UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
Departamento de Botânica
Programa de Pós-Graduação em Biologia Vegetal

UFMG

LAÍSA CORRÊA BRAGA MARQUES

TAXONOMY AND NITROGEN FIXATION IN *RAPHIDIOPSIS RACIBORSKII* (WOLOSZYNSKA) AGUILERA, BERRENDERO GÓMEZ, KASTOVSKY, ECHENIQUE & SALERNO

Dissertação apresentada ao Programa de Pós-Graduação em Biologia Vegetal do Departamento de Botânica do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Biologia Vegetal.

Área de Concentração Ecologia Vegetal

BELO HORIZONTE – MG

2020

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
Departamento de Botânica
Programa de Pós-Graduação em Biologia Vegetal

UFMG

LAÍSA CORRÊA BRAGA MARQUES

TAXONOMY AND NITROGEN FIXATION IN *RAPHIDIOPSIS RACIBORSKII* (WOLOSZYNSKA) AGUILERA, BERRENDERO GÓMEZ, KASTOVSKY, ECHENIQUE & SALERNO

Dissertação apresentada ao Programa de Pós-Graduação em Biologia Vegetal do Departamento de Botânica do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Biologia Vegetal.

Área de Concentração Ecologia Vegetal

Orientadora: Profa. Dra. Alessandra Giani
Universidade Federal de Minas Gerais

Co-orientadora: Profa. Dra. Juliana Pimentel
Faculdade Pitágoras

BELO HORIZONTE – MG

2020

043 Marques, Laísa Corrêa Braga.
Taxonomy and nitrogen fixation in *Raphidiopsis raciborskii* (Woloszynska) Aguilera, Berrendero Gómez, Kastovsky, Echenique & Salerno [manuscrito] / Laísa Corrêa Braga Marques. – 2020.

117 f. : il. ; 29,5 cm.

Orientadora: Profa. Dra. Alessandra Giani. Coorientadora: Profa. Dra. Juliana Pimentel.
Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Biologia Vegetal.

1. Cianobactéria. 2. Nitrogênio. 3. Fixação de Nitrogênio. 4. Classificação. 5. Nitrogenase. I. Giani, Alessandra. II. Pimentel, Juliana. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título.

CDU: 581

Ficha catalográfica elaborada pela bibliotecária Rosilene Moreira Coelho de Sá – CRB 6 – 2726

Universidade Federal de Minas Gerais
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Biologia Vegetal

UFMG

ATA DE DEFESA DE DISSERTAÇÃO - MESTRADO
(Defesa MS nº 184)
LAÍSA CORRÊA BRAGA MARQUES

Data: 17/12/2020
Entrada: 2018/2

Matrícula: 2018720168
CPF: 13034128614

Às 9 horas do dia 17 de dezembro de 2020, reuniu-se a Comissão Examinadora de Dissertação indicada pelo Colegiado do Programa para julgar, em exame final, o trabalho intitulado **"Taxonomy and nitrogen fixation in *Raphidiopsis raciborskii* (Woloszynska) Aguilera, Berrendero Gómez, Kastovsky, Echenique & Salerno"**, requisito final para obtenção do grau de Mestre em Biologia Vegetal, área de concentração em Fisiologia Vegetal e Ecologia, pela discente **Laísa Corrêa Braga Marques** Abrindo a sessão, a Presidente da Comissão, professora Dra. Alessandra Giani, após dar conhecimento aos presentes do teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores, com a respectiva defesa da candidata. Logo após, a comissão reuniu-se, sem a presença da candidata e do público, para julgamento e expedição do resultado final. Foram atribuídas as seguintes indicações:

Membro da Comissão Examinadora	Instituição	Resultado
Dra. Alessandra Giani (orientadora)	Universidade Federal de Minas Gerais	Aprovada
Dra. Maria Kolman	Instituto de Biotecnologia de Misiones Argentina	Aprovada
Dr. Álvaro Cantini Nunes	Universidade Federal de Minas Gerais	Aprovada

Pelas indicações, a candidata foi considerada: __APROVADA____.

O resultado final foi comunicado publicamente à candidata pela Presidente da Comissão. Nada mais havendo a tratar, a Presidente encerrou a reunião e lavrou-se presente ata, que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 17 de dezembro de 2020.

Dra. Alessandra Giani (orientadora) Alessandra Giani

Dra. Maria Kolman

Dr. Álvaro Cantini Nunes Álvaro Cantini Nunes

Prof. Alexandre Salino
PPG Biologia Vegetal - UFMG
Coordenador

Agradecimentos

Gostaria de agradecer a professora Alessandra Giani por sua companhia, seu encorajamento e orientação.

À profa. Juliana Pimentel pela dedicação em me orientar e ensinar.

Agradeço pela ajuda dos meus colegas de laboratório Bruna, Estêvão, Juliana Dutra, Natalia que foram uns doces e criaram um ambiente feliz no qual eu amei trabalhar

Agradeço à minha querida amiga Hortênsia que sempre esteve ao meu lado e ajudou em todas as etapas deste trabalho

À minha amada família, em especial, mãe, pai, vovó Maristela, Jéssica, Lorrana, tia Marlian e vovó Laura.

À Elenice por sua estimada amizade

Aos técnicos Paulão e Wagner pela ajuda

Aos profs. Álvaro Cantini e Maria Kolman por aceitarem prontamente fazer parte da banca examinadora, por sua disponibilidade e contribuíções

Aos meus colegas da pós-graduação pela amizade, principalmente Humberto pela ajuda.

Aos professores (as), especialmente ao prof. Joni Lima que, neste trabalho, me ajudou e ensinou e à profa. Queila por me emprestar seu equipamento

Á Fiocruz pelo sequenciamento

Á Pós-graduação de Biologia Vegetal, ao ICB e à UFMG pela oportunidade e ensino.

A todas as pessoas que direta ou indiretamente contribuíram para a realização deste trabalho

E a CNPq, cujo auxílio financeiro proporcionou toda a pesquisa feita neste trabalho.

Sumário

RESUMO GERAL

Cianobactérias são organismos fotosssintetizantes que possibilitaram a existência de oxigênio no ambiente, e com isso a vida de outros seres no planeta. Esses organismos têm um papel essencial no passado evolutivo e nos ecossistemas modernos por estarem na base dos ciclos biogeoquímicos globais de carbono e nitrogênio.

Neste trabalho estudamos a capacidade da espécie *Raphidiopsis raciborskii* de fixar nitrogênio atmosférico. Este organismo tem se adaptado aos ambientes aquáticos de várias partes do mundo e causado algumas preocupações devido ao seu potencial de produzir toxinas e porque é uma espécie invasora. Apesar de seu sucesso ecológico, a diazotrofia, isto é a capacidade de fixar nitrogênio atmosférico, ainda não é tão estudada nessa espécie. Seus aspectos relacionados à produção de toxinas e ao potencial de propagação foram foco de muitos estudos, mas ainda precisamos saber mais sobre sua ecologia, posição taxonômica e fisiologia.

Para contribuir com mais informações, neste trabalho, tivemos três principais objetivos. O primeiro foi esclarecer as relações taxonômicas e filogenéticas dentro de cepas com e sem heterocitos (células especializadas com função de fixação de nitrogênio) de *Raphidiopsis raciborskii*. O segundo foi avaliar sua resposta à restrição de nitrogênio, por meio de experimentos de expressão gênica. E o terceiro foi revisar brevemente a rota evolutiva dos genes que se traduzem na enzima nitrogenase, não apenas na espécie *R.raciborskii*, mas nas bactérias diazotróficas em geral.

Para alcançarmos o primeiro objetivo foram feitas análises filogenéticas de quatro regiões do genoma de *R.raciborskii*. A partir dos resultados, concluímos que os gêneros *Raphidiopsis* e *Cylindrospermopsis* são sinônimos e cepas produtoras ou não produtoras de heterócitos co-ocorrem normalmente, essa característica não sendo suficiente para distinguir entre gêneros.

O segundo objetivo foi abordado utilizando experimento de restrição de nitrato em que

se compararam cepas cultivadas em meio com depleção total de nitrogênio e em meio rico em nitrogênio dissolvido. A produção de heterócitos e a expressão gênica de *nifH*, *hetR* e *ntcA* (genes relevantes para que a fixação de nitrogênio ocorra) foram avaliados. Foi possível concluir que a produção de heterócitos não é totalmente inibida pela presença de nitrogênio dissolvido no meio e *R.raciborskii* pode expressar a nitrogenase mesmo em um meio rico em nitrato, ao contrário do que ocorre com outras cianobactérias diazotróficas.

Para o terceiro objetivo foi feito o levantamento bibliográfico juntamente com análise de relações filogenéticas entre as proteínas da família Nif, responsáveis pela fixação de nitrogênio. Foi possível ver que houve várias duplicações genicas para poder formar a família *nif* tal qual conhecemos hoje. Uma duplicação notável é a do óperon *nifDK* que gerou os genes *nifE* e *nifN*. O surgimento da enzima nitrogenase ainda é incerto e existem sugestões que o primeiro organismo capaz de sintetizá-la seja uma Archea, ou até o Último Ancestral Comum de toda a vida.

Palavras-chave: Cianobactéria, Nitrogênio, Fixação de nitrogênio, Heterócito, Taxonomia, nif, nitrogenase

ABSTRACT

Cyanobacteria are photosynthetic organisms that made possible the existence of oxygen rich environments and the development of other living organisms in the planet. These organisms play an essential role in the evolutionary past and in modern ecosystems because they are at the basis of the global biogeochemical cycles of carbon and nitrogen.

In this work we studied the ability of *Raphidiopsis raciborskii* to fix atmospheric nitrogen. This organism has adapted to aquatic environments in various parts of the world and causes concerns due to its potential for toxin production and its invasiveness. Despite its ecological success, diazotrophy, i.e. the capacity to fix N_2, is still not well studied in this species. The aspects related to the production of toxins and the potential invasiveness were the focus of many studies, but we still need to know more about its ecology, taxonomy and physiology.

To increase this knowledge, this work had three main objectives. The first was to clarify the phylogenetic and taxonomic relationships within strains with and without heterocytes (nitrogen fixation specialized cells) of *Raphidiopsis raciborskii*. The second was to assess their response to nitrogen restriction, through gene expression experiments. And, the third was to briefly review the evolutionary route of the genes that translate into the nitrogenase enzyme, not only in the species *R.raciborskii*, but in diazotrophic bacteria in general.

To achieve the first objective, phylogenetic analyses of four regions of the *R.raciborskii* genome were performed. From the results, we concluded that the genera *Raphidiopsis* and *Cylindrospermopsis* are synonymous, that the production of heterocytes is not enough to distinguish between genera and heterocyte producing or non-producing strains co-occur normally in the environment.

The second objective was addressed using a nitrate restriction experiment in which strains grown in medium with total nitrogen depletion and in medium rich in dissolved

nitrogen were compared. In this experiment the production of heterocytes and the gene expression of *nifH, hetR* and *ntcA* (N_2 fixation relevant genes) were evaluated. It was possible to conclude that the production of heterocytes is not completely inhibited by the presence of dissolved nitrogen and *R.raciborskii* can express nitrogenase in a medium rich in nitrate, unlike what happens with other diazotrophic cyanobacteria.

For the third objective, a bibliographic survey was carried out together with an analysis of phylogenetic relationships between the Nif family proteins, responsible for nitrogen fixation. It was possible to note that several genetic duplications occur in order to form the *nif* family as seen today. A notable duplication is that of the *nifDK* operon that generated the *nifE* and *nifN* genes. A lot of uncertainty results from the origin of the nitrogenase enzyme. Some suggest that the first organism capable of its synthesis was an Archea, others suggest that it was the Last Common Ancestor of all living organisms.

Key-words: Cyanobacteria, Nitrogen, Nitrogen-fixation, Heterocyst, Taxonomy, nif, nitrogenase

1. INTRODUÇÃO GERAL

1.1. CIANOBACTÉRIAS: O QUE SÃO?

As cianobactérias são organismos procariontes que pertencem ao domínio Bacteria e foram os primeiros a desenvolver a capacidade de realizar fotossíntese com liberação de oxigênio. Isso levou ao acúmulo de oxigênio na atmosfera e desempenhou um papel crucial na evolução do nosso planeta (Kopp *et al*., 2005). Esses microorganismos possuem inúmeras camadas de membranas dentro da célula, paralelas umas às outras, chamadas tilacóide, onde ocorrem as reações fotossintéticas.

As cianobactérias têm uma ampla distribuição ambiental e ocupam uma variedade de habitats, incluindo vastas áreas oceânicas, solos, lagos, e até mesmo habitats extremos, como desertos áridos ou fontes termais.

A parede celular das cianobactérias é Gram-negativa e compreende duas membranas, uma basal e outra externa, e há entre elas uma camada de peptidoglicano mais espessa do que em outras bactérias Gram-negativas (Hoiczyk & Hansel, 2000).

A parte mais externa do revestimento de células cianobacterianas é o glicocálice. Essa camada tem várias funções: tende a atuar como barreira protetora, armadilha molecular / iônica e filtro, e está envolvida na adesão celular e no reconhecimento de superfície (Smarda & Komrska, 1993). Essa camada protege as células da desidratação e provavelmente de fagos e predadores.

1.2. FLORAÇÕES

As chamadas "florações" (*blooms*) são formadas pelo crescimento massivo e descontrolado de microrganismos fotossintéticos em ecossistemas aquáticos, e esse crescimento é desencadeado por mudanças ambientais. O *bloom* de cianobactérias é um dos fenômenos mais comumente associados à degradação dos ecossistemas aquáticos, causado principalmente por compostos fosfatados e nitrogenados.

O desenvolvimento rápido da floração pode ocorrer quando as taxas de crescimento e

divisão celular são altas ou quando a estratificação da coluna d'água promove o acúmulo de espécies flutuantes na superfície da água (Paerl *et al*., 2001). Florações com uma grande densidade de células podem ocasionar o aumento da turbidez, (diminuição da transparência) e a mudança da aparência da água que se torna lamacenta, odorífera. Frequentemente, também há a deposição de células senescentes e formação de espuma na região litorânea e nas margens de lagos (Paerl *et al.*, 2001). Esse ambiente degradado hospeda uma variedade de bactérias decompositoras que utilizam oxigênio para o processo de decomposição, isso acaba tirando oxigênio das camadas do fundo, causando hipóxia, anóxia, liberação de sulfeto de hidrogênio tóxico e liberação acelerada de nutrientes dos sedimentos, que agrava ainda mais a eutrofização e as florações. As águas impactadas desta forma tem, em geral, seu uso restrito ou proibido para recreação (banho, pesca). As florações interrompem o equilíbrio ecológico dos ecossistemas aquáticos, mas também podem ser prejudiciais aos seres vivos. Um *bloom* recebe o título de 'prejudicial' quando espécies formadoras de floração produzem compostos tóxicos para a biota (Paerl & Otten, 2013). Essas toxinas são extremamente potentes e atingem organismos muito além daqueles presentes nas comunidades aquáticas.

1.3. NITROGÊNIO

As cianobactérias estão na base dos ciclos biogeoquímicos globais de carbono e nitrogênio. Isso proporcionou a elas papéis essenciais no passado evolutivo e nos ecossistemas modernos (Kopp *et al.*, 2005).

As fontes de nitrogênio mais comumente usadas por esses organismos são nitrato, amônio, uréia e nitrogênio atmosférico (N_2) (Flores & Herrero, 1999). Muitas cianobactérias fixam N_2, e um grande número delas o faz em condições aeróbicas. Já que a nitrogenase, complexo enzimático que executa a fixação de nitrogênio, é reativa ao oxigênio, muitas espécies dividem os processos de fotossíntese oxigenada e de fixação de nitrogênio, tanto

espacial quanto temporariamente. Algumas cianobactérias filamentosas restringem a nitrogenase a heterócitos, células diferenciadas fixadoras de nitrogênio. Essas células estabelecem condições micro-óxicas para a atividade ótima da nitrogenase, tais como a perda do fotossistema II e operações de fixação de carbono, e a deposição de camadas extras de polissacarídeo e glicolipídeo (Flores & Herrero, 2010). Como essas células não podem fazer fotossíntese, elas dependem do carbono fornecido pelas células vegetativas vizinhas (Wolk, 1968). Já outras cianobactérias fixadoras de nitrogênio, unicelulares e filamentosas, expressam a atividade da nitrogenase nas fases escuras do ciclo claro-escuro de crescimento (Mullineaux *et al*., 1991).

1.4. TAXONOMIA

A taxonomia das cianobactérias pode ser um grande desafio. Wayne (1987) propôs a classificação que atualmente é utilizada para as espécies bacterianas. Ele postulou que cepas bacterianas contendo 70% ou mais de parentesco DNA-DNA e compartilhando traços fenotípicos característicos devem ser consideradas da mesma espécie. Hoje, com o advento da genômica, muitas cepas são sequenciadas e com isso podemos juntar todos os genomas em um só, com a combinação de todos os genes presentes em todas as cepas de uma espécie. Esse genoma formado a partir de um conjunto de genomas é chamado de genoma *core*. Quando uma cepa apresenta 70% ou mais de parentesco com esse genoma esta cepa é considerada da mesma espécie. No entanto, no domínio Cyanobacteria o conceito de espécie não é tão uniforme. Recentemente, a taxonomia deste grupo vem sendo questionada por causa de um conjunto crescente de dados genômicos que permitem, em princípio, classificar as espécies bacterianas com mais precisão.

O uso de dados moleculares na taxonomia de cianobactérias é relativamente recente e vários grupos ainda carecem dessa informação, alguns grupos são descritos apenas por traços

morfológicos e ecológicos. Tais traços, usados para classificar a níveis supragenéricos e de gênero, talvez não tenham valor taxonômico. Komárek (2010) afirmou que inúmeros caracteres tradicionais não são compatíveis com resultados moleculares, tais como presença e ausência de bainha de mucilagem, tamanho de célula e ramificação falsa e verdadeira.

A interpretação da história evolutiva de certos grupos de cianobactéria pode ter erros. Por isso, estudos de filogenia molecular e caracters morfológicos que coincidam com sua posição filogenética são tão importantes.

Outro problema dentro da taxonomia é a interpretação errônea de dados. Ao analisar a variabilidade molecular, é importante considerar cuidadosamente as alternativas antes de propor razões *ad hoc* para padrões irregulares, porque isso leva a grupos parafiléticos e polifiléticos, o que seria um erro na filogenia. Parafilia, de acordo com Hennig (1966), refere a grupos que têm um ancestral comum, mas não incluem todos os seus descendentes. Já polifilia é o resultado de uma série de ancestrais distintos e não está especificamente relacionada à homologia (Farris, 1974). De acordo com Funk & Omland (2003) uma das causas da polifilia é a informação filogenética inadequada, que pode resultar em árvores de genes imprecisas, e essas árvores se tornam um artefato quando é feita a reconstrução filogenética. Algorítimos filogenéticos podem construir topologias independentemente do volume e da qualidade da informação. Portanto, se um gene evolui muito lentamente em relação à taxa de especiação dos táxons, ou se a região analisada é muito pequena, os dados obtidos teriam apenas algumas sinapomorfias para dar suporte à árvore filogenética. É por isso que, ao fazer árvores filogenéticas, é importante usar mais de uma região genômica.

1.5. DESCRIÇÃO DA ESPÉCIE ALVO

A cianobactéria *Raphidiopsis raciborskii* (Woloszynska) Aguilera, Berrendero Gómez, Kastovsky, Echenique & Salerno (Aguilera *et al*., 2018), até recentemente descrita

como *Cylindrospermopsis raciborskii* (Wolozyńska) Seenayya & Subba, é uma espécie invasiva, potencialmente tóxica, de origem subtropical-tropical, capaz de fixar N_2, produzir heterócitos e formar florações. Tem sido encontrada se expandindo para rios e corpos d'água em regiões temperadas (Padisák, 1997; Briand *et al*., 2004; Stüken *et al*., 2006) favorecida por um conjunto de características, como adaptação a temperaturas mais baixas e aos efeitos do aquecimento global (Briand *et al*., 2004; Sinha *et al*., 2012).

Devido ao seu predomínio frequente em florações de água doce e sua capacidade de sintetizar cianotoxinas, *R. raciborskii* tem recebido atenção nas últimas décadas. É relatado que essa cianobactéria produz cylindrospermopsina (CYN) e saxitoxina (SXT), porém em cepas diferentes (Hawkins *et al*., 1985; Lagos *et al*., 1999). SXT é uma neurotoxina alcalóide que bloqueia os canais neuronais de sódio, potássio e cálcio, afetando a propagação dos impulsos nervosos causando paralisia neuromuscular (Wang *et al*., 2003; Su *et al*., 2004). CYN é uma hepatotoxina alcalóide de guanidina sulfatada cíclica que causa danos e necrose celular principalmente no fígado, rins, timo e coração de vertebrados (Terao *et al*., 1994; Humpage *et al*., 2000; Foscio *et al*., 2003). Ambas as toxinas, encontradas em ambientes aquáticos, são um risco em potencial para a saúde humana.

