaena variabilis</i> ATCC 29413 nif	WP_01132092 7.1	WP_01132092 8.1	WP_01132092 9.1
<i>Azotobacter vinelandii</i> DJ nif	WP_01269883 1.1	WP_01269883 2.1	WP_01269883 3.1
<i>Azotobacter vinelandii</i> DJ anf	WP_01270336 2.1	WP_01270336 1.1	WP_01270335 9.1
<i>Azotobacter vinelandii</i> DJ vnf	WP_01269895 5.1	WP_01269895 0.1	WP_01269894 8.1
<i>Beijerinckia indica</i> subsp. <i>indica</i> ATCC 9039	WP_01238348 4.1	WP_01238348 5.1	WP_01238348 6.1
<i>Caldicellulosiruptor saccharolyticus</i> DSM 8903	WP_01191796 4.1	WP_01191796 1.1	WP_01191796 0.1
Candidatus <i>Azobacteroides pseudotrichonymphae</i> CFP2	WP_01257356 9.1	WP_01257357 2.1	WP_01257357 3.1
Candidatus <i>Desulforudis audaxviator</i> MP104C	WP_01230130 5.1	WP_01230130 8.1	WP_01230130 9.1
Candidatus <i>Methanosphaerula palustris</i> E1-9c	WP_01261725 3.1	WP_01261725 0.1	WP_01261724 9.1
<i>Chlorobium tepidum</i> TLS000	NP_662417.1	NP_662420.1	NP_662421.1
<i>Chloroherpeton thalassium</i> ATCC 35110	WP_01250036 9.1	WP_01250037 2.1	WP_01250037 4.1
<i>Chloroherpeton thalassium</i> ATCC 35110 nif	WP_01249958 3.1	WP_01249958 0.1	WP_01249957 9.1
<i>Clostridium acetobutylicum</i> ATCC 824	NP_346894.1	NP_346897.1	NP_346898.1
<i>Clostridium beijerinckii</i> NCIMB 8052	ABR34169.1	ABR34172.1	ABR34173.1
<i>Clostridium kluyveri</i> NBRC 12016	WP_01210214 0.1	WP_01210213 9.1	WP_01210213 7.1
<i>Dehalococcoides mccartyi</i> 195	WP_01093685 0.1	WP_01093684 7.1	WP_01093684 6.1
<i>Desulfitobacterium hafniense</i> Y51	WP_00581352 9.1	WP_00581353 1.1	WP_01146172 0.1
<i>Desulfotomaculum reducens</i> MI-1	WP_01187912 0.1	WP_01187911 7.1	WP_01187911 6.1
<i>Desulfovibrio vulgaris</i> DP4	ABM30101.1	WP_01178730 9.1	ABM30105.1
<i>Geobacter lovleyi</i> SZ	WP_01246873 2.1	WP_01246873 1.1	WP_01246873 0.1
<i>Geobacter</i> sp. M21	WP_01583743 6.1	WP_01583743 5.1	WP_01583743 4.1
<i>Hellobacterium modesticaldum</i> Ice1	WP_01228221 8.1	WP_01228221 9.1	WP_01228222 0.1
<i>Klebsiella pneumoniae</i> 342	WP_00420355 3.1	WP_01254111 2.1	WP_00880413 3.1

ANEXO A – Lista de espécies utilizadas para construção da árvore concatenada dos genes *nifHDK* e homólogos, com número de acesso dos genes (continuação)