As cepas de *R.raciborskii* sul-americanas são conhecidas por produzirem derivados de saxitoxina, já cilindrospermopsina, amplamente encontrada em isolados da Austrália, Nova Zelândia e Ásia, até agora não foi identificada em cepas da América do Sul (Hawkins *et al*., 1985; Saker *et al*., 1999; Li *et al*., 2001; Wood *et al*., 2003; Chonudomkul *et al*., 2004; Everson *et al*., 2011). Algo também muito importante é o risco de *R. raciborskii* produzir metabólitos tóxicos ainda não caracterizados, como já foi mostrado em algumas cepas isoladas na Europa (Poniedziałek *et al*., 2015), algumas delas neurotoxicas (Vehovszky *et al*., 2013). Isso traz ênfase para a necessidade do registro da presença de *R. raciborskii* e sua propagação para novas áreas.

1.6. OBJETIVOS GERAIS

Este trabalho teve como objetivo estudar a fixação de nitrogênio em várias cepas da cianobactéria *Raphidiopsis* (=*Cylindrospermopsis*) *raciborskii* e foi fundamentado em três objetivos principais:

- eclarecer a relação taxonômica entre cepas heterocitadas e não-heterocitadas de *Raphidiopsis raciborskii*;

- avaliar a resposta de diferentes cepas à limitação de nitrogênio;

- revisar a rota evolutiva dos genes que se traduzem na enzima nitrogenase.

Estes objetivos serão apresentados em três capítulos.

2. REFERÊNCIAS

Aguilera, A., Gómez, E. B., Kaštovský, J., Echenique, R. O., & Salerno, G. L. 2018. The polyphasic analysis of two native *Raphidiopsis* isolates supports the unification of the genera *Raphidiopsis* and *Cylindrospermopsis* (Nostocales, Cyanobacteria). *Phycologia.* 57:130-46.

Briand, J. F., Leboulanger, C., Humbert, J. F., Bernard, C., & Dufour, P. 2004. *Cylindrospermopsis raciborskii* (Cyanobacteria) invasion at mid-latitudes: selection, wide physiological tolerance, or global warming?. *Journal of Phycology*. 40:231-8.

Chonudomkul, D., Yongmanitchai, W., Theeragool, G., Kawachi, M., Kasai, F., Kaya, K., & Watanabe, M. M. 2004. Morphology, genetic diversity, temperature tolerant and toxicity of *Cylindrospermopsis raciborskii* (Nostocales, Cyanobacteria) strains from Thailand and Japan. *FEMS Microbiology Ecology* 48: 345-355.

Everson, S., Fabbro, L., Susan Kinnear, S., Wright, P. 2011. Extreme differences in akinete, heterocyte and cylindrospermopsin concentrations with depth in a successive bloom involving *Aphanizomenon ovalisporum* (Forti) and *Cylindrospermopsis raciborskii* (Woloszynska) Seenaya and Subba Raju. *Harmful Algae* 10: 265-276.

Farris, J. S. 1974. Formal definitions of paraphyly and polyphyly. *Systematic Zoology*, 23: 548-554.

Flores, E., & Herrero, A. 2010. Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nature Reviews Microbiology*, 8:39-50.

Flores, E., Muro-Pastor, A. M. & Herrero, A. 1999. Cyanobacterial nitrogen assimilation genes and NtcA-dependent control of gene expression. *In* G. A. Peschek, W. Loffelhardt, and G. Schmetterer [EDs.] *The phototrophic prokaryotes*. Plenum Publishing Corporation, New York, NY, pp. 463-477.

Froscio, S. M., Humpage, A. R., Burcham, P. C., Falconer, I. R. 2003.

Cylindrospermopsin induced protein synthesis inhibition and its dissociation from acute toxicity in mouse hepatocytes. *Environmental Toxicology*, 18: 243-251.

Funk, D. J., & Omland, K. E. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, 34: 397-423.

Hawkins, P. R., Runnegar, M. T. C., Jackson, A. R. B., Falconer, I. R. 1985. Severe hepatotoxicity caused by the tropical cyanobacterium (blue-green-alga) *Cylindrospermopsis raciborskii* (Woloszynska) Seenaya and Subba Raju isolated from a domestic water-supply reservoir. *Applied and Environmental Microbiology*, 50: 1292-1295.

Hoiczyk, E., & Hansel, A. 2000. Cyanobacterial cell walls: News from an unusual prokaryotic envelope. *Journal of Bacteriology*, 182:1191-1199.

Hennig, W. 1999. Phylogenetic systematics. University of Illinois Press.

Humpage, A. R., Fenech, M., Thomas, P., Falconer, I. R. 2000. Micronucleus induction and chromosome loss in transformed human white cells indicate clastogenic and aneugenic action of the cyanobacterial toxin, cylindrospermopsin. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 472: 155-161.

Komárek, J. 2010. Recent changes (2008) in cyanobacteria taxonomy based on a combination of molecular background with phenotype and ecological consequences (genus and species concept). *Hydrobiologia*, 639:245-259.

Kopp, R. E., Kirschvink, J. L., Hilburn, I. A., Nash, C. Z. 2005. The Paleoproterozoic snowball Earth: A climate disaster triggered by the evolution of oxygenic photosynthesis. *Proceedings of the National Academy of Sciences*, 102:11131-11136.

Lagos, N., Onodera, H., Zagatto, P. A., Andrinolo, D., Azevedo, S. M., & Oshima, Y. 1999. The first evidence of paralytic shellfish toxins in the fresh water cyanobacterium *Cylindrospermopsis raciborskii*, isolated from Brazil. *Toxicon*, 37: 1359-1373.

Li, R., Carmichael, W. W., Brittain, S., Eaglesham, G. K., Shaw, G. R., Liu, Y., & Watanabe, M. M. 2001. First report of the cyanotoxins cylindrospermopsin and deoxycylindrospermopsin from *Raphidiopsis curvata* (Cyanobacteria). *Journal of Phycology,* 37: 1121-1126.

Mullineaux, P. M., Gallon, J. R., Chaplin, A. E. 1981. Acetylene reduction (nitrogen fixation) by cyanobacteria grown under alternating light-dark cycles. *FEMS Microbiol Letters,* 10:245-247.

Padisák, J. 1997. *Cylindrospermopsis raciborskii* (Woloszynska) Seenayya et Subba Raju, an expanding, highly adaptive cyanobacterium: worldwide distribution and review of its ecology. *Archiv Für Hydrobiologie Supplementband. Monographische Beitrage.* 107:563-93.

Paerl, H. W., Fulton, R. S., Moisander, P. H., & Dyble, J. 2001. Harmful Freshwater Algal Blooms, With an Emphasis on Cyanobacteria. *The Scientific World Journal*, 1:76-113.

Paerl, H. W., & Otten, T. G. 2013. Harmful cyanobacterial blooms: causes, consequences, and controls. *Microbial ecology*, 65:995-1010.

Poniedziałek, B., Rzymski, P., Kokociński, M., & Karczewski, J. 2015. Toxic potencies of metabolite(s) of non-cylindrospermopsin producing *Cylindrospermopsis raciborskii* isolated from temperate zone in human white cells. *Chemosphere*, 120:608-614.

Saker, M. L., Thomas, A. D., Norton, J. H. 1999. Cattle mortality attributed to the toxic cyanobacterium *Cylindrospermopsis raciborskii* in an outback region of north Queensland. *Environmental Toxicology,* 14: 179-182.

Šmarda, J., & Komrska, J. 1993. Advances in S-layer research of chroococcal cyanobacteria. *In* Advances in Bacterial Paracrystalline Surface Layers. Springer, Boston, MA, pp. 77-84.

Sinha, R., Pearson, L. A., Davis, T. W., Burford, M. A., Orr, P. T., & Neilan, B. A. 2012. Increased incidence of *Cylindrospermopsis raciborskii* in temperate zones-is climate

change responsible?. *Water Res*. 46:1408-19.

Stüken, A., Rücker, J., Endrulat, T., Preussel, K., Hemm, M., Nixdorf, B., Karsten, U., & Wiedner, C. 2006. Distribution of three alien cyanobacterial species (Nostocales) in northeast Germany: *Cylindrospermopsis raciborskii*, *Anabaena bergii* and *Aphanizomenon aphanizomenoides*. *Phycologia*. 45:696-703.

Su, Z., Sheets, M., Ishida, H., Li, F., Barry, W. H. 2004. Saxitoxin blocks L-type Ica. *Journal of Pharmacology and Experimental Therapeutics*, 308: 324-329.

Terao, K., Ohmori, S., Igarashi, K., Ohtani, I., Watanabe, M. F., Harada, K. I., Ito E., Watanabe, M. 1994. Electron microscopic studies on experimental poisoning in mice induced by cylindrospermopsin isolated from blue-green alga *Umezakia natans*. *Toxicon*, 32: 833-843.

Vehovszky, Á., Kovács, A. W., Farkas, A., Győri, J., Szabó, H., & Vasas, G. 2015. Pharmacological studies confirm neurotoxic metabolite(s) produced by the bloom-forming *Cylindrospermopsis raciborskii* in Hungary. *Environmental toxicology*, 30:501-512.

Wang, J., Salata, J. J., Bennett, P. B. 2003. Saxitoxin is a gating modifier of HERG K+ channels. *The Journal of General Physiology*, 121: 583-598.

Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., Moore, L. H. , Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., Starr, M. P., & Truper, H. G. 1987. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *International Journal of Systematic and Evolutionary Microbiology*, 37: 463-464.

Wolk, C. P. 1968. Movement of carbon from vegetative cells to heterocysts in *Anabaena cylindrica*. *Journal of Bacteriology*, 96:2138-2143.

Wood, S. A., Stirling, D. J. 2003. First identification of the cylindrospermopsin-producing cyanobacterium *Cylindrospermopsis raciborskii* in New Zealand. *New Zealand Journal of Marine and Freshwater Research,*, 37: 821-828.

CHAPTER ONE

PHYLOGENETIC RELATIONSHIPS AND HETEROCYTE PRODUCTION IN *RAPHIDIOPSIS* (=*CYLINDROSPERMOPSIS*) *RACIBORSKII* (CYANOBACTERIA)[1] STRAINS

Laísa Corrêa Braga Marques - Departamento de Botânica, Laboratório de Ficologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, 31270-901 Belo Horizonte, MG, Brazil

Alessandra Giani[2] - Departamento de Botânica, Laboratório de Ficologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, 31270-901 Belo Horizonte, MG, Brazil

1- Submitted on XX/XX/202X

2- Corresponding author: agiani@ufmg.br; phone: 55 31 34092681; fax: 55 31 34092671.

Running title: Phylogeny and heterocytes in *Raphidiopsis*

1. ABSTRACT

Raphidiopsis (=*Cylindrospermopsis*) *raciborskii* is a planktonic, bloom forming and potentially toxic cyanobacterial species. This species is known for its invasiveness and ability to produce toxins. It was first described as subtropical-tropical species, but it was later reported expanding towards waterbodies in temperate regions. The ability to produce heterocyte is what used to distinguish *Cylindrospermopsis* from the very similar *Raphidiopsis* genus. Recently, the two genera were recognized as being just one. However, detailed work about their main distinguishing trait, the heterocyte differentiation, is still scarce. In this study, we characterized 16 strains isolated from Brazilian environments using a polyphasic approach that included morphological, physiological and phylogenetic analyses. Strains with or without heterocytes were all submitted to nitrogen (N) deprivation experiments. Some strains isolated from a hypereutrophic waterbody never differentiated heterocytes even after prolonged N starvation periods. Others always produced heterocytes even under N rich conditions. A phylogenetic tree was constructed based in sequences from four DNA regions (16SrRNA, ITS, *rpoC1*, *nifH*) and showed some biogeographic patterns. Non-heterocytous Brazilian strains clustered together with heterocytous South American strains and their intermixed position in the tree suggest heterocyte loss during their evolution in South America. This study presents new and important insights into the phylogeography of *R. raciborskii*, including the presence of nitrogen-fixing and non-nitrogen-fixing strains in the species. It also provides a confirmation based in a larger number of strains and more genomic regions that *Raphidiopsis* and *Cylindrospermopsis* represent the same genus. Our experiments also showed that heterocyte development does not always depend on environmental nitrogen limitation.

Keywords: Filamentous cyanobacteria, Nostocales, biogeography, nitrogen limitation, heterocyst.

2. INTRODUCTION

The cyanobacterium *Raphidiopsis raciborskii* (Woloszynska) Aguilera, Berrendero Gómez, Kastovsky, Echenique & Salerno (Aguilera *et al*., 2018), until recently known and described in the literature as *Cylindrospermopsis raciborskii* (Wolozyńska) Seenayya & Subba, is an invasive, bloom forming, potentially toxic, N_2-fixing heterocytous cyanobacterium of seeming subtropical-tropical origin. It is also found expanding into rivers and waterbodies in temperate regions (Padisák, 1997; Briand *et al*., 2004; Stüken *et al*., 2006) helped by a set of characteristics such as high physiological tolerance, adaptation to both lower temperatures and the effects of global warming (Briand *et al*., 2004; Sinha *et al*., 2012).

This species exhibits broad temperature tolerance and is able to sustain biomass growth at temperatures as low as 12-17 ° C (Chonudomkul *et al*., 2004; Piccini *et al*., 2011; Dokulil *et al*., 2016) and as high as 35°C (Briand *et al*., 2004; Chonudomkul *et al*., 2004). An interesting and unusual characteristic for a cyanobacterium of tropical origin is that this species has shade tolerance and preference for low light intensities (Padisák & Reynolds, 1998; Briand *et al*., 2002), which is very advantageous in the self-shading situation of a dense bloom (Shafik *et al*., 2001; Briand *et al*., 2004). The role of buoyancy is also important to optimize access to light and nutrients uptake in a vertical gradient of nutrients in the water column (Padisák, 1997). This species has both a high uptake affinity (Wu *et al*., 2009) and high storage capacity for phosphorus (Istvánovics *et al*., 2000). This can be an advantage when there are sporadical addictions of phosphorus in the water column. Its worldwide dispersal is very concerning, not only because blooms cause a decrease in diversity in the aquatic environment, but also because this cyanobacterium can produce harmful toxins, such as cylindrospermopsin and saxitoxin (Lagos *et al.*, 1999; Mohamed 2007; Li *et al.*, 2001; McGregor *et al.*, 2011).

In *Raphidiopsis* (= *Cylindrospermopsis*) *raciborskii* morphology is variable. The filaments can be straight, curved, flexuous or coiled. The trichomes have attenuated terminal cells and aerotopes are present. As a member of Nostocales group, it can develop specialized cells for nitrogen fixation, called heterocytes, which are always terminal and solitary. The taxonomy of this species has been frequently reviewed. Wolozyńska (1912) described it as *Anabaena raciborskii*. Later, when the genus *Anabaenopsis* was established for species with terminal heterocytes (Miller, 1923), *Anabaena raciborskii* became *Anabaenopsis raciborskii*. But the observation that *Anabaenopsis raciborskii* had a different heterocyte differentiation process than all the other *Anabaenopsis* species, culminated in the separation of the genus *Cylindrospermopsis* from *Anabaenopsis* (Seenayya & Subba Raju, 1972).

Cylindrospermopsis was very similar to another bloom-forming genus, *Raphidiopsis* Fritsch & Rich (Komárková *et al*., 1999; Li *et al*., 2001; Moustaka-Gouni *et al*., 2009; Alster *et al*., 2010; Wu *et al*., 2011; Komárek, 2013). The main morphological difference between the two genera was the total absence of heterocytes in *Raphidiopsis*. They also shared similar ecological niches, since *Raphidiopsis* species had often been found to co-occur with *Cylindrospermopsis raciborskii* (Moustaka-Gouni *et al*., 2009; Wu *et al*., 2011). The taxonomic validity of both genera was intensively discussed because of their resemblance (Moustaka-Gouni *et al*., 2009; Stucken *et al*., 2010; Wu *et al*., 2011; Li *et al*., 2016). Recently, Aguilera *et al.* (2018) proposed the unification of *Cylindrospermopsis* and *Raphidiopsis* in a single genus, based on phylogeny and secondary structure of the 16S-23S ribosomal RNA intergenic spacer region. Following this publication, the name *Raphidiopsis* was adopted based in the taxonomic rules of earlier description of the genus and priority principle. However, the genus *Cylindrospermopsis*, and especially the species *C. raciborskii*, is better known worldwide and is found in a significantly higher number of publications.

In this study, we aimed to characterize 16 strains of *Raphidiopsis*

(*Cylindrospermopsis*), isolated from waterbodies in southeastern Brazil, using a polyphasic approach that included morphological, physiological and phylogenetic analyses. Heterocytous and non-heterocytous forms were selected. The four genes used for the phylogeny were the housekeeping gene 16S rRNA that has been extensively applied for studies in phylogenetic relationships among cyanobacteria (Fox *et al.*, 1992); the variable region of the 16S-23S rRNA larger fragment (ITS-L) (Iteman *et al.*, 2000); the RNA polymerase *rpoC1* (Wilson *et al.*, 2000); and the *nifH* gene that encodes the dinitrogenase reductase enzyme and is of environmental importance. The use of multigene analyses provides better accuracy and reliability of the phylogenetic inference, since more nucleotide data are placed into the analysis and consequently more genetic information is available and can be investigated (Gadagkar *et al.*, 2005). The sequencing results were compared with molecular data from *Raphidiopsis* and *Cylindrospermopsis* strains from around the world. At the same time, experiments were run with the 16 strains under N-deprived conditions to observe the response of the heterocytous and non-heterocytous forms. This research seeks to further clarify the phylogenetic relationships between the two forms, to assess their response to nitrogen limitation and to characterize the Brazilian strains according to their original environment.

3. MATERIALS AND METHODS

A total of 16 strains with and without heterocytes of *Raphidiopsis*, isolated from Brazilian waterbodies located in the State of Minas Gerais, were provided by the culture collection of the Phycology Laboratory of the Universidade Federal de Minas Gerais (Belo Horizonte, Brazil). Strain descriptions can be found in table S1 and S2. Cultures were grown in WC medium (Guillard, 1975) at 21°C temperature, 40 µmol of photons m^{-2} s^{-1} light intensity, and a 12h:12h light and dark photoperiod.

To simulate a gradual nitrogen limitation, as would happen in the environment, to

attempt inducing heterocyte differentiation in non-heterocytous forms and observe potential changes in heterocyte production, the 16 strains were submitted in replicates to experiments with decreasing nitrogen source according to the following procedure. First, the strains were cultivated on N-complete medium and were labeled as "Phase I" of the experiment, which also corresponds to the “control” condition. During their exponential growth, a 1 mL sub-sample of these cultures was taken and added into 10 mL of nitrogen free medium (=WC0). This step corresponded to the experimental phase II (Phase II). After 15 days, 1 mL from the "Phase II" culture was added into fresh 10 mL of WC0 medium (Phase III). After 15 more days, 1 mL of the "Phase III" culture was added into fresh 10 mL of WC0 medium (Phase IV). At every step, one sub-sample of each strain was also collected for cell counting and preserved with Lugol iodine solution. The vegetative cells and heterocytes in each sample were counted on a Fuchs-Rosenthal hemocytometer. The filtrate was collected for analysis of nitrate, which is the nitrogen form used in WC culture medium, and the concentration was determined using standard colorimetric techniques (APHA, 2005). In parallel experiments, the 16 strains were cultivated and harvested for genomic DNA extraction. Extraction followed the phenol-chloroform methodology based in Kurmayer *et al*. (2003), as previously described in Pimentel & Giani (2013). For the phylogenetic analysis the following primers were chosen. The 16S rRNA region was amplified using the primer pair 356F (de la Torre *et al.,* 2003) and 1391R (Turner *et al.,* 1999) generating an 890 bp amplicon. The ITS 16S-23S used the 322F primer to amplify the 16S rRNA region and the primer 340R to amplify the 23S rRNA (Iteman *et al.,* 2000). For the *nifH* gene, nifHF and nifHR primers (Gugger *et al*., 2005) amplified a 350 bp product and the *rpoC1* amplicon of 380 bp was obtained using the rpoC1F and rpoC1R primers (Gugger *et al*., 2005). Additionally, even though not used for the phylogeny, the presence of *hetR* gene (regulatory gene involved in heterocyte differentiation) was also tested in all strains, using the primer set described by Willis *et al.* (2016) that give an

amplicon of 182 bp. All reactions were carried out with a final volume of 50 µL mixture containing 10 ng DNA, Buffer IC 1X (Phoneutria), 80 µg/mL BSA, 200 µM dNTPs (Promega), 0,2 µM of each primer, 1U Taq Phoneutria. The PCR was performed in a C-1000 Touch Thermal Cycler (Bio-Rad) and then purified with (EDTA: ETOH). After purification, the amplification products were sequenced (Macrogen, South Korea).

The ITS 16S-23S region amplifies two fragments of 485 bp (ITS-S) and 710 bp (ITS-L). The ITS-L and ITS-S fragments are almost identical in sequence except for the regions encoding $tRNA^{Ile}$ and $tRNA^{Ala}$, which are absent in the shorter ITS-S (Iteman *et al.,* 2000). The ITS-L fragment is more frequently used in filamentous cyanobacteria phylogeny (Iteman, *et al.* 2000) and was chosen in this study. Its respective bands were cut from 0.8% agarose gel and purified. Thereafter, the purified product was ligated with TOPO-TA vector system (Thermo Scientific) and cloned into DH5α chemocompetent *Escherichia coli* cells. After separation by agarose gel electrophoresis, the amplification products were sequenced (Instituto René Rachou- Fiocruz sequencing platform, Brazil).

The resulting sequences were deposited in NCBI GenBank, and their accession numbers can be found in table S3. Sequences were compared with those available in public databases (National Center for Biotechnology Information and Joint Genome Institute) using BlastN search (Table S3). Selected sequences were aligned by Muscle (Edgar, 2004). The matrices were merged using Mesquite v.3.61. Phylogenetic trees were calculated using maximum parsimony (MP) and maximum likelihood (ML) analyses, both with 1000 replications of bootstrap, and Bayesian inference (BI) in MrBayes v3.2.2 (Ronquist *et al*., 2012), with two runs of four Markov chains for 40 million generations. The evolutionary model was chosen by jModelTest v2 software (Darriba *et al.,* 2012), using the Akaike information criterion that was TIM2+I+G. The CIPRES supercomputing facilities (Miller *et al.*, 2010) were used for calculation of the Bayesian trees. The single-region analysis of each

genomic region was also performed (Figure S1, S2, S3 and S4).

4. RESULTS

Morphological description and nitrogen depletion experiments

All strains exhibited solitary planktic trichomes, some were straight or flexuous with or without heterocytes. These phenotypes were represented in Figure 1. They were subsymmetric, unconstricted, with tapering terminal cells, without mucilaginous sheaths. Flexuous trichomes were observed only in a few strains isolated from Pampulha Reservoir. Terminal cells were conical or gradually attenuated, with rounded or sharply pointed ends. The strains from Lagoa Santa Lake were all heterocytous and had straight trichomes, with sharply pointed to needle-like terminal cells. Aerotopes were present in all strains.

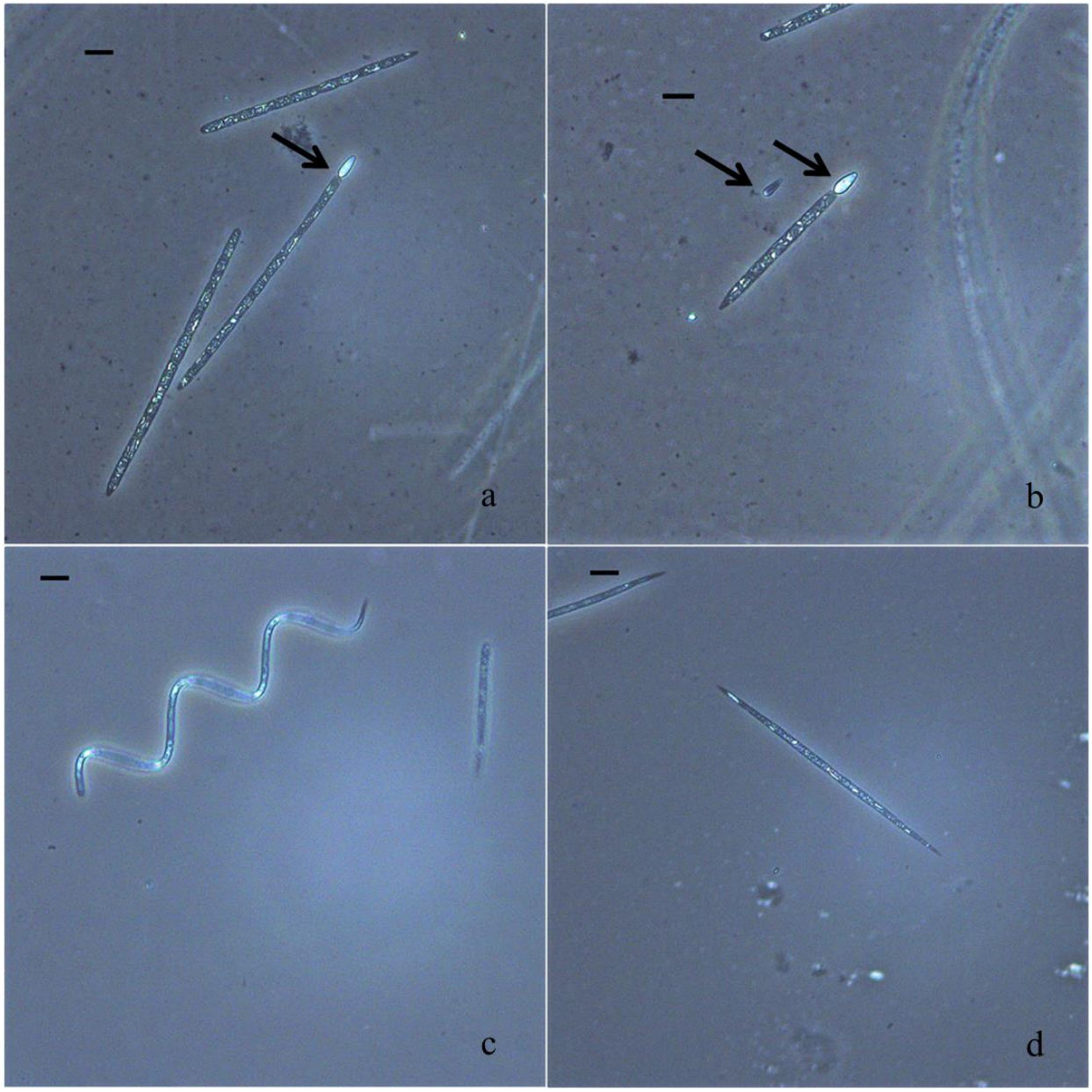


Fig. 1- Light micrographs that represent heterocytous (a-b) and non-heterocytous (c-d) trichomes of *Raphidiopsis* (=*Cylindrospermopsis*) strains. Figures a,b,e: straight trichomes; figure c: flexuous trichome. Arrows point to heterocytes. Scale bars = 10 µm.

In Phase I, nitrate concentration was according to WC medium protocol of 58000 µg/L. In Phase II, nitrate concentration decreased to 3172.59 µg/L, in Phase III to 30.68 µg/L and in Phase IV just mean values of 2.14 µg/L of nitrate were recorded as seen in table 1

Table 1- Nitrate concentration (µg/L), mean and standard error in four experimental phases.

Phase	Mean	Standard Error
I	5.80E+04	-
II	3.17E+03	249.71
III	3.07E+01	6.58
IV	2.14E+00	0.89

Figure 2 represents the percentage of heterocytes relative to vegetative cells observed in the trichomes of the 16 strains submitted to progressive N depletion. Some strains (UFMG-37, 40 and 42) only presented heterocytes in Phase III and IV, i.e, after severe nitrogen restriction. Other strains (UFMG-34, 36, 44, 182, 185, 186, 187, 188, 232) showed heterocyte presence even when the N source was abundantly available (Phase I and II). Of all strains analyzed only four (UFMG-181, 183, 233 and 234) could never differentiate heterocytes, even in the last experimental phase when growing in N-free medium (WC0).

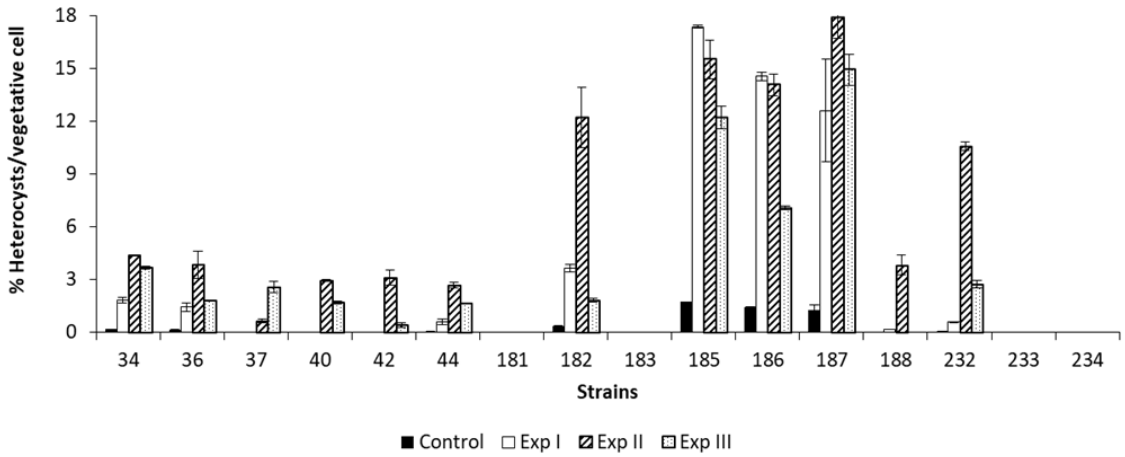


Fig.2- Percentage of heterocytes relative to vegetative cells in each strain and experimental phase. For details see text.

After the decrease in nitrogen concentration in Phase II, the percentage of heterocytes increased in strains UFMG-34, 36, 44, 185, 186 and 188 (Fig. 2). During Phase III, the percentage of heterocytes also increased in strains UFMG-34, 36, 187, 232, 182 and 188. But in Phase IV, after 45 days growth in nitrogen depleted medium, most heterocytous strains

showed a decrease in heterocyte number relative to vegetative cells, except for strain 37 (Fig.2). It is worth noting that strain 182 that produced heterocytes in all four experimental phases and strain183 that never produced any were collected in the same bloom, in other words, co-occurred in the same environment (Pampulha Reservoir).

The genes chosen to identify nitrogen fixation (*nifH*) and heterocyte production (*hetR*) were observed in all strains except in those four that never produced heterocytes (Fig. 3), even under severe N starvation (Fig. 2).

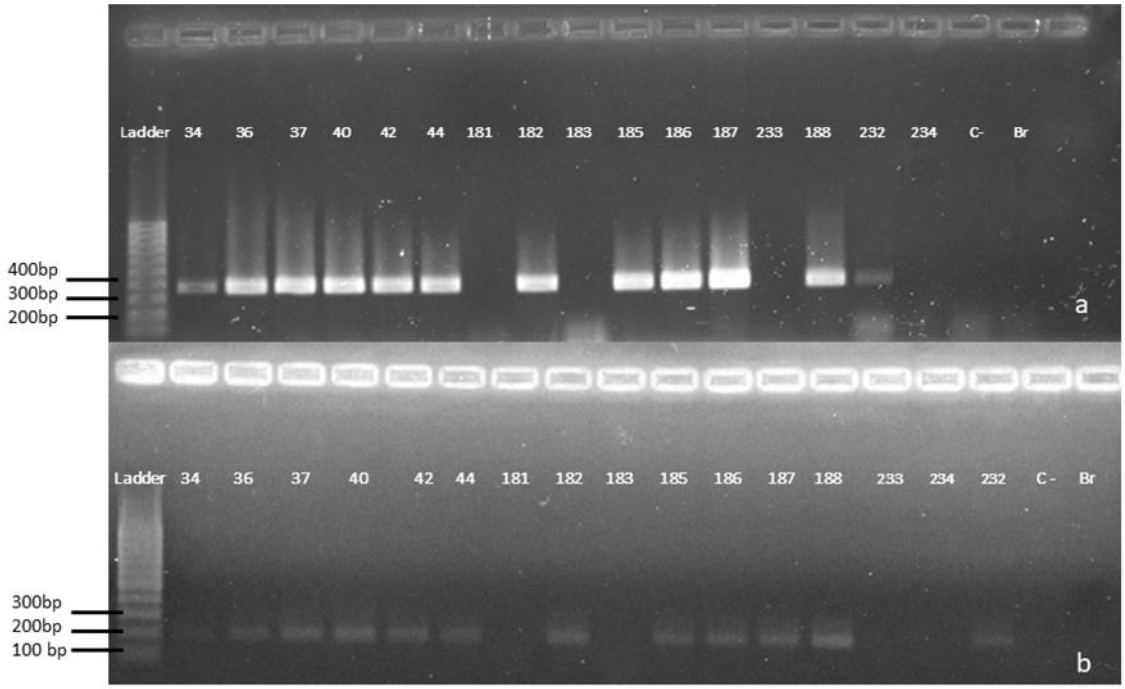
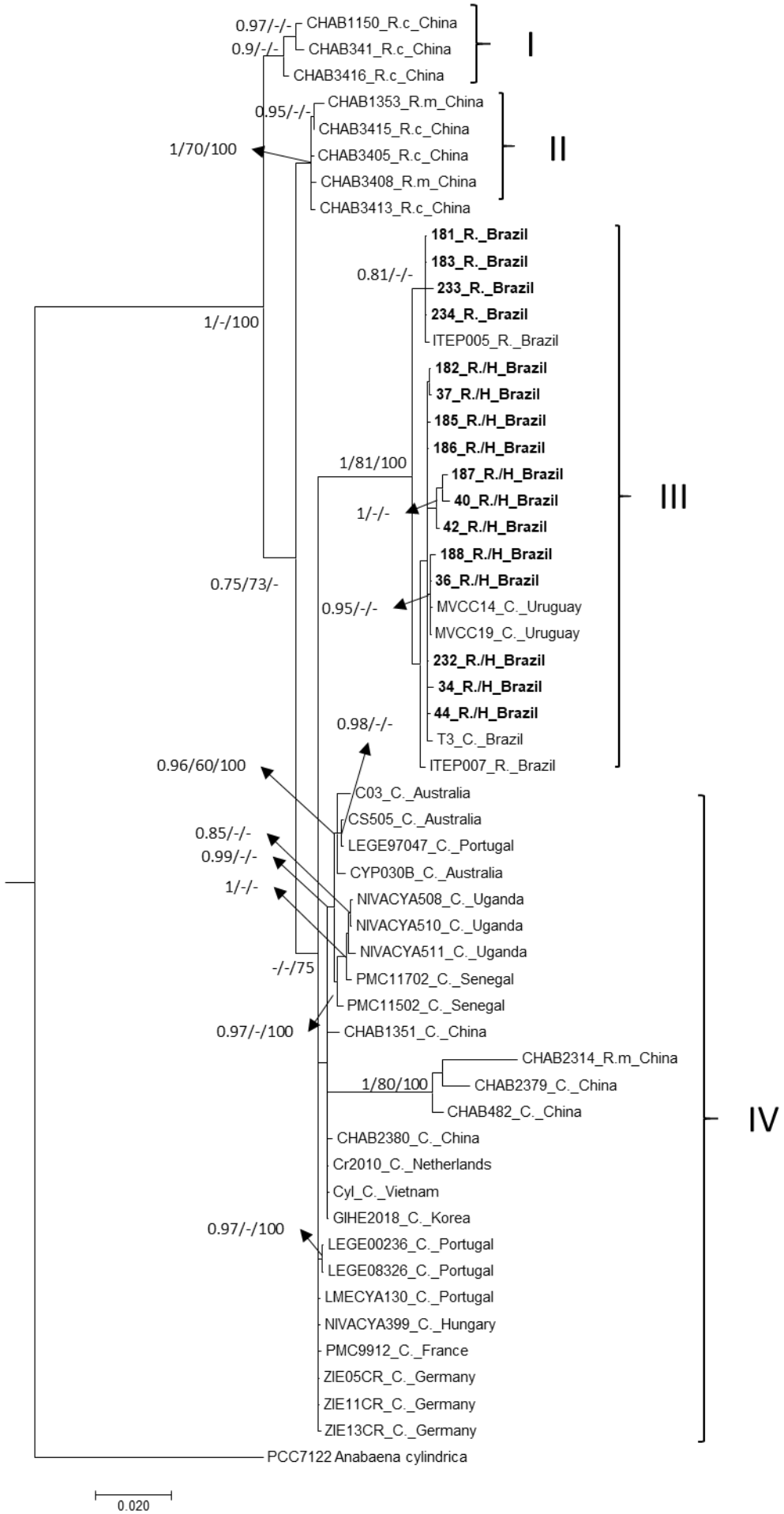


Fig.3- Ethidium bromide-stained gels (1.2% agarose) showing PCR amplification products corresponding to *nifH* (a) and *hetR* (b) genes. First lane: molecular weight marker. Numbers on top of the other lanes represent the correspondent strain. Negative control (C-) and blank (Br).

Concatenated phylogenetic analysis

A total of 58 strain sequences, 42 of which were obtained from the NCBI GenBank, were included in the phylogenetic analysis based on the 16S rRNA, 16S-23S rRNA ITS-L, *rpoC1* and *nifH* gene sequences. The concatenated tree showed that the genera *Raphidiopsis* and *Cylindrospermopsis* were not assembled in separated clusters, but grouped and mixed together (Fig.4). The analysis showed that the strains formed four phylogenetic distinct

clades. *Raphidiopsis* from China clustered in two separated clades (I and II), while *Cylindrospermopsis* from China grouped with strains from Europe, Asia, Africa and Australia in clade IV. Clade III comprised only strains from America and all the Brazilian strains of the present work. Interestingly, the four Brazilian strains from Pampulha Reservoir that could never produce heterocytes were grouped together and were in the same cluster of *Raphidiopsis* ITEP005, also a Brazilian strain from Tapacurá, Pernambuco State. The phylogeny showed no association to the morphological features of the 16 studied strains, namely the straight or flexuous trichomes and the presence or not of heterocyte.


0.97/-/-
CHAB1150_R.c_China
0.9/-/-
CHAB341_R.c_China
CHAB3416_R.c_China
I
0.95/-/-
CHAB1353_R.m_China
CHAB3415_R.c_China
1/70/100
CHAB3405_R.c_China
CHAB3408_R.m_China
CHAB3413_R.c_China
II
181_R._Brazil
183_R._Brazil
0.81/-/-
233_R._Brazil
234_R._Brazil
1/-/100
ITEP005_R._Brazil
182_R./H_Brazil
37_R./H_Brazil
185_R./H_Brazil
186_R./H_Brazil
1/81/100
187_R./H_Brazil
1/-/-
40_R./H_Brazil
42_R./H_Brazil
III
188_R./H_Brazil
0.75/73/-
36_R./H_Brazil
0.95/-/-
MVCC14_C._Uruguay
MVCC19_C._Uruguay
232_R./H_Brazil
34_R./H_Brazil
44_R./H_Brazil
0.98/-/-
T3_C._Brazil
0.96/60/100
ITEP007_R._Brazil
C03_C._Australia
CS505_C._Australia
0.85/-/-
LEGE97047_C._Portugal
0.99/-/-
CYP030B_C._Australia
1/-/-
NIVACYA508_C._Uganda
NIVACYA510_C._Uganda
NIVACYA511_C._Uganda
-/-/75
PMC11702_C._Senegal
PMC11502_C._Senegal
0.97/-/100
CHAB1351_C._China
CHAB2314_R.m_China
1/80/100
CHAB2379_C._China
CHAB482_C._China
IV
CHAB2380_C._China
Cr2010_C._Netherlands
Cyl_C._Vietnam
GIHE2018_C._Korea
0.97/-/100
LEGE00236_C._Portugal
LEGE08326_C._Portugal
LMECYA130_C._Portugal
NIVACYA399_C._Hungary
PMC9912_C._France
ZIE05CR_C._Germany
ZIE11CR_C._Germany
ZIE13CR_C._Germany
PCC7122 Anabaena cylindrica
0.020

1

Fig.4- Phylogenetic analysis of *Raphidiopsis* / *Cylindrospermopsis* strains derived from concatenated data using 16S rRNA, ITS-L, *rpoC1* and *nifH* sequences and based on topology given by Bayesian analysis. KT290325.1 *A.cylindrica* was selected as the outgroup taxon. Support values (BI/ML/MP) for the branches are indicated at the nodes of the tree when bootstrap values are above 60% and BI is above 0.8. Strains sequenced in this study are indicated in bold.

5. DISCUSSION

Raphidiopsis (=*Cylindrospermopsis*) *raciborskii* is known to be capable to use multiple forms of combined nitrogen, such as ammonium (NH_4^+), nitrate (NO_3^-), and urea (CH_4N_2O,), and it is also diazotrophic, process that occurs in heterocytes (Saker & Neilan, 2001; Sinha *et al.*, 2014). In general, it is acknowledged that the presence of ammonium and nitrate should repress N_2 fixation (Meeks *et al.*, 1983; Ohmori & Hattori, 1972; Haselkorn *et al.*, 1998) and heterocyte formation (Bottomley *et al.*, 1979; Adams & Duggan, 1999; Herrero *et al.*, 2001). However, in the present study, we observed that some strains produced heterocytes all the time, even in N rich medium. On the other hand, some strains would never produce heterocytes, and for some their production was only observed when nitrogen was depleated from the medium (Phase IV). This suggests that the strains are in fact representing different ecotypes, some even co-occurring in the same environment.