LINHAGENS	Número de acesso GenBank		
	<i>nifH</i>	<i>nifD</i>	<i>nifK</i>
<i>Magnetospirillum magnetotacticum</i> MS-1	WP_01138398 4.1	ZP_00054386. 1	ZP_00054385. 1
<i>Methanobacterium thermoautotrophicum</i> strain DeltaH	WP_01087716 9.1	WP_01087717 2.1	WP_01087717 3.1
<i>Methanococcus aeolicus</i> Nankai-3	WP_01197414 2.1	WP_01197413 9.1	WP_01197413 8.1
<i>Methanococcus maripaludis</i> strain S2	WP_01117079 7.1	WP_01117080 0.1	WP_01117080 1.1
<i>Methanococcus vannielii</i> SB	WP_01197188 3.1	WP_01197188 0.1	WP_01197187 9.1
<i>Methanosarcina acetivorans</i> str.C2A <i>nif</i>	WP_01102379 1.1	WP_01102379 4.1	WP_01102379 5.1
<i>Methanosarcina acetivorans</i> str.C2A <i>vnf</i>	WP_01102122 7.1	WP_01102123 7.1	WP_01102123 9.1
<i>Methanosarcina acetivorans</i> str.C2A <i>anf</i>	WP_01102122 7.1	WP_01102123 2.1	WP_01102123 0.1
<i>Methylococcus thermophilus</i> V4	WP_01246421 2.1	ACD83928.1	ACD83927.1
<i>Methylocella silvestris</i> BL2	WP_01259258 5.1	WP_01259258 4.1	WP_01259258 3.1
<i>Nostoc</i> sp. PCC 7120	WP_01099562 6.1	WP_01099561 3.1	WP_01099561 2.1
<i>Diplosphaera colitermitum</i> TAV2	ZP_03723745. 1	ZP_03723746. 1	ZP_03723747. 1
<i>Pelobacter carbinolicus</i> DSM 2380	YP_357508.1	YP_357509.1	YP_357510.1
<i>Rhodobacter sphaeroides</i> ATCC 17025	ABP70147.1	ABP70148.1	ABP70149.1
<i>Rhodopseudomonas palustris</i> BisA53	WP_01166527 4.1	WP_01166527 3.1	WP_01166527 1.1
<i>Rhodopseudomonas palustris</i> HaA2	WP_01143986 8.1	WP_01143986 9.1	WP_01143987 0.1
<i>Rhodospirillum centenum</i> SW	WP_01256880 5.1	WP_01143986 9.1	WP_01256880 3.1
<i>Rhodospirillum rubrum</i> ATCC 11170	YP_426483.1	YP_426482.1	YP_426480.1
<i>Roseiflexus castenholzii</i> DSM 13941	WP_01212249 7.1	WP_01212249 5.1	WP_01212249 4.1
<i>Roseiflexus</i> sp. RS-1	WP_01195596 1.1	WP_01195595 9.1	WP_01195595 8.1
<i>Sinorhizobium medicae</i> WSM 419	YP_001314762 .1	YP_001314761 .1	YP_001314760 .1
<i>Synechococcus</i> sp. JA-2-3B'a	WP_01143065 1.1	WP_01143208 2.1	WP_01143208 3.1
<i>Syntrophobacter fumaroxidans</i> MPOB	WP_01169788 4.1	WP_01169788 1.1	WP_01169788 0.1
<i>Thermodesulfobacterium yellowstonii</i> DSM 11347	YP_002249508 .1	YP_002249507 .1	YP_002249506 .1
<i>Zymomonas mobilis</i> subsp. <i>mobilis</i> ATCC 10988	AEH63154.1	WP_01124155 7.1	AEH63152.1
<i>Methanocaldococcus infernus</i> ME	WP_01309945 9.1	WP_01309945 6.1	WP_01309945 5.1
<i>Syntrophothermus lipocalidus</i> DSM 12680	WP_01317627 5.1	WP_01317627 2.1	WP_01317627 1.1

ANEXO A – Lista de espécies utilizadas para construção da árvore concatenada dos genes nifHDK e homólogos, com número de acesso dos genes (conclusão)

LINHAGENS	Número de acesso GenBank		
	<i>nifH</i>	<i>nifD</i>	<i>nifK</i>
<i>Methanocaldococcus vulcanius</i> M7	WP_015733159.1	WP_015733162.1	WP_015733163.1
<i>Methanosarcina mazei</i> strain Goe1	WP_011032670.1	WP_011032673.1	WP_011032674.1
<i>Methanocaldococcus</i> sp. FS406-22	WP_012979647.1	WP_012979650.1	WP_012979651.1
<i>Methanothermococcus okinawensis</i> IH1	WP_013867273.1	WP_013867270.1	WP_013867269.1
<i>Oscillochloris trichoides</i> DG-6	WP_006560621.1	EFO82064.1	EFO82063.1
	ChIL/BchL	ChIN/BchN	ChIB/BchB
<i>Bradyrhizobium</i> sp. BTAi1	WP_012046265.1	WP_012046268.1	WP_012046267.1
<i>Methylobacterium populi</i> BJ001	(WP_012457046.1)	(WP_012457049.1)	(WP_012457048.1)
<i>Halorhodospira halophila</i> SL1	(WP_011814420.1)	(WP_011814423.1)	(WP_011814422.1)
<i>Prochlorococcus</i> sp. CC9311	(WP_011619890.1)	(WP_011619888.1)	(WP_011619889.1)
<i>Prochlorococcus marinus</i> MIT 9515	(WP_011819923.1)	(WP_011819925.1)	(WP_011819924.1)
<i>Chloroflexus aggregans</i> DSM 9485	(WP_015942165.1)	(WP_015942168.1)	(WP_015942167.1)
<i>Roseiflexus castenholzii</i> HLO8, DSM 13941	(WP_012120059.1)	(WP_012120057.1)	(WP_012120058.1)
<i>Chlorobium phaeobacteroides</i> DSM 266	(WP_011746157.1)	(WP_011746159.1)	(WP_011746158.1)
<i>Gloeobacter violaceus</i> PCC 7421	(NP_925316.1)	(NP_925315.1)	(NP_923161)
<i>Heliobacterium modesticaldum</i> Ice1	(WP_012282383.1)	(WP_012282384.1)	WP_012282385.1
<i>Chlorobium limicola</i> DSM 245	(WP_012467084.1)	(WP_012467086.1)	(WP_012467085.1)