Actually, some studies have demonstrated that heterocytes can differentiate in the presence of nitrogen. For example, *Anabaena variabilis,* growing on glutamine as the sole nitrogen source, gave rise to patterned heterocytes although nitrogenase activity was not detected (Thiel & Leone, 1986). Sprőber *et al.* (2003) also found that heterocytes in *C.raciborskii* (*R. raciborskii*) could differentiate even when ammonium was supplemented to the growth medium and Willis *et al.* (2016) recorded similar observations with ammonium and nitrate. This shows that heterocyte development can occur even with a nitrogen supply in the environment. However, the presence of heterocytes does not seem to imply that the cell is actively fixing nitrogen. Vintila *et al.* (2007), for example, observed nitrogenase activity in the heterocytous *Nodularia spumigena* only when it was growing in a medium lacking

combined nitrogen.

From Phase I to Phase II, the percentage of heterocytes relative to vegetative cells increased in most strains. After long nitrogen restriction in Phase III, however, most heterocytous strains showed a decrease in the percentage of produced heterocytes, possibly because the lack of abundant nitrogen was already impacting the cell growth and metabolism. According to Duval *et al*. (2013) nitrogenase reduction of N_2 to NH_4^+ results in a reduction in growth rates, since this is an energy-expensive process that reduces the available energy for growth. The ATP-driven reduction of one N_2 with evolution of one H_2 requires a minimum of $8e^-$ and the hydrolysis of 16 ATP molecules. For this reason, N fixation cannot fulfill the cellular requirement to reach maximum growth rates even when all other nutrients are replete (Willis *et al.*, 2016), which contradicts the common misconception that fixating nitrogen is an advantage that helps cyanobacteria in bloom formation.

In the phylogeny tree (Fig. 4), Brazilian non-heterocytous strains clustered together with heterocytous South American strains and their intermixed position in the tree indicates that several heterocyte losses may have occurred during their evolution in South America. Also, *Raphidiopsis* from China formed a single cluster, whereas former Chinese *Cylindrospermopsis* strains grouped with African, Australian and European strains. This suggests the existence of two lines of Chinese strains sufficiently separated from each other, one capable of forming heterocytes, more closely related to strains found in other parts of the world, and another that cannot. It also points to the existence of different non-heterocytous *Raphidiopsis* genotypes, one from South America and one from Asia, as Aguilera *et al.* (2019) also concluded in their work. Our data reinforce their study entirely, confirming that to consider *Raphidiopsis* a monophyletic group, *Raphidiopsis* and *Cylindrospermopsis* as a single genus.

In addition, the single-region analysis (Figs. S1, S2, S3 and S4) and the multigene

analysis proved to be inefficient to identify and separate sequences at the species level. This may be due to the regions used or due to the lack of sufficient genetic database that these species have. In fact, *Raphidiopsis* / *Cylindrospermopsis* species have been mostly identified based on morphological characters (Komárková 1998, Komárek 1999) and most species still lack sequenced genomes to really define whether they are real. Although the ITS and *rpoC1* regions are said to have a better discriminatory analysis power at the species level (Wilson *et al.,* 2000), the present study indicates that these regions were not able to separate strains at species level for *Raphidiopsis* / *Cylindrospermopsis*. The concatenated gene tree with 16S rRNA, ITS, *rpoC1* and *nifH* regions was compared with a second concatenated gene tree without *nifH* (Fig. S5) to assess if the *nifH* sequences were grouping heterocytous strains together, since non-heterocytous strains would not have this gene. But, we observed that the absence of *nifH* did not interfere with the results. The tree kept the same topology.

In this study, the phylogenetic analysis showed a biogeographic clustering of strains. Previous studies on phylogeny of *R. raciborskii* (*C. raciborskii*) have already observed biogeographic variation associated to genetic variability (Neilan *et al*., 2003; Gugger *et al*., 2005; Moreira *et al*., 2014; Abreu *et al*., 2018) and, especially, the existence of a South American cluster (Aguilera *et al.,* 2018; Abreu *et al*., 2018) and a Chinese *Raphidiopsis* cluster. This may suggest that these strains went through a longer evolutionary time, i.e, they diverged from the others long ago. Clades I and II in our tree topology suggest that the Chinese *Raphidiopsis* strains came from a lineage that diverged sooner. The dispersion to the American continent is also very ancient, while the dispersion to the rest of the world seem to be more recent. Because some Chinese *Cylindrospermopsis* strains are mixed within cluster IV with strains from all over the world, we suggest that this could indicate a recent secondary dispersion to China.

Various hypotheses have been put forward to describe the spread routes of *R.*

raciborskii (*C. raciborskii*) and its origin. Padisák (1997) suggested primary roots in African tropical lakes, and eventual spread to other equatorial regions such as Indonesia and Central America. The dispersion into tropical, subtropical, and temperate regions would have happened from a secondary radiation center starting in Australia. Neilan *et al.* (2003) further investigated *R. raciborskii* (*C. raciborskii*) phylogeography and their findings support the presence of distinct Australian, European and North/South American phylotypes. Vico *et al.* (2020) investigated toxin production and distribution around the world and agreed with Padisák's proposal of tropical Africa as a primary evolutionary center of *R. raciborskii*. They hypothesize that the last common ancestor to all current *R. raciborskii* would be African and unable to produce any toxin. Cylindrospermopsin (CYL) production, but not saxitoxin (PSP), would arise later, and thereafter the species would spread to America and Oceania. They suggest that later a single case of horizontal gene transfer occurred in the development of PSP production. Their research however lacks support as their ITS region phylogenetic tree bootstrap values were very low.

Contrary to Pádisak's proposal Gugger *et al.* (2005) hypothesized that the glaciations or other climatic events in the Pleistocene led to the extinction of *Raphidiopsis*/*Cylindrospermopsis* in most of its geographical distribution and allowed its survival only in warm refuge areas. Later, the rise in temperature allowed the spread from warm refuge areas inside Europe and America towards other continents. Haande *et al.* (2008) acknowledges the hypothesis of warm refuge areas, and also infers that the American population diverged earlier than others. More recently, Moreira *et al.* (2014) suggested that *Raphidiopsis*/*Cylindrospermopsis*' initial dispersal core was in America, and that the cyanobacterium had spread to Africa when the two continents were united. Later, they suggest, there was a migration to Asia and Australia and lastly to Europe. Their assumption was based on the genetic similarity between European, Asian and Australian isolates, and on

the fact that the American strains are the most divergent, indicating that European, Asian and Australian strains emerged later in this species' evolution. These evidences are consistent with the relatively recent description of this species in Europe (Saker *et al.*, 2003; Briand *et al.*, 2002; De Hoyos *et al.*, 2004; Manti *et al.*, 2005; Stüken *et al*., 2006; Sukenik *et al.*, 2012). But their work does not take into consideration the Asian *Raphidiopsis* strains.

Our study found that strains UFMG-181, 183, 233 and 234 are phylogenetically close and do not possess either *nifH* or *hetR* genes. These strains were isolated from Pampulha Reservoir, a shallow hypereutrophic lake with high levels of dissolved nitrogen (Batista *et al.*, 2018; Silva *et al*., 2019). However, these genes were present in other strains isolated from the same environment, during the same period (for example UFMG182 and UFMG183). This seems to suggest that the loss of these genes may have occurred in some strains, which lost the capacity to produce heterocytes and fix nitrogen. This finding would support the hypothesis that populations without heterocytes might have arisen from heterocytous ones multiple times, by natural selection, in nitrogen-rich environments.

Similar to this study, Moissander *et al*. (2012) found that the loss of heterocytes and N_2 fixation in *R. brookii* could be due to a consistent availability of dissolved inorganic or organic nitrogen sources in its natural environment. This agrees with Stucken *et al.* (2010) that observed synteny surrounding the N_2-fixation cluster of *R.brooki* D9 and *C.raciborskii* CS-505 and, since D9 did not have the cluster, they concluded that it might have been selectively lost in this strain. Hess (2011) considered this loss as another case of reductive genome evolution. Thus, the N_2-fixation cluster is not part of the core genome (i.e., the combination of genes present in all strains) and we can conclude and agree with Aguilera *et al*. (2018) that its phenotype should not be used to characterize the main diagnostic difference between *Cylindrospermopsis* and *Raphidiopsis*.

This study provides new insights into the phylogeography of the species *R. raciborskii*,

including nitrogen-fixing and non-nitrogen-fixing strains. We also observed that both strains may co-occur in nature and suggest that populations without heterocytes may likely have arisen from heterocytous ones, by natural selection in nitrogen-rich environments, like for example Pampulha Reservoir. In the present study, by adding more strains and more genomic regions to the findings of Aguilera *et al.* (2018), we further confirm that *Raphidiopsis* and *Cylindrospermopsis* are the same. Finally, our experiments also showed that heterocytes can develop even in individuals growing in a nitrogen rich environment.

6. ACKNOWLEDGMENTS

We thank CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) for a scholarship to L.C.B.M and FAPEMIG (Fundação de Apoio a Pesquisa de Minas Gerais) for a research grant to A.G. A.G is a CNPq fellow. We also thank Instituto René Rachou- Fiocruz (Belo Horizonte, Brazil) for the sample sequencing. We are especially grateful to Prof. Joni Lima (UFMG) for his dedicated support with the cloning methods.

7. REFERENCES

Abreu, V. A. C., Popin, R. V., Alvarenga, D. O., Schaker, P. D. C., Hoff-Risseti, C., Varani, A. M., & Fiore, M. F. 2018. Genomic and genotypic characterization of *Cylindrospermopsis raciborskii*: Toward an intraspecific phylogenetic evaluation by comparative genomics. *Front. Microbiol.* 9:1-12.

Adams, D. G., & Duggan, P. S. 1999. Tansley Review No. 107. Heterocyst and akinete differentiation in cyanobacteria. *New Phytol.* 144:3-33.

Aguilera, A., Gómez, E. B., Kaštovský, J., Echenique, R. O., & Salerno, G. L. 2018. The polyphasic analysis of two native *Raphidiopsis* isolates supports the unification of the genera *Raphidiopsis* and *Cylindrospermopsis* (Nostocales, Cyanobacteria). *Phycologia.* 57:130-46.

Alster, A., Kaplan-Levy, R. N., Sukenik, A., & Zohary, T. 2010. Morphology and phylogeny of a non-toxic invasive *Cylindrospermopsis raciborskii* from a Mediterranean Lake. *Hydrobiologia*. 639:115-28.

APHA (American Public Health Association) .1998. Standard methods for the examination of water and wastewater, 20th ed. American Public Health Association Publications, Washington DC.

Batista, A. M. M., Figueredo, C. C., & Giani, A. 2018. Variability in a permanent cyanobacterial bloom: species-specific responses to environmental drivers. *FEMS Microbiol. Ecol.* 94: fiy197.

Bottomley, P. J., Grillo, J. F., Van Baalen, C., & Tabita, F. R. 1979. Synthesis of nitrogenase and heterocysts by *Anabaena* sp. CA in the presence of high levels of ammonia. *J. Bacteriol*. 140:938-43.

Briand, J. F., Leboulanger, C., Humbert, J. F., Bernard, C., & Dufour, P. 2004. *Cylindrospermopsis raciborskii* (Cyanobacteria) invasion at mid-latitudes: selection, wide physiological tolerance, or global warming?. *Journal of Phycology*. 40:231-8.

Briand, J. F., Robillot, C., Quiblier-Lloberas, C., Humbert, J. F., Couté, A., & Bernard, C. 2002. Environmental context of *Cylindrospermopsis raciborskii* (Cyanobacteria) blooms in a shallow pond in France. *Water Res*. 36:3183-92.

Chonudomkul, D., Yongmanitchai, W., Theeragool, G., Kawachi, M., Kasai, F., Kaya, K., & Watanabe, M. M. 2004. Morphology, genetic diversity, temperature tolerance and toxicity of *Cylindrospermopsis raciborskii* (Nostocales, Cyanobacteria) strains from Thailand and Japan. *FEMS Microbiol. Ecol.* 48:345-55.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods.* 9:772.

Hoyos, C. D., Negro, A. I., & Aldasoro Martín, J. J. 2004. Cyanobacteria distribution

and abundance in the Spanish water reservoirs during thermal stratification. *Limnetica*. 23:119-32.

De la Torre, J. R., Goebel, B. M., Friedmann, E. I., & Pace, N. R. 2003. Microbial diversity of cryptoendolithic communities from the McMurdo Dry Valleys, Antarctica. *Appl. Environ. Microbiol.* 69:3858-67.

Dokulil, M.T. 2016. Vegetative survival of *Cylindrospermopsis raciborskii* (Cyanobacteria) at low temperature and low light. *Hydrobiologia.* 764:241-7.

Duval, S., Danyal, K., Shaw, S., Lytle, A. K., Dean, D. R., Hoffman, B. M., Antony, E., & Seefeldt, L. C. 2013. Electron transfer precedes ATP hydrolysis during nitrogenase catalysis. *Proc. Natl. Acad. Sci. USA*. 110:16414-9.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792-7.

Fox, G. E., Wisotzkey, J. D., & Jurtshuk Jr, P. 1992. How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *Int. J. Syst. Evol. Microbiol.* 42:166-70.

Gadagkar, S. R., Rosenberg, M. S., & Kumar, S. 2005. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. *J. Exp. Zool. B.* 304:64-74.

Gugger, M., Molica, R., Le Berre, B., Dufour, P., Bernard, C., & Humbert, J. F. 2005. Genetic diversity of *Cylindrospermopsis* strains (Cyanobacteria) isolated from four continents. *Appl. Environ. Microbiol.* 71:1097-100.

Guillard, R. R. L. 1975. Culture of phytoplankton for feeding marine invertebrates. *In* Smith, W.L. & Chanley, M.H. [Eds.] *Culture of marine invertebrate animals*. Springer, Boston, MA, pp. 29-60.

Haande, S., Rohrlack, T., Ballot, A., Røberg, K., Skulberg, R., Beck, M., & Wiedner,

C. 2008. Genetic characterisation of *Cylindrospermopsis raciborskii* (Nostocales, Cyanobacteria) isolates from Africa and Europe. *Harmful algae.* 7:692-701.

Haselkorn, R., Schlictman, D., Jones, K., & Buikema, W. J. 1998. Heterocyst differentiation and nitrogen fixation in cyanobacteria. *In* Elmerich, C., Kondorosi, A., Newton, W.E. [Eds.] *Biological Nitrogen Fixation for the 21st Century*. Springer, Dordrecht, pp. 93-6.

Herrero, A., Muro-pastor, A. M., & Flores, E. 2001. MINIREVIEW Nitrogen Control in Cyanobacteria. *J. Bacteriol.* 183:411-25.

Hess, W. R. 2011. Cyanobacterial genomics for ecology and biotechnology. *Curr. Opin. Microbiol.* 14:608-14.

Isvánovics, V., Shafik, H. M., Présing, M., & Juhos, S. 2000. Growth and phosphate uptake kinetics of the cyanobacterium, *Cylindrospermopsis raciborskii* (Cyanophyceae) in throughflow cultures. *Freshwater biology,* 43:257-75.

Iteman, I., Rippka, R., De Marsac, N. T., & Herdman, M. 2000. Comparison of conserved structural and regulatory domains within divergent 16S rRNA-23S rRNA spacer sequences of cyanobacteria. *Microbiology.* 146:1275-86.

Komárek, J. 1999. Übersicht der planktischen Blaualgen (Cyanobakterien) im Elbe Flussgebiet. *IKSE/MKOL, Magdeburg*. 133:1-53.

Komárek, J. 2013. *Süßwasserflora von Mitteleuropa, Bd. 19/3: Cyanoprokaryota*. Springer, Berlin, 1131 pp.

Komárková, J. 1998. The tropical planktonic genus *Cylindrospermopsis* (Cyanophytes, cyanobacteria). *In* Azevedo, M.T.P., Santos, D.P., Pinto, L.S.C., Menezes, M., Fujii, M.T., Yokoya, N.S., Senna, P.A.C. & Guimarães, S.M.P.B., [Eds.] *Anais do IV Congresso Latino-Americano de Ficologia*. Sociedade Ficológica da America Latina e Caribe. São Paulo, Brasil, pp. 327-40.

Komárková, J., Laudares-Silva, R., & Senna, P. A. C. 1999. Extreme morphology of *Cylindrospermopsis raciborskii* (Nostocales, Cyanobacteria) in the Lagoa do Peri, a freshwater coastal lagoon, Santa Catarina, Brazil. *Algological Studies/Arch. Hydrobiol. Suppl*. 129:207-22.

Kurmayer, R., Christiansen, G., & Chorus, I. 2003. The abundance of microcystin-producing genotypes correlates positively with colony size in *Microcystis* sp. and determines its microcystin net production in Lake Wannsee. *Appl. Environ. Microbiol.* 69:787-95.

Lagos, N., Onodera, H., Zagatto, P. A., Andrinolo, D., Azevedo, S. M., & Oshima, Y. 1999. The first evidence of paralytic shellfish toxins in the freshwater cyanobacterium *Cylindrospermopsis raciborskii*, isolated from Brazil. *Toxicon*. 37:1359-73.

Li, R., Carmichael, W. W., Brittain, S., Eaglesham, G. K., Shaw, G. R., Mahakhant, A., Noparatnaraporne, N., Yongmanitchai, W., & Watanabe, M. M. 2001. Isolation and identification of the cyanotoxin cylindrospermopsin and deoxy-cylindrospermopsin from a Thailand strain of *Cylindrospermopsis raciborskii* (Cyanobacteria). *Toxicon*. 39:973-80.

Li, X., Li, S., Kong, R., & Li, R. 2016. Molecular separation of two long taxonomically debated cyanobacterial genera *Cylindrospermopsis* and *Raphidiopsis* (Nostocales) based on the ITS-L phylogeny. *Harmful Algae*. 57:88-97.

Manti, G., Mattei, D., Messineo, V., Melchiorre, S., Bogialli, S., Sechi, N., Casiddu, P., Luglie, A., Di Brizio, M., Bruno, M. 2005. First report of *Cylindrospermopsis raciborskii* in Italy. *Harmful Algae News*. 28:8-9.

McGregor, G. B., Sendall, B. C., Hunt, L. T., & Eaglesham, G. K. 2011. Report of the cyanotoxins cylindrospermopsin and deoxy-cylindrospermopsin from *Raphidiopsis mediterranea* Skuja (Cyanobacteria/Nostocales). *Harmful Algae*. 10:402-10.

Meeks, J. C., Wycoff, K. L., Chapman, J. S., & Enderlin, C. S. 1983. Regulation of expression of nitrate and dinitrogen assimilation by *Anabaena* species. *Appl. Environ.*

Microbiol. 45:1351-9.

Miller, M. A., Pfeiffer, W., & Schwartz, T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In 2010 gateway computing environments workshop (GCE), Ieee, pp. 1-8.

Miller, V. V. 1923. K sistematike sistematike roda *Anabaena*. *Bory. Arch. Russk. Protistol*. 2:116-26.

Mohamed, Z. A. 2007. First report of toxic *Cylindrospermopsis raciborskii* and *Raphidiopsis mediterranea* (Cyanoprokaryota) in Egyptian fresh waters. *FEMS Microbiol. Ecol*. 59:749-61.

Moisander, P.H., Cheshire, L.A., Braddy, J., Calandrino, E.S., Hoffman, M. 2012. Facultative diazotrophy increases *Cylindrospermopsis raciborskii* competitiveness under fluctuating nitrogen availability. *FEMS Microbiol. Ecol.* 79:800-11.

Moreira, C., Fathalli, A., Vasconcelos, V., & Antunes, A. 2014. Phylogeny and biogeography of the invasive cyanobacterium *Cylindrospermopsis raciborskii*. *Archi. Microbiol.* 197:47-52.

Moustaka-Gouni, M., Kormas, K. A., Vardaka, E., Katsiapi, M., & Gkelis, S. 2009. *Raphidiopsis mediterranea* Skuja represents non-heterocytous life-cycle stages of *Cylindrospermopsis raciborskii* (Woloszynska) Seenayya et Subba Raju in Lake Kastoria (Greece), its type locality: evidence by morphological and phylogenetic analysis. *Harmful Algae*. 8:864-72.

Neilan, B. A., Saker, M., Fastner, J., Törökné, A., & Burns, B. P. 2003. Phylogeography of the invasive cyanobacterium *Cylindrospermopsis raciborskii*. *Mol. Ecol*. 12:133-40.

Ohmori, M., & Hattori, A. 1972. Effect of nitrate on nitrogen-fixation by the blue-green alga *Anabaena cylindrica*. *Plant Cell Physiol.* 13:589-99.

Padisák, J. 1997. *Cylindrospermopsis raciborskii* (Woloszynska) Seenayya et Subba Raju, an expanding, highly adaptive cyanobacterium: worldwide distribution and review of its ecology. *Archiv Für Hydrobiologie Supplementband. Monographische Beitrage.* 107:563-93.

Padisák, J., & Reynolds, C. S. 1998. Selection of phytoplankton associations in Lake Balaton, Hungary, in response to eutrophication and restoration measures, with special reference to the cyanoprokaryotes. *Hydrobiologia*. 384:41-53.

Piccini, C., Aubriot, L., Fabre, A., Amaral, V., González-Piana, M., Giani, A., Figueredo, C. C., Vidal, L., Kruk, C., & Bonilla, S. 2011. Genetic and eco-physiological differences of South American *Cylindrospermopsis raciborskii* isolates support the hypothesis of multiple ecotypes. *Harmful Algae*. 10:644-53.

Pimentel, J. S., & Giani, A. 2013. Estimating toxic cyanobacteria in a Brazilian reservoir by quantitative real-time PCR, based on the microcystin synthetase D gene. *J. Appl. Phycol*. 25:1545-54.

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., & Huelsenbeck, J. P. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61:539-42.

Saker, M. L., & Neilan, B. A. 2001. Varied diazotrophies, morphologies, and toxicities of genetically similar isolates of *Cylindrospermopsis raciborskii* (Nostocales, Cyanophyceae) from northern Australia. *Appl. Environ. Microbiol*. 67:1839-45.

Saker, M. L., Nogueira, I. C., Vasconcelos, V. M., Neilan, B. A., Eaglesham, G. K., & Pereira, P. 2003. First report and toxicological assessment of the cyanobacterium *Cylindrospermopsis raciborskii* from Portuguese freshwaters. *Ecotoxicol. Environ. Saf.* 55:243-50.

Seenayya, G., & Raju, N. S. 1972. On the ecology and systematic position of the alga

known as *Anabaenopsis raciborskii* (Wolosz.) Elenkin and a critical evaluation of the forms described under the genus *Anabaenopis*. *In* T.V. Desikachary [Eds.] *International Symposium on Taxonomy and Biology of Blue-Green Algae.* University of Madras, Madras, India, pp.52-7

Shafik, H. M., Herodek, S., Presing, M., & Vörös, L. 2001. Factors effecting growth and cell composition of cyanoprokaryote *Cylindrospermopsis raciborskii* (Wołoszyńska) Seenayya et Subba Raju. *Algological Studies/ Arch. Hydrobiol. Suppl*. 103:75-93.

Silva, T.F.G., Vinçon-Leite, B., Lemaire, B.J., Petrucci, G., Giani, A., Figueredo, C.C., Nascimento, N.O. 2019. Impact of urban stormwater runoff on cyanobacteria dynamics in a tropical urban lake. *Water*. 11:946.

Sinha, R., Pearson, L. A., Davis, T. W., Burford, M. A., Orr, P. T., & Neilan, B. A. 2012. Increased incidence of *Cylindrospermopsis raciborskii* in temperate zones-is climate change responsible?. *Water Res*. 46:1408-19.

Sinha, R., Pearson, L. A., Davis, T. W., Muenchhoff, J., Pratama, R., Jex, A., Burford, M. A., & Neilan, B. A. 2014. Comparative genomics of *Cylindrospermopsis raciborskii* strains with differential toxicities. *BMC Genomics.* 15:1-14.

Sprober, P., Shafik, H. M., Présing, M., Kovács, A. W., & Herodek, S. 2003. Nitrogen uptake and fixation in the cyanobacterium *Cylindrospermopsis raciborskii* under different nitrogen conditions. *Hydrobiologia.* 506:169-74.

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. 2010. The smallest known genomes of multicellular and toxic cyanobacteria: comparison, minimal gene sets for linked traits and the evolutionary implications. *PLoS One*. *5*.

Stüken, A., Rücker, J., Endrulat, T., Preussel, K., Hemm, M., Nixdorf, B., Karsten, U., & Wiedner, C. 2006. Distribution of three alien cyanobacterial species (Nostocales) in northeast Germany: *Cylindrospermopsis raciborskii*, *Anabaena bergii* and *Aphanizomenon*

aphanizomenoides. *Phycologia*. 45:696-703.

Sukenik, A., Hadas, O., Kaplan, A., & Quesada, A. 2012. Invasion of Nostocales (cyanobacteria) to subtropical and temperate freshwater lakes-physiological, regional, and global driving forces. *Front. Microbiol.* 3:86.

Thiel, T., & Leone, M. 1986. Effect of glutamine on growth and heterocyst differentiation in the cyanobacterium *Anabaena variabilis*. *J. Bacteriol*, 168:769-74.

Turner, S., Pryer, K. M., Miao, V. P., & Palmer, J. D. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *J. Eukaryot. Microbiol.* 46:327-38.

Vico, P., Bonilla, S., Cremella, B., Aubriot, L., Iriarte, A., & Piccini, C. 2020. Biogeography of the cyanobacterium *Raphidiopsis* (*Cylindrospermopsis*) *raciborskii*: integrating genomics, phylogenetic and toxicity data. *Mol. Phylogenetics Evol.* 106824.

Vintila, S., & El-Shehawy, R. 2007. Ammonium ions inhibit nitrogen fixation but do not affect heterocyst frequency in the bloom-forming cyanobacterium *Nodularia spumigena* strain AV1. *Microbiology.* 153:3704-12.

Willis, A., Chuang, A. W., & Burford, M. A. 2016. Nitrogen fixation by the diazotroph *Cylindrospermopsis raciborskii* (Cyanophyceae). *Journal of Phycology*. 52:854-62.

Wilson, K. I. M. M., Schembri, M. A., Baker, P. D., & Saint, C. P. 2000. Molecular characterization of the toxic cyanobacterium *Cylindrospermopsis raciborskii* and design of a species-specific PCR. *Appl. Environ. Microbiol.* 66:332-8.

Wołoszynska, J. 1912. Das Phytoplankton einiger Javanian Seen mit berücksichtigung des Sawa-Planktons. *Bull. Int. Acad. Sci. Crac. Ser. B.* 6:649-709.

Wu, Z., Shi, J., & Li, R. 2009. Comparative studies on photosynthesis and phosphate metabolism of *Cylindrospermopsis raciborskii* with *Microcystis aeruginosa* and *Aphanizomenon flos-aquae*. *Harmful Algae*. 8:910-5.

Wu, Z., Shi, J., Xiao, P., Liu, Y., & Li, R. 2011. Phylogenetic analysis of two cyanobacterial genera *Cylindrospermopsis* and *Raphidiopsis* based on multi-gene sequences. *Harmful Algae*. 10:419-25.

CHAPTER TWO

EXPRESSION OF NITROGEN FIXATION GENES IN *RAPHIDIOPSIS* (=*CYLINDROSPERMOPSIS*) *RACIBORSKII*

ABSTRACT

Raphidiopsis raciborskii is an invasive, potentially toxic, N_2-fixing heterocytic cyanobacterium that gained attention because of its apparently recent spread around the world. Despite its ecological success, nitrogen fixation in terminal heterocytous cyanobacteria such as *Raphidiopsis* has remained largely unstudied until now. This study investigated the production of heterocytes in *R.raciborskii* and the gene expression of *nifH* and *hetR* in nitrogen depleted medium. The experiments comparing strains cultivated in complete mediun versus nitrogen-free medium showed that heterocyte development may not be controlled by nitrogen insufficiency and by using reverse transcription qPCR it was possible to understand that *R.raciborskii* can express nitrogenase even in a nitrogen rich medium, contrary to what occurs with other diazotrophic cyanobacteria

Key-words: diazothrophy, nitrate, gene expression, RT-qPCR, *Cylindrospermopsis*, cyanobacteria

1. INTRODUCTION

Nitrogen is a quantitatively important element very abundant as N_2 gas in the atmosphere. Some bacteria are capable of fixing N_2 into ammonia molecules (Kumar *et al.*, 2010). Nitrogen is incorporated into the biosphere through assimilatory processes carried out by microorganisms, algae and plants. Numerous nitrogen-containing compounds can be used by different organisms as nitrogen source. These include inorganic ions like nitrate or ammonium and simple organic compounds like urea, amino acids, and some nitrogen-containing bases. Microorganisms are able to control assimilation pathways of some nitrogen sources when some other, more easily assimilated nitrogen form is available to the cells (Herrero *et al.*, 2001).

Biological reduction of dinitrogen (N_2) to ammonia (NH_3) can only occur in a select group of Bacteria and Archea that contain the enzyme nitrogenase. Nitrogenase is an ATP-hydrolyzing, redoxactive complex of two component proteins, the dinitrogenase $\alpha_2\beta_2$ heterotetramer (where a ¼ NifD and b ¼ NifK proteins) and the dinitrogenase reductase γ_2 homodimer (NifH protein). To catalyze this difficult reaction, nitrogenase requires electrons, protons, and ATP.

Biological nitrogen fixation catalyzed by the Mo-dependent nitrogenase has a limiting reaction showed below (Burgess *et al.,* 1996; Seefeldt *et al.,* 2009):

$$N_2 + 8e^- + 16MgATP + 8H^+ \rightarrow 2NH_3 + H_2 + 16MgADP + 16Pi$$

The ATP-driven reduction of one N_2 with evolution of one H_2 requires a minimum of $8e^-$ and the hydrolysis of 16 ATP molecules.

Initial research of cyanobacteria in reservoirs stated that N_2 fixation was a competitive advantage for nitrogen-fixing species (Harris & Baxter, 1996). However, nitrogenase reduction of N_2 to NH_4^+ is an energy-expensive process (Duval *et al.,* 2013) reducing the available energy for growth and resulting in a reduction in growth rate. Actually, N_2 fixation

levels are insufficient in many freshwater systems to mitigate N deficiencies within ecosystems (Lewis & Wurtsbaugh, 2008; Scott & McCarthy, 2010; Scott & Grantz, 2013).

1.1. HOW DO HETEROCYTES WORK?

For biological nitrogen fixation, nitrogenase needs to be protected from oxygen. Some filamentous cyanobacteria restrict nitrogenase to heterocytes, differentiated cells that that establish micro-oxic conditions for optimum nitrogenase activity. This involves a loss of photosystem II and carbon-fixing operations, and the deposition of extra polysaccharide and glycolipid wall layers. (Flores & Herrero, 2010). They obtain photosynthate, most likely sucrose (Curatti *et al.*, 2002), from nearby vegetative cells and, in return, supply those cells with fixed nitrogen as amino acids.

Usually, like *Anabaena* PCC7120, cyanobacteria from the Nostocales order can develop high numbers of heterocytes, including evenly spaced intercalated heterocytes, supporting about 7-15 vegetative cells (Neunuebel & Golden, 2008; Corrales-Guerrero *et al.*, 2013). But, unlike many Nostocales, these heterocytes are always terminal in *Cylindrospermopsis raciborskii* (Plominsky *et al.*, 2013). This brings a challenge to the species, to transfer the N_2 fixed in its two terminal heterocytes to 50 or more associated vegetative cells.

The presence of heterocytes in *Raphidiopsis* stand out because this was the main morphological difference between *Cylindrospermopsis* and *Raphidiopsis* genera. The absence of heterocytes was a characteristic of *Raphidiopsis* and their presence of *Cylindrospermopsis* (Komárek & Anagnostidis, 1989). More recently. the taxonomic validity of both genera was discussed and based on phylogeny and secondary structure of the 16S-23S ribosomal RNA intergenic spacer region, Aguilera *et al*. (2018) proposed the unification of the two genera in the single genus *Raphidiopsis*.

1.2. NITROGEN SOURCES

According to Flores & Herrero (1994), ammonium and nitrate inhibits nitrogenase and heterocyte formation and directly inactivates nitrogenase through the modification of the nitrogenase-Fe protein. Energetically, ammonium and nitrate are more efficient sources for the growth of cyanobacteria than N_2. It is therefore not surprising that dissolved sources of nitrogen are preferably used when available.

1.3. MAJOR GENES IN THE NITROGEN FIXATION PROCESS

Nitrogen control in cyanobacteria is mediated by NtcA, a transcriptional regulator. This protein is required for gene expression in ammonium and nitrate assimilation pathways, as well for heterocyte development (Wei *et al.* 1993; Ramasubramanian *et al.*, 1994). The *ntcA* gene is induced soon after nitrogen deprivation and is autoregulated (Ramasubramanian *et al.*, 1994; Ramasubramanian *et al.,* 1996; Muro-Pastor *et al.,* 2002). Frias *et al*. (1994) and Wei *et al*. (1994) investigated *ntcA* insertional mutants (*ntcA* was inactivated) and found that they were unable to use nitrate as the sole source of nitrogen and also were blocked from initiating heterocyte development.

HetR is a master regulator of heterocyte development and plays a key role in differentiation and pattern formation (Buikema & Haselkorn, 1991). Null mutants of *hetR* fail to produce heterocytes and overexpression of *hetR* (Buikema & Haselkorn, 1991; Buikema & Haselkorn, 2001) results in increased heterocyte frequency. *hetR* is one of the earliest genes induced in differentiating cells.

The conventional nitrogenase enzyme is encoded by the *nifHDK* operon, which comprises genes in contiguous arrangement within the genome. The *nifH* gene encodes the dinitrogenase reductase component of the nitrogenase enzyme complex, and is traditionally used as a marker for the N_2-fixation process. In Nostocales, the heterocytes are the sole sites

for *nifH* expression and NifH biosynthesis (Elhai & Wolk, 1990; Plominsky *et al*., 2013).

1.4. A LITTLE ABOUT THE METHOD (RT-qPCR)

Reverse transcriptase (RT) technique followed by quantification using the polymerase chain reaction (qPCR) in real time or, simply reverse transcription-qPCR (RT-qPCR), is an in vitro method for enzymatic amplification of defined complementary DNA sequences obtained from RNA. The RT-qPCR technique is one of the most used in gene expression quantification experiments as it is considered one of the most sensitive and specific methods of gene quantification.

2. OBJECTIVES

Despite the ecological success of *R. raciborskii* (*C. raciborskii*) and the large number of publications on this species, nitrogen fixation in terminal heterocytous cyanobacteria has until now largely remained unstudied. Our main goal was to investigate the production of heterocytes and the expression of genes related to nitrogen metabolism and fixation (*ntcA*, *nifH* and *hetR*) in *Raphidiopsis raciborskii*, also using a gene related to phycobilin pigment production (*cpcβ*) as metabolic control.

3. METHODOLOGY

Initially, two *R. raciborskii* strains (UFMG40 and UFMG188) were grown in WC complete medium (Guillard, 1975) in light / dark regime 12h / 12h at 20° C. To prepare the cultures for the nitrogen reduction experiments, 150 mL of the cultures were carefully filtered, at low pressure, through a cellulose membrane to remove the N-rich medium. Next, the cells were suspended in 300 mL WC medium, with and without nitrogen, and were allowed to grow for 3 days. After the acclimatization period, 90 mL of each of these cultures was transfered into 150 mL of WC medium in triplicate, for 4 days growth, as follows:

- Control: Cultures grown in complete nitrogen-rich medium;

• Treatment: Cultures grown in a nitrogen-free medium.

At the end of the experiment, each culture was filtered through microfiber glass filters where: 2 filters of 60 mL were for chorophyll-α analysis and 4 filters of 20 mL were for gene expression analysis (RT-qPCR). Unfiltered 5 mL samples were preserved in lugol solution for quantification of vegetative cells and heterocytes on a Fuchs-Rosenthal counting chamber.

3.1. CHLOROPHYLL-α ANALYSIS

Analysis of chlorophyll-α was carried out using the method defined by Nusch (1980). The chlorophyll was extracted in duplicate from previously frozen filters. The quantification of chlorophyll-α was conducted in a spectrophotometer at wavelengths of 665 nm and 730 nm respectively. Readings were conducted at 730 nm to correct the turbidity.

3.2. RT-qPCR

The filters were submerged with TRIzol (Invitrogen), later they were quickly frozen with liquid nitrogen and triturated. After filter trituration, the RNA was extracted according to TRIzol manufacturer's recommendations. Total RNA was suspended in 30 μL of DEPC-H_2O, and RNA was treated with 1 U/g of DNase (Promega) at 37°C for 30 min. The reaction was stopped by the addition of a stop solution and by heating for 10 min at 65°C. Removal of DNA traces was confirmed by PCR.

RNA was quantified using a Qubit RNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, United States). About 100 ng of RNA was used for RT-PCR. The cDNA was generated using a High Capacity kit (Applied Biosystems) with RT random primers. Concentrations and PCR cycling conditions were established according to the manufacturer's recommendations. Real-time PCR was performed using a StepOne system (Applied Biosystems) with 1 μL of cDNA sample, 0.3 μL of each primer (10 pmol/ μL), 5 μL of Power SYBR green I (Applied Biosystems), and sterile Milli-Q water for a final volume of 10

μL. The reactions were done in duplicate, and specifications for the PCR cycle followed the manufacturer's guidelines. Primers used in this study are described in Table 1.

The new primers were designed with the help of Primer-BLAST tools (NCBI). The primer amplification efficiencies for 16S rRNA, *nifH*, *hetR*, *ntcA* and *cpcβ* genes were calculated, and they provided E values of 1.97, 2.13, 1.98, 1.96, 1.95 respectively, where an E value of 2 indicates 100% PCR efficiency.

Relative quantification of the *nifH*, *hetR*, *ntcA* and *cpcβ* target genes and 16S-rRNA gene, used as a reference gene, was represented as the change in transcription, which was calculated according with the formula $Log_2(2^{-\Delta\Delta Ct})$, and was compared to the results observed under the control conditions.

Table 1- Primers information.

Primer name	Sequence	Region	Amplicon size	Reference
16SF	AGAAAAGAGGTTTACGACCCAAGAGC	16S rRNA	267	Wu *et al.*, 2013
16SR	TGAAAGATTTATTGCCTGGAGATGAGC	16S rRNA	267	Wu *et al.*, 2013
NifHf	CGTAGGTTGCGACCCTAAGGCTGA	*nifH*	297	Gugger *et al.*, 2005
NifHr	GCATACATCGCCATCATTTCACC	*nifH*	297	Gugger *et al.*, 2005
hetR_F	GATGAAGGACTGGGCAGAAA	*hetR*	182	Willis *et al.*, 2016
hetR_R	CCTGTTGACCAAAAGCCATT	*hetR*	182	Willis *et al.*, 2016
ntcA_F	TGCGGTGGAATTGCTCTCTT	*ntcA*	113	This work
ntcA_R	CTGTTTGCAGAATCCGCGAG	*ntcA*	113	This work
PCB_F	TCCGTAGCTGTTGGTGTTGG	PC Subunit B	88	This work
PCB_R	GACTGCAATCACCCTTGGTA	PC Subunit B	88	This work

3.3. STATISTICAL ANALYSES

Statistical analyses were performed by using R software. Shapiro-Wilk normality test and Levene's Test for Homogeneity of Variance were applied. ANOVA was run on a Gaussian (normal) distribution and independent t-test was used to test for difference among treatments. Significant difference was accepted for $P \leq 0.05$. Correlation analyses were run between ΔCt values of *nifH* and *hetR*, and *hetR* and *ntcA* genes.

4. RESULTS

The treatment with the complete medium was always used as a reference for the quantification of the gene expression of the treatment with nitrogen depletion.

The gene expression between control and treatment for both strains was significantly different ($F_{(1,11)}=89.48$; P-value=$1.285e^{-06}$). In both strains, the number of heterocytes and the expression of *hetR*, *nifH* and *ntcA* genes increased with nitrogen depletion (Fig. 1).

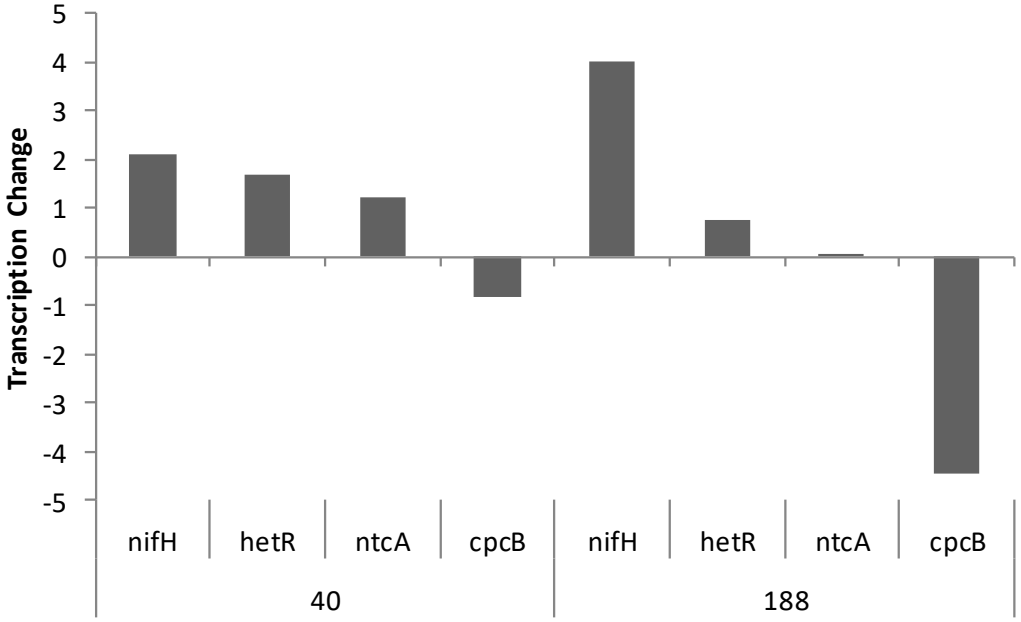


Fig. 1- Fold change relative quantification of *nifH*, *hetR*, *ntcA* and *cpcβ* gene expression for two cyanobacterial strains (UFMG-40 and 188). The treatment with nitrate was used as a standard (control) to estimate changes in gene expression under depletion of nitrate.

At the same time, it was observed that in the nitrogen depletion conditions the strains became yellowish (Fig. 2), and this was reflected in lower chlorophyll α concentration (Fig. 3), a lower cell numbers and a decrease in the expression of the *cpcβ* gene associated with the production of the phycocyanin pigment (Fig. 1). Discarded heterocytes were absent.

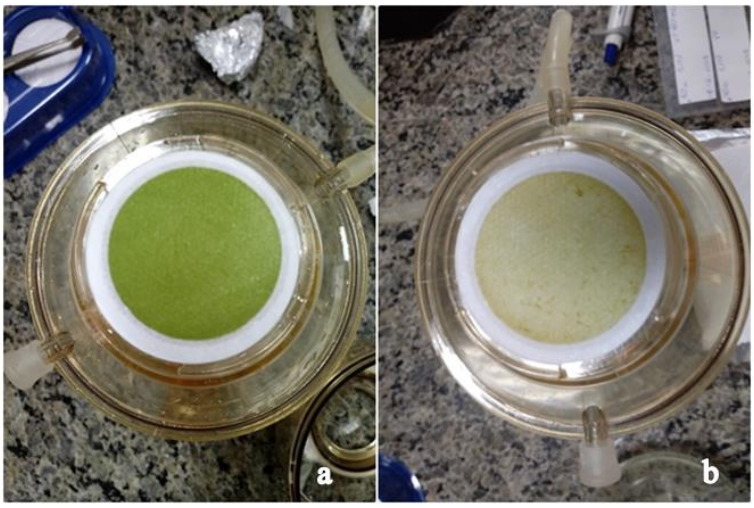


Fig.2- Filters from Control (a) and Treatment under nitrate depletion (b).

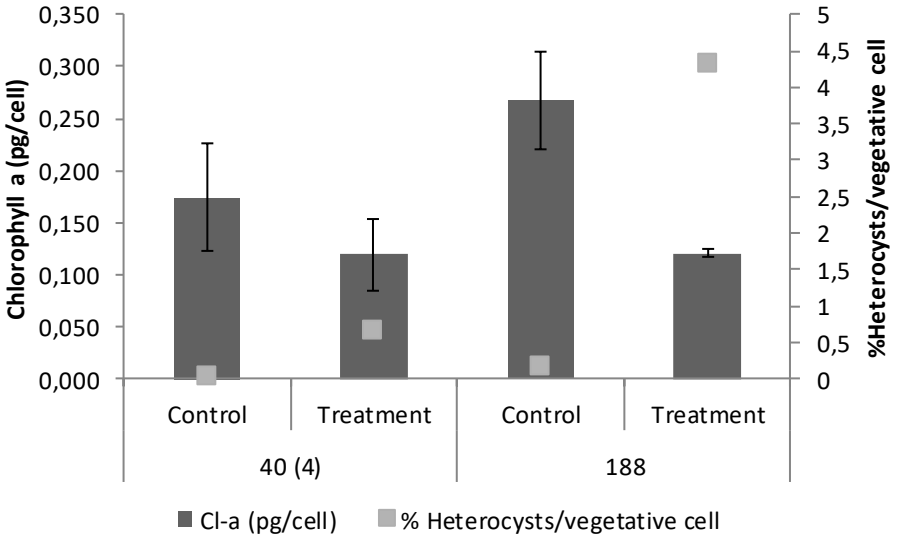
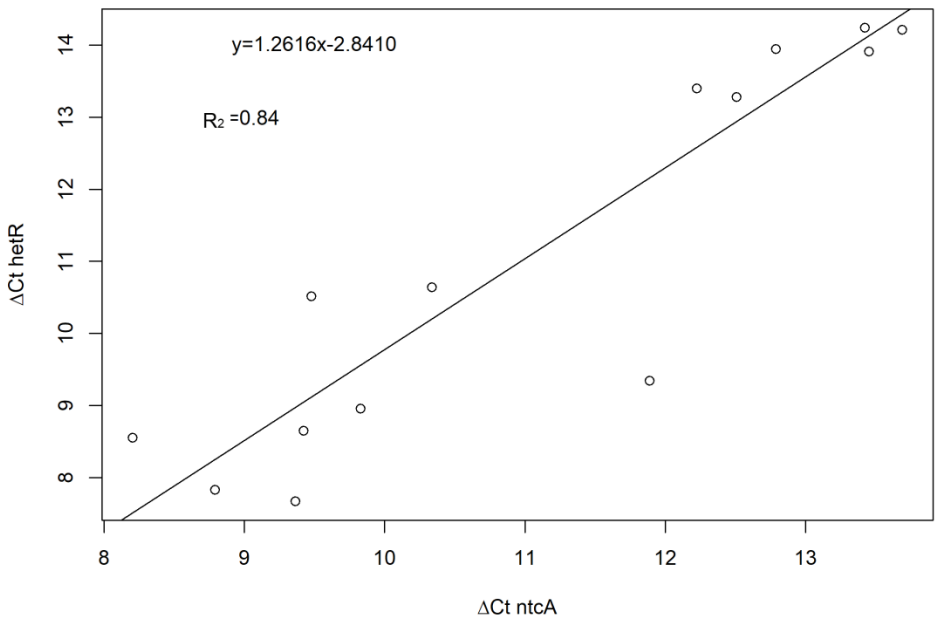


Fig.3- Chlorophyll α and heterocytes percentage per total of vegetative cells for each strain (UFMG40 and 188) at two nutrient concentrations. Bars represent standard error of the means. The primary axis is represented by dark gray bars. The secondary axis is represented with lighter gray dots.

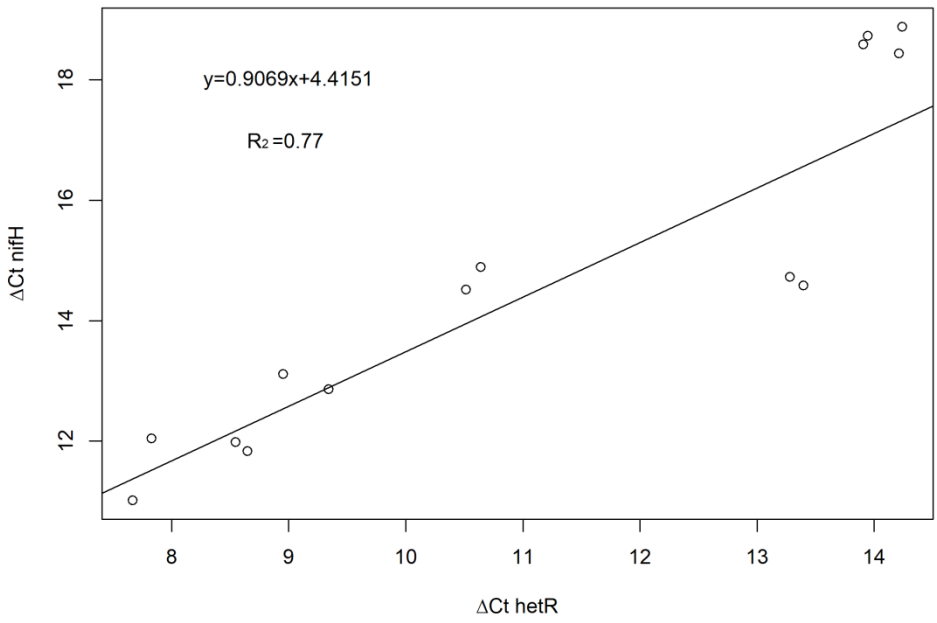
There was a decrease in chlorophyll α content and an increase in the production of heterocytes in the treatment compared to the control, for both strains. However, for strain UFMG40 the difference was not significant for chlorophyll concentration (t = 0.3397, df = 4, p-value = 0.3756), but it was for heterocytes (t = -3.9105, df = 2.0094, p-value = 0.02958). For strain UFMG188, there was a significant difference between treatments for chlorophyll (t = 5.4757, df = 2.0265, p-value = 0.01543) and heterocytes (t = -5.56, df = 2.0236, p-value = 0.01503).

Linear regression analysis was performed to find out if there was a correlation between ΔCt *hetR* and ΔCt *ntcA* (Fig. 4) and between ΔCt *nifH* and ΔCt *hetR* (Fig. 5). The expression of *hetR* increased with *ntcA* expression.And *nifH* expression increased with *hetR* expression.



1

2 Fig.4- Correlation analysis between ΔCt *hetR* and ΔCt *ntcA*.



3

4 Fig.5- Correlation analysis between ΔCt *nifH* and ΔCt *hetR*.

5. DISCUSSION

The heterocyte differentiation process induced by nitrogen depletion is known to have multiple steps (Nierzwicki-Bauer *et al.*, 1984). The first event is perception of nitrogen stress, that can induce uptake of alternative sources of nitrogen in cells, which leads to an increase in general proteolysis and, in particular, to degradation of the phycobiliproteins, photosynthetic accessory pigments that may account for up to 50% of the cellular protein. This is perceived as chlorosis of the cell, the loss of pigmentation by discoloration is an important sign of the occurrence of oxidative stress. This is probably what happened with our cultures in nitrogen depletion as we detected the decrease of *cpcβ* gene expression (phycocyanin-related gene), of chlorophyll α concentration, of number of vegetative cells, and loss of culture pigmentation.

Different strains of *Raphidiopsis raciborskii* showed different levels of expression of the *nifH* gene and *hetR*. These genes were expressed even in the control treatment, under rich nitrogen environment, and not only when nitrogen was depleated. It is known that the presence of heterocytes is not always indicative of nitrogen fixation. Haselkorn *et al*. (1998) observed that overexpression of *hetR* results in heterocyte development independent of the nitrogen source, but only in a medium without combined nitrogen the heterocytes can express nitrogenase. And according to Flores *et al.* (1999) and Huang *et al*. (1999) the *nif* operon and other nitrogen fixation-related genes are not expressed in cultures supplemented with combined nitrogen. This is contrary to our findings, in which *nifH* was expressed even when in nitrate rich medium. The *nifH, ntcA* and *hetR* genes were expressed in both conditions, but expression was higher in the nitrogen-free treatment.

In earlier studies, it was believed that diazotrophic bacteria are generally subjected to immediate inhibition of nitrogen fixation activity or repression of nitrogenase synthesis in response to external ammonium or nitrate ions (Hartmann *et al*., 1986, Streeter and Wong, 1988). Seldin *et al*. (1984) also found that species such as *Paenibacillus macerans* and

Paenibacillus polymyxa can totally inhibit the nitrogen fixation activity in the presence of 0.5 % nitrate. More recently, Vintila & El-Shehawy (2007) found that *Nodularia spumigena* strain AV1 maintains heterocyte frequency in the presence of ammonium, but ammonium suppressed *nifH* expression. Martin-Neto *et al*. (1991) found that nitrate is an effective inhibitor of both nitrogenase synthesis and heterocyte development in *Anabaena* sp. PCC 7120. In addition, Elhai & Wolk (1990) and Holland & Wolk (1990) reported that nitrate was used as a source of combined nitrogen able to keep repressed *nifHDK* and *hetA* expression. In accordance with our findings Sprober *et al*. (2003) reported in their experiments that ammonium did not repress N_2 fixation of *C. raciborskii* and Kenesi *et al.* (2008) found that N_2 fixation was completely inhibited by ammonium and partially inhibited by nitrate.

It appears to us that the inhibition of nitrogenase synthesis by nitrate depends on the strain/species studied. In *R. raciborskii* (*C.raciborskii*) strains nitrogenase enzyme system can still function within a nitrogen-rich medium (Sprober *et al*., 2003; Kenesi *et al.*, 2008; this study). This species may have a level of nitrate reductase that is not high enough to inhibit nitrogen fixation, as nitrate reductase activity is required to inhibit nitrogenase synthesis (Martin-Nieto *et al*., 1991).

Our study shows that *R.raciborskii* heterocytes can be formed in nitrogen rich medium and can also express nitrogenase, contrary to what occurs with other diazotrophic cyanobacteria. As expected, the expression of *hetR* was linked to the expression of *ntcA*, since both are autoregulated and mutually dependent (Muro-Pastor *et al*., 2002; Herrero *et al*., 2004) (Fig. 6). And the expression of *hetR* was positively correlated to the expression of *nifH*, probably because the heterocyte is the site of expression of the *nifHDK* operon. After *hetR* is expressed and the heterocyte is formed, nitrogenase is also produced, i.e., *nifH* is expressed. Previous studies observed that the *nifHDK* operon is not expressed in *ntcA* (Frías *et al*., 1994) or *hetR* (Valladares *et al*., 1999) mutants.

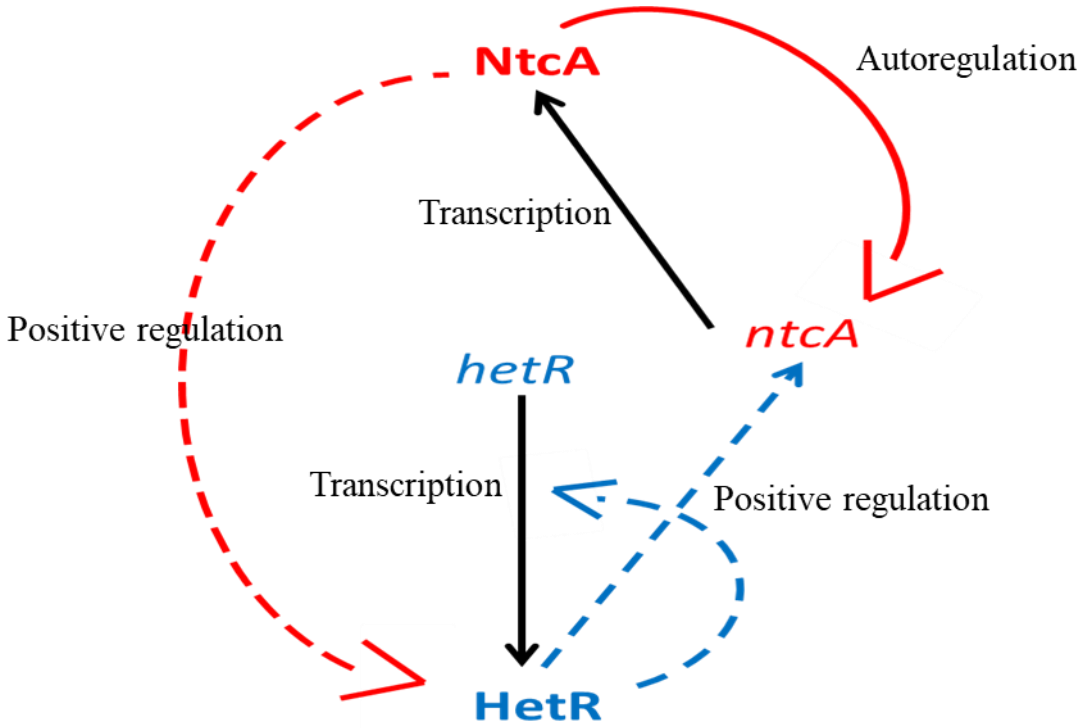


Fig.6- Activation pathway of gene expression at the beginning of heterocyte differentiation. Black arrows represent transcription to the corresponding mature protein. Red solid arrows indicate NtcA promoted transcription activation. Dashed arrows indicate a positive action exerted by NtcA (red) or HetR (blue) on gene expression. (Image based on Herrero *et al.*, 2004)

6. CONCLUSION

This study shows that the effect of nitrate may be variable and dependent on the strain studied, that heterocyte development may not be controlled by nitrogen insufficiency and that *R.raciborskii* can express nitrogenase even in nitrogen rich medium, contrary to what occurs with other diazotrophic cyanobacteria.

7. ACKNOWLEDGMENTS

We thank CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) for a scholarship to L.C.B.M and FAPEMIG (Fundação de Apoio a Pesquisa de Minas Gerais) for a research grant to A.G.

8. REFERENCES

Aguilera, A., Gómez, E. B., Kaštovský, J., Echenique, R. O., & Salerno, G. L. 2018. The polyphasic analysis of two native *Raphidiopsis* isolates supports the unification of the genera *Raphidiopsis* and *Cylindrospermopsis* (Nostocales, Cyanobacteria). *Phycologia.* 57:130-146.

Buikema, W. J., & Haselkorn, R. 1991. Characterization of a gene controlling heterocyst differentiation in the cyanobacterium *Anabaena* 7120. *Genes & development*, 5:321-330.

Buikema, W. J., & Haselkorn, R. 2001. Expression of the Anabaena *hetR* gene from a copper-regulated promoter leads to heterocyst differentiation under repressing conditions. *Proceedings of the National Academy of Sciences*, 98:2729-2734.

Burgess ,B. K., Lowe, D. J. 1996. Mechanism of molybdenum nitrogenase. *Chem Rev*, 96:2983-3012.

Corrales-Guerrero, L., Mariscal, V., Flores, E., & Herrero, A. 2013. Functional dissection and evidence for intercellular transfer of the heterocyst-differentiation PatS morphogen. *Molecular microbiology*, 88:1093-1105.

Curatti, L., Flores, E., Salerno, G. 2002. Sucrose is involved in the diazotrophic metabolism of the heterocyst- forming cyanobacterium Anabaena sp. *FEBS Lett*, 513:175-178.

Duval, S., Danyal, K., Shaw, S., Lytle, A. K., Dean, D. R., Hoffman, B. M., Antony, E., & Seefeldt, L. C. 2013. Electron transfer precedes ATP hydrolysis during nitrogenase catalysis. *Proceedings of the National Academy of Sciences of the United States of America*, 110:16414-16419.

Elhai, J., Wolk, C. P. 1990. Developmental regulation and spatial pattern of expression of the structural genes for nitrogenase in the cyanobacterium *Anabaena*. *EMBO*, J 9:3379-

3388

Flores, E., & Herrero, A. 1994. Assimilatory nitrogen metabolism and its regulation. *In* The molecular biology of cyanobacteria. Springer, Dordrecht, pp. 487-517.

Flores, E., & Herrero, A. 2010. Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nature Reviews Microbiology*, 8:39-50.

Flores, E., Muro-Pastor, A. M. & Herrero, A. 1999. Cyanobacterial nitrogen assimilation genes and NtcA-dependent control of gene expression. *In* G. A. Peschek, W. Loffelhardt, and G. Schmetterer [EDs.] *The phototrophic prokaryotes*. Plenum Publishing Corporation, New York, NY, pp. 463-477.

Frías, J. E., Flores, E. & Herrero, A. 1994. Requirement of the regulatory protein NtcA for the expression of nitrogen assimilation and heterocyst development genes in the cyanobacterium *Anabaena* sp. PCC 7120. *Mol. Microbiol*. 14:823-832.

Gugger, M., Molica, R., Le Berre, B., Dufour, P., Bernard, C., & Humbert, J. F. 2005. Genetic Diversity of *Cylindrospermopsis* strains (Cyanobacteria) Isolated from Four Continents. *Applied and Environmental Microbiology*, 46:515-521.

Guillard, R. R. L. 1975. Culture of phytoplankton for feeding marine invertebrates. *In* Smith, W.L. & Chanley, M.H. [Eds.] *Culture of marine invertebrate animals*. Springer, Boston, MA, pp. 29-60.

Harris, G. P., & Baxter, G. 1996. Interannual variability in phytoplankton biomass and species composition in a subtropical reservoir. *Freshwater biology*, 35:545-560.

Hartmann, A., Fu, H., & Burris, R. H. 1986. Regulation of nitrogenase activity by ammonium chloride in *Azospirillum* spp. *Journal of Bacteriology*, 165:864-870.

Haselkorn, R., Schlictman, D., Jones, K., & Buikema, W. J. 1998. Heterocyst differentiation and nitrogen fixation in cyanobacteria. *In* Elmerich, C., Kondorosi, A., Newton, W.E. [Eds.] *Biological Nitrogen Fixation for the 21st Century*. Springer, Dordrecht,

pp. 93-6.

Herrero, A., Muro-Pastor, A. M., & Flores, E. 2001. Nitrogen control in cyanobacteria. *Journal of Bacteriology*, 183:411-425.

Herrero, A., Muro-Pastor, A. M., Valladares, A., & Flores, E. 2004. Cellular differentiation and the NtcA transcription factor in filamentous cyanobacteria. *FEMS Microbiology Reviews*, 28:469-487.

Holland, D., Wolk, C. P. 1990. Identification and characterization of *hetA*, a gene that acts early in the process of morphological differentiation of heterocysts. *J Bacteriol*, 172:3131-3137.

Huang, T. C., Lin, R. F., Chu, M. K., & Chen, H. M. 1999. Organization and expression of nitrogen-fixation genes in the aerobic nitrogen-fixing unicellular cyanobacterium *Synechococcus* sp. strain RF-1. *Microbiology*, 145:743-753.

Kenesi, G., Shafik, H. M., Kovács, A. W., Herodek, S., & Présing, M. 2009. Effect of nitrogen forms on growth, cell composition and N2 fixation of *Cylindrospermopsis raciborskii* in phosphorus-limited chemostat cultures. *Hydrobiologia*, 623:191-202.

Komarek, J., & Anagnostidis, K. 1989. Modern approach to the classification system of Cyanophytes 4-Nostocales. *Archiv für Hydrobiologie. Supplementband. Monographische Beiträge*, 82:247-345.

Kumar, K., Mella-Herrera, R. A., & Golden, J. W. 2010. Cyanobacterial heterocysts. *Cold Spring Harbor Perspectives in Biology*, 2:1-20.

Lewis Jr, W. M., & Wurtsbaugh, W. A. 2008. Control of lacustrine phytoplankton by nutrients: erosion of the phosphorus paradigm. *International Review of Hydrobiology*, 93:446-465.

Martin-Nieto, J., Herrero, A., & Flores, E. 1991. Control of nitrogenase mRNA levels by products of nitrate assimilation in the cyanobacterium *Anabaena* sp. strain PCC 7120.

Plant physiology, 97:825-828.

Muro-Pastor, A. M., Valladares, A., Flores, E., & Herrero, A. 2002. Mutual dependence of the expression of the cell differentiation regulatory protein HetR and the global nitrogen regulator NtcA during heterocyst development. *Molecular microbiology*, 44:1377-1385.

Neunuebel, M. R., & Golden, J. W. 2008. The *Anabaena* sp. strain PCC 7120 gene all2874 encodes a diguanylate cyclase and is required for normal heterocyst development under high-light growth conditions. *Journal of bacteriology*, 190:6829-6836.

Nierzwicki-Bauer, S. A., Balkwill, D. L., & Stevens, S. E. 1984. Heterocyst differentiation in the cyanobacterium Mastigocladus laminosus. *Journal of bacteriology*, 157:514-525.

Nusch, E. A. 1980. Comparison of different methods for chlorophyll and pheopigment determination. *Archiv für Hydrobiologie-Beiheft Ergebnisse der Limnologie*,14:14-36

Plominsky, Á. M., Larsson, J., Bergman, B., Delherbe, N., Osses, I., & Vásquez, M. 2013. Dinitrogen fixation is restricted to the terminal heterocysts in the invasive cyanobacterium *Cylindrospermopsis raciborskii* CS-505. *PloS one*, 8:e51682.

Ramasubramanian, T. S., Wei, T. F., & Golden, J. W. 1994. Two *Anabaena* sp. strain PCC 7120 DNA-binding factors interact with vegetative cell-and heterocyst-specific genes. *Journal of Bacteriology*, 176:1214-1223.

Ramasubramanian, T. S., Wei, T. F., Oldham, A. K., & Golden, J. W. 1996. Transcription of the *Anabaena* sp. strain PCC 7120 *ntcA* gene: multiple transcripts and NtcA binding. *Journal of bacteriology*, 178:922-926.

Scott, J. T., & Grantz, E. M. 2013. N_2 fixation exceeds internal nitrogen loading as a phytoplankton nutrient source in perpetually nitrogen-limited reservoirs. *Freshwater Science*, 32:849-861.

Scott, J. T., & McCarthy, M. J. 2010. Nitrogen fixation may not balance the nitrogen pool in lakes over timescales relevant to eutrophication management. *Limnology and Oceanography*, 55:1265-1270.

Seefeldt, L. C., Hoffman, B. M., Dean, D. R. 2009. Mechanism of Mo-dependent nitrogenase. *Annu Rev Biochem* 78:701-722.

Seldin, L., Van Elsas, J. & Penido, E. 1984. *Bacillus azotofixans* sp. nov., a nitrogen fixing species from Brazilian soils and grass roots. *International Journal of Systematic and Evolutionary Microbiology*, 34:451-456.

Sprober, P., Shafik, H. M., Présing, M., Kovács, A. W., & Herodek, S. 2003. Nitrogen uptake and fixation in the cyanobacterium *Cylindrospermopsis raciborskii* under different nitrogen conditions. *Hydrobiologia*. 506:169-74.

Streeter, J. & Wong, P. P. 1988. Inhibition of legume nodule formation and N2 fixation by nitrate. *Critical Reviews in Plant Sciences*, 7:1-23.

Valladares, A., Muro-Pastor, A. M., Fillat, M. F., Herrero, A., & Flores, E. 1999. Constitutive and nitrogen-regulated promoters of the petH gene encoding ferredoxin:NADP+ reductase in the heterocyst-forming cyanobacterium *Anabaena* sp. *FEBS Letters*, 449:159-164.

Vintila, S., & El-Shehawy, R. 2007. Ammonium ions inhibit nitrogen fixation but do not affect heterocyst frequency in the bloom-forming cyanobacterium *Nodularia spumigena* strain AV1. *Microbiology*, 153:3704-3712.

Wei, T. F., Ramasubramanian, T. S., & Golden, J. W. 1994. *Anabaena* sp. strain PCC 7120 *ntcA* gene required for growth on nitrate and heterocyst development. *Journal of bacteriology*, 176:4473-4482.

Wei, T. F., Ramasubramanian, T. S., Pu, F., & Golden, J. W. 1993. *Anabaena* sp. strain PCC 7120 *bifA* gene encoding a sequence-specific DNA-binding protein cloned by in

vivo transcriptional interference selection. *Journal of bacteriology*, 175:4025-4035.

Willis, A., Chuang, A. W., & Burford, M. A. 2016. Nitrogen fixation by the diazotroph *Cylindrospermopsis raciborskii* (Cyanophyceae). *Journal of Phycology*, 52: 854-862.

Wu, Z., Shi, J., & Yang, S. 2013. The effect of pyrogallic acid on growth, oxidative stress, and gene expression in *Cylindrospermopsis raciborskii* (Cyanobacteria). *Ecotoxicology*, 22:271-278.

CHAPTER THREE

BRIEF REVIEW ON EVOLUTION OF NITROGENASE GENES

1. ABSTRACT

The evolution of the nitrogen-fixing enzyme nitrogenase represented a major breakthrough in the radiation of life. However, information about the nitrogenase enzyme complex evolution is still inconclusive. There are three types of nitrogenase, a molybdenum (Mo) nitrogenase (Nif), a vanadium (V) nitrogenase (Vnf) and a nitrogenase-3 (Anf) that lacks molybdenum and vanadium. The alternative nitrogenases (Anf and Vnf) could be derived paralogs to the most common Nif nitrogenase or they may instead represent early nitrogenases. Our understanding of the nitrogenase system and the evolution of the *nif* genes relies on knowledge of the genomic events that produced and stabilized the genetic mechanisms and structure of the operon. In this study we sought to review the evolutionary route of the *nif* genes and nitrogenase enzyme. A literature survey was carried out together with an analysis of phylogenetic relationships between the Nif family proteins, responsible for nitrogen fixation. It was possible to observe that several genetic duplications occurred in order to form the *nif* family as we see it today. A notable duplication is that of the *nifDK* operon that generated the *nifE* and *nifN* genes. Nif nitrogenase is probably the ancestor of Vnf and Anf nitrogenases, although it is still uncertain if there was an even erlier nitrogenase than Nif.

Key-words: Mo-Fe nitrogenase, vanadium, nitrogenase, evolution

Notice: this chapter is not intended to be conclusive, but a brief review of *nif* family genes evolution.

2. INTRODUCTION

One of the key nutrients essential for the survival of all living species is nitrogen, more specifically bioavailable nitrogen, also called fixed nitrogen. It is a necessary component of many biomolecules, including proteins, DNA, and chlorophyll. In the early Earth, abiotic processes like lightning discharges could have converted atmospheric N_2 into assimilable NH_4^+, HCN, or NOx species (Yung & McElroy, 1979). These sources were, however, scarce, and this dramatically restricted the size of the biosphere. Decreasing CO_2 concentrations and at the same time the decrease in abiotic N_2 reduction to nitrous oxide led to a nitrogen crisis at approximately 2.2 billion years ago (Navarro-González *et al*., 2001). This may have triggered the innovation of biological N_2 fixation. The evolution of the nitrogen-fixing enzyme nitrogenase, which reduces atmospheric N_2 to NH_4^+, represented a major breakthrough in the radiation of life. The ability to fix nitrogen, also called diazotrophy, is exclusively distributed in the bacterial and archaeal domains.

The proper function of nitrogenase and nitrogen-fixation is related to a diverse group of proteins encoded by *nif* genes. The *nifHDK* operon encodes the subunits of nitrogenase. These subunits are: dinitrogenase reductase (iron protein) and dinitrogenase (molybdenum-iron protein). Adjacent to the *nifH*, *nifD*, and *nifK* genes in most *nif* operons are the *nifB*, *nifE*, *nifN*, *nifS* and *nifU* genes like in Fig.1.

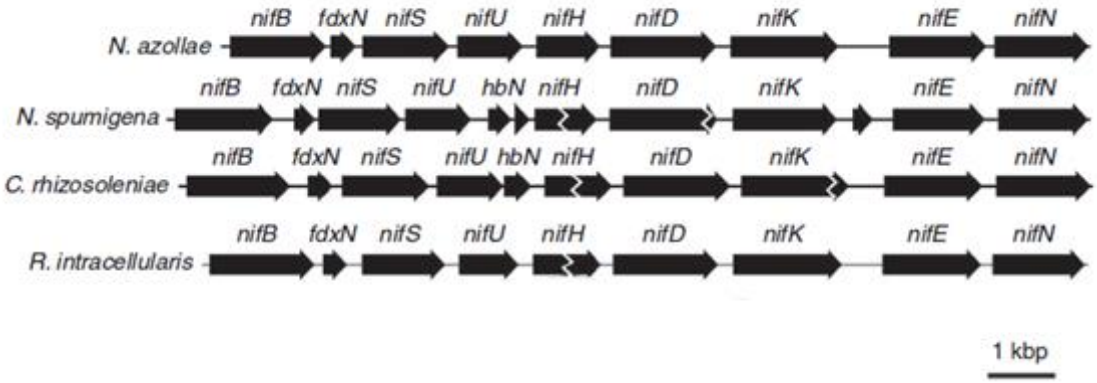


Fig.1- Image adapted from Hilton *et al*. (2013). The *nif* operon arrangement and surrounding genes in four strains: *Nostoc azollae* 0708, *Nodularia spumigena* CCY9414, *Calothrix rhizosoleniae* SC01 and *Richelia intracellularis* HH01.

The *nifH* gene encodes dinitrogenase reductase which is composed of two identical subunits (Mevarech *et al*., 1980), while *nifD* and *nifK* encode dinitrogenase which is a tetramer composed of two subunits that are α-subunit (NifD) and β-subunit (NifK). Dinitrogenase reductase (NifH) contains an iron-sulfur (4F-4S) cofactor, which binds the subunits of nitrogenase, and is responsible for binding atmospheric dinitrogen (N_2) and mediating the ATP-dependent transfer of electrons to the dinitrogenase tetramer. NifS transfers sulfur from cysteine to NifU, which interacts with multiple members of the signaling pathway (acts as a scaffolding protein) for (Fe-S) cluster assembly. The (Fe-S) clusters are transferred to NifB to make NifB-co, a (Fe6-S9) cluster that serves as the precursor to FeMo-cofactor. NifE and NifN, function as a scaffold for FeMo-cofactor assembly.

There are three types of nitrogenase, a molybdenum (Mo) nitrogenase (Nif), a vanadium (V) nitrogenase (Vnf) and a nitrogenase-3 (Anf) that lacks molybdenum and vanadium (Dos Santos *et al*., 2012). The alternative nitrogenases are expressed only when Mo concentrations are limiting, the Mo nitrogenase has been found to be more specific for and more efficient in binding dinitrogen and reducing it to ammonia than either of the alternative nitrogenase. The Vnf nitrogenase is expressed when V is present but Mo is not, and if neither

is, then the Anf nitrogenase is expressed if the organism possesses all three types of nitrogenase. In other words, the order of preferential expression is the Nif nitrogenase, followed by Vnf and at last Anf (Joerger & Bishop 1988).

Although these alternative nitrogenases could be derived paralogs of the Mo-requiring enzyme, they may instead represent early nitrogenases that have been maintained in several diverse lineages of prokaryotes (Anbar & Knoll 2002).

Our understanding of the nitrogenase system relies on knowledge of the genomic events that produced and stabilized the genetic mechanisms and structure of the operon. In this study we seek to review the evolutionary route of the *nif* genes.

3. METHODOLOGY

The NCBI databank was used to obtain the Nif protein sequences of *Nostoc* sp. PCC 7120 (model organism for nitrogen fixation studies) and the annotations of the 28 genomes used in this study (Table 1). With these data, a local Blastp of the Nif sequences of *Nostoc* sp. was performed against the annotations of the chosen genomes. Among these sequences, all those smaller than 80% of the size of the smallest Nif protein of *Nostoc* sp. PCC 7120 were eliminated, which contained 299 aa, in order to remove non-alignable sequences such as pseudogenes or incomplete sequenced genes. Subsequently, sequences that contained 100% similarity were also eliminated manually, always maintaining the largest sequence. Each sequence had their conserved domains verified manually using the NCBI conserved domains plataform.

To compare the gene evolution with the species evolution we used the 16S rRNA region of the 28 genomes and as outgroup the Archeae *Methanococcus voltae* PS strain.

The online platform program Mafft version 7.0 (Katoh *et al*., 2019) was used to build the alignment, under the following parameters: –thread 10 –threadtb 5 –threadit 0 –reorder –

leavegappyregion –auto. The phylogenetic tree of Maximum likelyhood was built using the PhyML 3.0 software (http://www.atgc-montpellier.fr/phyml/) with default settings. The best fitted model was chosen using the Akaike Information Criterion (AIC) witch was LG+G+F for the Nif proteins and GTR+G+I for the 16S rRNA region.

4. RESULTS

4.1. 16S PHYLOGENY

16S-rRNA maximum likelihood tree showed that diazotrophy is a characteristic that is spread across the tree (Fig.2). *Methanococcus* an Archea capable of fixing nitrogen was used as an outgroup. The Nostocales cyanobacteria are all diazotrophyc and grouped together. Non-diazotrophyc Oscilatorialles grouped together, with the exception of *Crinalium epipsammum* PCC 9333 that exhibited a close relationship to Nostocales. And the only diazotrophyc Oscillatoriales *Cyanothece* sp. PCC7425 grouped together with Synechococcales. Synechococcales did not form a single group but the species were closely related to each other. *Pleurocapsa* a Pleurocapsales and the Chroococcales *Microcystis* and *Crocosphaera*, grouped together. *Gloeobacter*, as expected, positioned as a sister group of all other cyanobacteria. Proteobacteria formed a separated group from the rest. This 16S-rRNA tree was used as the basis to analyze the Nif family tree.

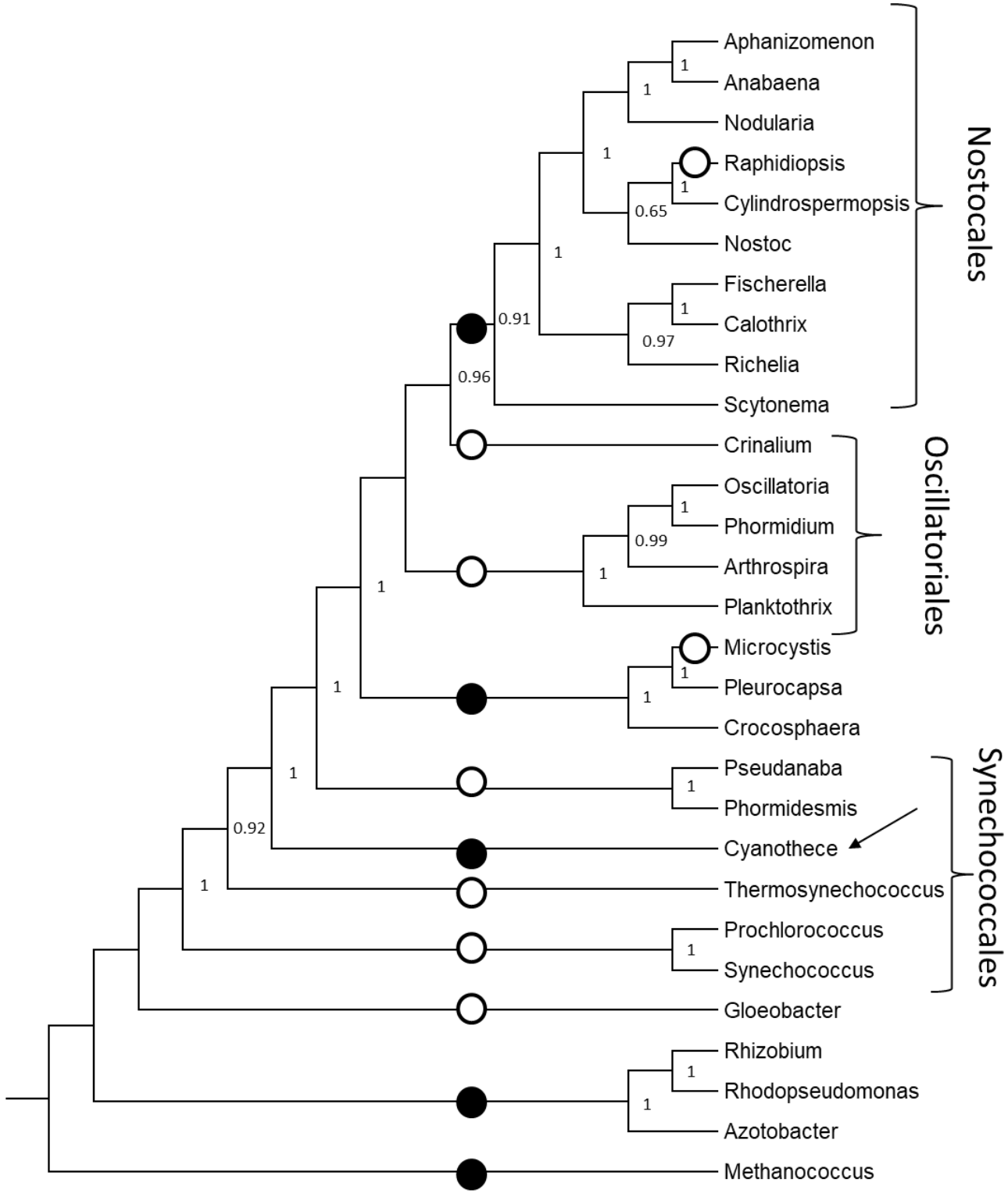


Fig.2- Maximum likelihood phylogenetic tree using the 16S-rRNA region. Nodules indicate branch support. Empty circles indicate non-fixing strains and filled circles indicate diazotrophic strains. Curly brackets indicate orders. The arrow indicates an Oscillatoriales genus out of place. See text for detailed explanation.

4.2. NIF FAMILY PHYLOGENY

Blastp returned 320 sequences as seen in table 1 and their accession number can be seen in table S1.

Table1- Strains that presented *nif* family genes and their respective number of sequences returned after blastp.

Strain	n° of sequences returned
Anabaena sp. 90	14
Aphanizomenon flos-aquae NIES81	13
Calothrix sp. PCC 6303	13
Cylindrospermopsis raciborskii CS-505	13
Nodularia spumigena CCY9414	15
Nostoc punctiforme PCC73102	18
Raphidiopsis brooki D9	5
Richelia intracellularis HH01	11
Scytonema tolypothrichoides VB-61278	16
Fischerella thermalis PCC7521	15
Crinalium epipsammum PCC 9333	5
Oscillatoria acuminata PCC 6304	7
Arthrospira platensis NIES39	7
Phormidium tenue NIES30	5
Planktothrix agardhii NIVA-CYA 1268	4
Cyanothece sp. PCC7425	15
Pseudanabaena biceps PCC 7429	5
Phormidesmis priestleyi BC1401	7
Thermosynechococcus elongatus BP1	5
Prochlorococcus marinus CCMP1375	6
Synechococcus elongatus PCC 6301	5
Microcystis aeruginosa NIES843	8
Crocosphaera watsonii WH8501	12
Pleurocapsa sp. PCC 7327	23
Gloeobacter violaceus PCC 7421	5
Azobacter vinelandii CA	23
Rhizobium leguminosarum WSM1689	9
Rhodopseudomonas palustris CGA009	36

Nif phylogenetic tree showed 4 main clades (Fig.3, Fig. S1). Clade I reunited NifK, VnfK, AnfK, NifN, VnfN, NifD, VnfD, AnfD, NifE and VnfE proteins. The NifK, VnfK, AnfK grouped together but VnfK and AnfK were external to NifK. NifN formed a second group and *Rhodopseudomonas* VnfN proteins became a third group. NifD, VnfD and AnfD

grouped together. NifN proteins formed another group. *Rhodopseudomonas* VnfE proteins were outside to all these groups.

Clade II reunited NifU, MoaA and NifB proteins. NifU proteins grouped together and a second group was formed by MoaA and NifB proteins.

Clade III consists of a NifH group and a chlL group of proteins.

Clade IV consists of NifS, IscS, SufS and CsdA proteins, all of them are cysteine desulfurases /aminotransferases involved in the iron-sulfur (FeS) cluster biosynthesis. NifS and IscS grouped together, SufS formed a group with a few CsdA sequences and the other CsdA sequences formed a separate group.

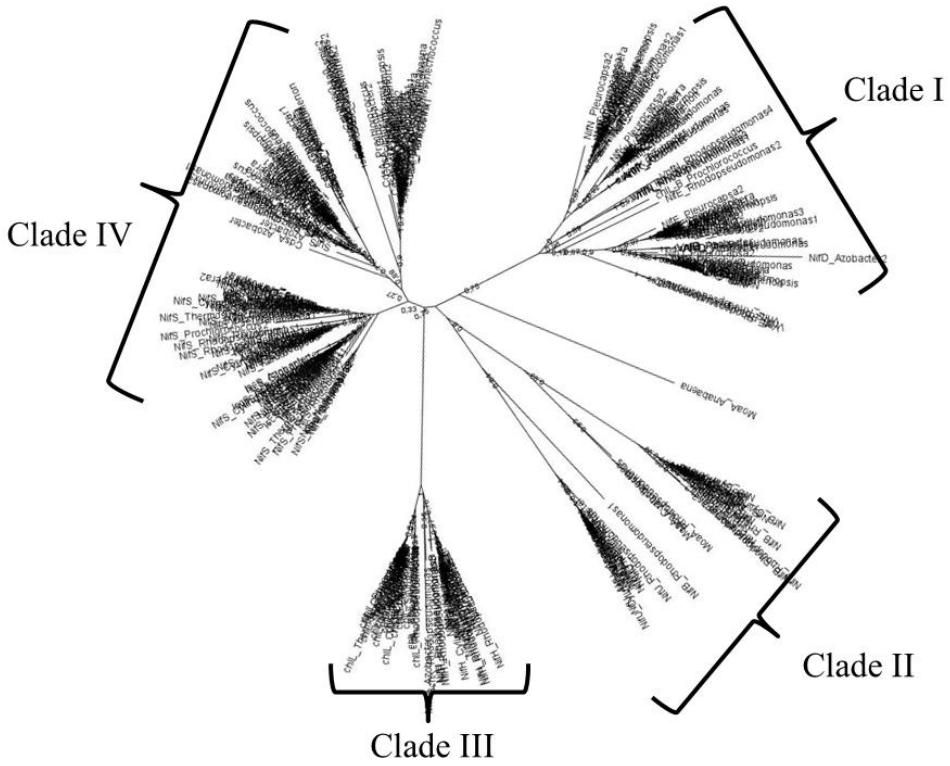


Fig.3- Phylogenetic relationships of bacterial Nif. The radiation tree containing 320 protein sequences is subdivided into four major clades. To better visualize this tree and the clades' content, see supplementary figure S1.

5. DISCUSSION

The 16S-rRNA maximum likelihood tree showed that the cyanobacterium *Cyanothece* was closer to Synechococcales than to Oscillatoriales. According to Rippka & Cohen-Bazire

(1983) several smaller unicellular cyanobacteria with oval cells were falsely designated as *Cyanothece*. Usually, their genetic position and morphological features are far from the type species *Cyanothece aeruginosa*. Probably this is the case of *Cyanothece* sp. PCC 7425. Bandyopadhyay *et al*. (2011) reported several anomalies in this strain; its cells are smaller than those of the other strains analyzed and are more cylindrical. The GC content, which is normally of ~40% for the genus *Cyanothece*, is of ~50% in *Cyanothece* sp. PCC 7425. Also, unlike the other *Cyanothece* strains, *Cyanothece* sp. PCC7425 fixes nitrogen only under anaerobic conditions. The authors believe that this strain is evolving independently of the other *Cyanothece* strains, but this can be just a misidentified strain. Another unexpected placement in the phylogenetic tree was the Oscillatoriales *Crinalium epipsammum*, found closely related to some Nostocales strains as also reported by Howard-Azzeh *et al*. (2014).

The 16S-rRNA phylogenetic relationships among the species analyzed were used to evaluate if the Nif proteins also followed the same evolution pattern. It was found that some proteins indeed followed the organism evolution, while others did not. This is because the evolution rate is different for each gene and many extra genes attained by duplication degenerate and become pseudogenes, but some may diverge to have new functions. One of the gene copies can degenerate, since the other copy can maintain the original function that is needed for the organism to live. Duplicated genes may distinguish in expression patterns (subfunctionalization) and even acquire new functions (neofunctionalization) (Landgraf *et al*., 1999; Dermitzakis & Clark, 2001; Ohta, 2013).

5.1. NIF FAMILY PHYLOGENY

The Phylogenetic relationships of Nif proteins congregated in 4 main clades. Clade I reunited NifDEKN, VnfDEKN and AnfDKN proteins. Its arrangement suggests that *nifN*, *nifK*, *nifD* and *nifE* have a common ancestor, that *nifN* and *nifK* share a more recent ancestor and *nifD*

and *nifE* also share one. According to Fani *et al*. (2000) the NifE and NifN proteins have significant similarity to NifD and NifK, respectively, and are believed to have originated from an ancient duplication of the *nifDK* operon. *Rhodopseudomonas* presented 4 copies of VnfN, 3 of VnfE and 3 of NifE, of which one copy of NifE was positioned more distant from the others, suggesting that it accumulated mutations for a longer time, that is, the duplication of the other two may be more recent. *Rhodopseudomonas* VnfE formed an external group to the NifE, NifD, AnfD and VnfD sequences. And *Rhodopseudomonas* VnfN formed an external group to NifN, NifK, AnfK and VnfK. This implies that the *nif* and *vnf* families in *Rhodopseudomonas* have expanded.

Clade II organization suggests that a duplication of an ancestral gene culminated in the *nifU* and *nifB* genes. Also, some MoaA proteins from *Rhodopseudomonas*, *Microcystis* and *Cyanothece* were similar to NifB proteins. In fact, the MoaA group was positioned as a sister-group of NifB. Their similarity probably comes from a common ancestor gene that duplicated and acquired new functions. MoaA catalyzes an early step in the biosynthesis of the molybdenum cofactor (MoCo) and NifB is required for the biosynthesis of the iron-molybdenum (or iron-vanadium) cofactor (Toyama *et al*., 1997).

Clade III shows that *nifH* is a sister-group of *chlL*. The ChlL enzyme is a light-independent protochlorophyllide reductase that contains a [4Fe-4S] cluster, and structurally resembles the Fe protein/MoFe protein complex of NifH. A number of proteins have structural and mechanistic similarities as well as evolutionary relationships with the NifH protein, notable among them being: light independent protochlorophyllide (Pchlide) reductase (ChlL/FrxC or bChL) and MinD that functions in spatial regulation of cell division (Thakur *et al*., 2013). Although involved in very diverse biological processes, these proteins share a common structural framework. Phylogenetic research suggests that NifH and ChlL are very similar and could have evolved from a common ancestor (Reinbothe *et al*., 2010), probably

MinD (Lahiri *et al.*, 2008; Thakur *et al*., 2013). There are also some instances of nitrogenase genes being recruited into the photosynthetic apparatus pathway (Xiong et al. 2000), which could be the case with *nifH* and *chlL*, for example by *nifH* duplication and the copy acquiring new function.

Clade IV consists of NifS, IscS, SufS and CsdA proteins, all of them are cysteine desulfurases/aminotransferases involved in the iron-sulfur (FeS) cluster biosynthesis. NifS and IscS grouped together, SufS formed a group with a few CsdA sequences and the other CsdA sequences formed a separate group. Probably SufS is derived from CsdA.

IscS was first identified in *A.vinelandii* as a NifS homolog. It was named IscS for its proposed role in iron-sulfur cluster biosynthesis (Zheng *et al*., 1998). IscS is thought to play a pivotal role in Fe-S cluster assembly in vivo, whereas NifS specializes in the maturation of nitrogenase. Some organisms have more than one copy of a *nifS* homolog, they are named *iscS*, *csdA* and *sufS* (also called *csdB*), and also several non-diazotrophic prokaryotes have *nifS*-like genes (Mihara *et al*., 2002).

The phylogenetic tree arrangement in this study suggests that the ancestral gene duplicated and originated *nifS* and *csdA*. Then, *nifS* duplicated and one copy gained a new function and became *iscS*, while *csdA* duplicated two times, and then one of these copies gained a new function and became *sufS* (Fig.4). Cyanobacteria seem to have, in general, two copies of CsdA and one of SufS, some strains may have lost a copy or a copy may have duplicated again.

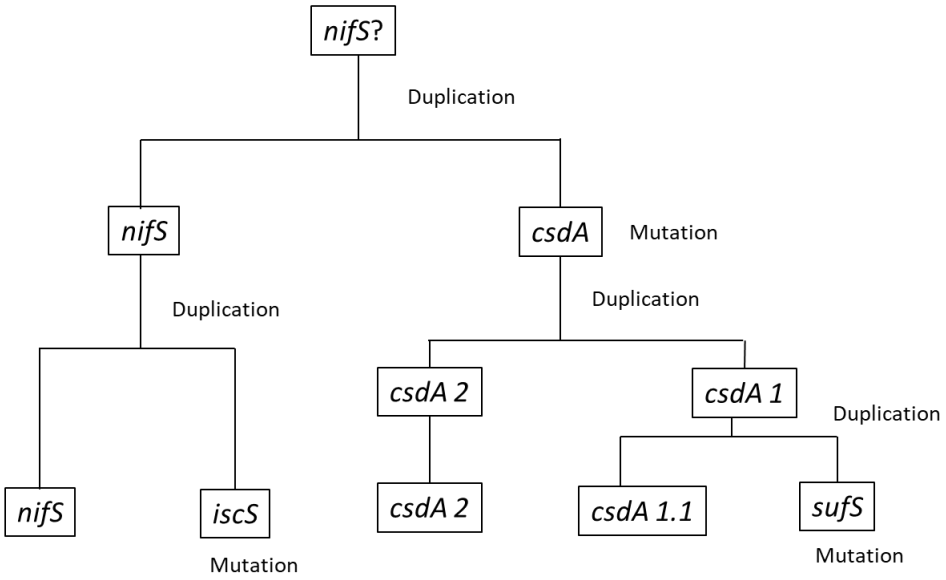


Fig.4- Probable evolution of *nifS, iscS, csdA* and *sufS* genes. At the top *nifS* is the probable ancestral. At the very bottom the current *nifS*, *iscS*, *csdA* and *sufS* genes, after a series of duplication events and mutations that culminated in neofunctionalization

5.2. DIAZOTROPHY ACROSS THE TREE

For some time, the evolutionary past of diazotrophy has been discussed. The key emphasis is on how the irregular distribution of nitrogen fixation in distantly related prokaryotes can be explained.

To explain that distribution three hypothesis were raised. The first says that diazotrophy arose multiple times through convergent evolution (Postgate & Eady, 1988). The second says that it arose once and it has been transferred lateraly to various lineages (Normand & Bousquet, 1989). And the third says that it arose early in prokaryotes and was commonplace among them, but it has since been lost by many lineages and retained by a few (Young 1992; Normand *et al*., 1992).

In this work some cyanobacteria that do not fix nitrogen showed the presence of *nif* genes. Until now, almost all nitrogen-fixing species contain a minimum of six genes dedicated to both nitrogenase catalytic components (*nifHDK*) and two FeMo-cofactor biosynthetic enzymes (*nifEN* and *nifB*) (Addo & Dos Santos, 2020). It can be suggested that these cyanobacteria that had *nif* genes but were not diazothrophic, had leftovers of nitrogen

fixing genes while most of the other *nif* genes were probably lost. There was no expansion of the Nif family in certain orders, like Nostocales, but there was the apparent gene loss in other strains.

It is unclear whether Horizontal Gene Transference (HGT) or vertical descent has had a larger impact in the spread of N_2 fixation. According to Latysheva *et al*. (2012) the rate of gain of N_2 fixation within the cyanobacterial clade is significant, but much lower than the rate of loss. Consistent with this, stochastic character mapping suggested a higher number of trait losses than gains. Zehr *et al*., (1997) and Latysheva *et al*. (2012) also suggested that any HGT in cyanobacterial N_2 fixation would have happened near the root of the phylogeny just after the split between *Gloeobacter violaceus* and all remaining cyanobacteria. Latysheva *et al*. (2012) analysis has shown that loss of N_2 fixation has been more prevalent than HGT in cyanobacteria. They concluded that the origin of N_2 fixation is pre-cyanobacterial, and probably in the Last Universal Common Ancestor, followed by repeated independent loss (Fani *et al.*, 2000; Normand *et al*., 1992; Young 1992). Boyd *et al*. (2011) however, concluded that Nif nitrogenase is unlikely to have been associated with the Last Universal Common Ancestor based on phylogenetic relationships of proteins required for Nif, Anf and Vnf nitrogenase.

5.3. NITROGENASE ANCESTOR

Azotobacter vinelandii and *Rhodopseudomonas palustris* are both proteobacteria that synthesize one of three nitrogenases depending on the metal content of the medium in which they are grown. The homologous polypeptides of the Vnf nitrogenase and Anf nitrogenase are encoded by the structural genes *vnfHDK* and *anfHDK*, respectively. The Nif phylogenetic tree could be interpreted as a Vnf nitrogenase being ancestral to Anf and Nif nitrogenases. In clade I VnfN branched before NifN, VnfK, AnfK and NifK, and also, VnfE branched earlier than

NifN, VnfD, AnfD and NifD. But this would be a mistaken approach, beacuse that result was only due to one anomalous strain, *Rhodopseudomonas palustris* CGA009, which had several gene duplications. Indeed, it has been suggested (Anbar & Knoll, 2002; Raymond *et al*., 2004 and Anbar, 2008) that alternative nitrogenases were responsible for N_2-fixation on early Earth because oceans were depleted of bioavailable Mo. However, alternative nitrogenases have only been found in species that contain the Nif nitrogenase and are so far known to be expressed only under Mo-deficient conditions. Also, the Methanogens, the only Archeas that can fix nitrogen possesses Nif nitrogenase. This is why such an evolutionary path is unlikely. According to Mus *et al*. (2018) the Vnf nitrogenase and the Anf nitrogenase are not the ancestors of the Nif nitrogenase. They suggested that nitrogenase arose within the methanogenic archaea and then gave rise to the alternative paralogous enzymes that employ alternative metals because of the selective pressure imposed by environmental metal availability. Stüeken *et al.* (2015) analyzed nitrogen isotope ratios from marine and fluvial sedimentary rocks between 3.2 and 2.75 billion years ago and came to the conclusion that the origin of biological nitrogen fixation was probably 3.2 billion years ago. They also suggested that molybdenum was bioavailable in the mid-Archaean ocean long before the Great Oxidation Event. Boyd *et al*. (2011) compared the substitution rates estimated for proteins required for the biosynthesis of the nitrogenase active site and came to the conclusion that Nif emerged approximately 1.5-2.2 billion years ago, after the origin of oxygenic photosynthesis and the widespread oxygenation of the biosphere.

6. REFERENCES

Addo, M. A., & Dos Santos, P. C. 2020. Distribution of Nitrogen-Fixation Genes in Prokaryotes Containing Alternative Nitrogenases. *ChemBioChem*, 21:1749-1759.

Anbar, A. D., & Knoll, A. H. 2002. Proterozoic ocean chemistry and evolution: a bioinorganic bridge?. *Science*, 297:1137-1142.

Anbar, A. D. 2008. Elements and evolution. *Science*, 1481-1483.

Bandyopadhyay, A., Elvitigala, T., Welsh, E., Stöckel, J., Liberton, M., Min, H., Sherman, L. A., & Pakrasi, H. B. 2011. Novel metabolic attributes of the genus Cyanothece, comprising a group of unicellular nitrogen-fixing cyanobacteria. *MBio*, 2:1-10.

Boyd, E. S., Anbar, A. D., Miller, S., Hamilton, T. L., Lavin, M., & Peters, J. W. 2011. A late methanogen origin for molybdenum-dependent nitrogenase. *Geobiology*, 9:221-232.

Dermitzakis, E. T., & Clark, A. G. 2001. Differential selection after duplication in mammalian developmental genes. Molecular biology and evolution, 18:557-562.

Dos Santos, P. C., Fang, Z., Mason, S. W., Setubal, J. C., & Dixon, R. 2012. Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics*, 13:1-12.

Fani, R., Gallo, R., & Lio, P. 2000. Molecular evolution of nitrogen fixation: the evolutionary history of the *nifD*, *nifK*, *nifE*, and *nifN* genes. *Journal of molecular evolution*, 51:1-11.

Hilton, J. A., Foster, R. A., Tripp, H. J., Carter, B. J., Zehr, J. P., & Villareal, T. A. 2013. Genomic deletions disrupt nitrogen metabolism pathways of a cyanobacterial diatom symbiont. *Nature communications*, 4:1-7.

Howard-Azzeh, M., Shamseer, L., Schellhorn, H. E., & Gupta, R. S. 2014. Phylogenetic analysis and molecular signatures defining a monophyletic clade of heterocystous cyanobacteria and identifying its closest relatives. *Photosynthesis*

research, 122:171-185.

Joerger, R. D., Bishop, P. E., & Evans, H. J. 1988. Bacterial alternative nitrogen fixation systems. *CRC Critical reviews in microbiology*, 16(1), 1-14.

Katoh, K., Rozewicki, J., & Yamada, K. D. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in bioinformatics*, 20:1160-1166.

Lahiri, S., Pulakat, L., & Gavini, N. 2008. NifH: Structural and Mechanistic Similarities with Proteins Involved in Diverse Biological Processes. *American Journal of Biochemistry and Biotechnology*, 4:304–316.

Landgraf, R., Fischer, D., & Eisenberg, D. 1999. Analysis of heregulin symmetry by weighted evolutionary tracing. *Protein engineering*, 12:943-951.

Latysheva, N., Junker, V. L., Palmer, W. J., Codd, G. A., & Barker, D. 2012. The evolution of nitrogen fixation in cyanobacteria. *Bioinformatics*, 28:603-606.

Mevarech, M., Rice, D., & Haselkorn, R. 1980. Nucleotide sequence of a cyanobacterial nifH gene coding for nitrogenase reductase. *Proceedings of the National Academy of Sciences*, 77:6476-6480.

Mihara, H., & Esaki, N. 2002. Bacterial cysteine desulfurases: Their function and mechanisms. *Applied Microbiology and Biotechnology*, 60:12-23.

Mus, F., Alleman, A. B., Pence, N., Seefeldt, L. C., & Peters, J. W. 2018. Exploring the alternatives of biological nitrogen fixation. *Metallomics*, 10:523-538.

Navarro-González, R., McKay, C. P., & Mvondo, D. N. 2001. A possible nitrogen crisis for Archaean life due to reduced nitrogen fixation by lightning. *Nature*, 412:61-64.

Normand, P., & Bouquet, J. 1989. Phylogeny of nitrogenase sequences in *Frankia* and other nitrogen-fixing microorganisms. *Journal of molecular evolution*, 29:436-447.

Normand, P., Gouy, M., Cournoyer, B., & Simonet, P. 1992. Nucleotide sequence of

nifD from Frankia alni strain ArI3: phylogenetic inferences. *Molecular biology and evolution*, 9:495-506.

Ohta, T. 2013. Evolution of Gene Families. *In* Brenner's Encyclopedia of Genetics: Second Edition (Vol. 2, Issue January 2001). *Elsevier Inc*, pp. 563-565. https://doi.org/10.1016/B978-0-12-374984-0.00497-6

Postgate, J. R., & Eady, R. R. 1988. Evolution of biological nitrogen fixation. *In* Bothe, H., de Bruijn, F. J. & Newton W. E. [Eds.] Nitrogen fixation: Hundred years after. *Proceedings of the 7th International Congress on N*. Stuttgart, Germany, pp. 31-40.

Raymond, J., Siefert, J. L., Staples, C. R., & Blankenship, R. E. 2004. The Natural History of Nitrogen Fixation. *Molecular Biology and Evolution*, 21:541–554.

Reinbothe, C., El Bakkouri, M., Buhr, F., Muraki, N., Nomata, J., Kurisu, G., Fujita, Y., & Reinbothe, S. 2010. Chlorophyll biosynthesis: spotlight on protochlorophyllide reduction. *Trends in plant science*, 15:614-624

Rippka, R., & Cohen-Bazire, G. 1983. The cyanobacteriales: a legitimate order based on the type strain Cyanobacterium stanieri?. *In* Annales de l'Institut Pasteur/Microbiologie. Elsevier Masson, pp. 21-63.

Stüeken, E. E., Buick, R., Guy, B. M., & Koehler, M. C. 2015. Isotopic evidence for biological nitrogen fixation by molybdenum-nitrogenase from 3.2 Gyr. *Nature*, 520:666-669.

Thakur, S., Bothra, A. K., & Sen, A. 2013. Functional divergence outlines the evolution of novel protein function in NifH/BchL protein family. *Journal of Biosciences*, 38:733-740.

Toyama, H., Chistoserdova, L., & Lidstrom, M. E. 1997. Sequence analysis of pqq genes required for biosynthesis of pyrroloquinoline quinone in Methylobacterium extorquens AM1 and the purification of a biosynthetic intermediate. *Microbiology*, 143:595-602.

Xiong, J., Fischer, W. M., Inoue, K., Nakahara, M., & Bauer, C. E. 2000. Molecular

evidence for the early evolution of photosynthesis. Science, 289:1724-1730.

Young, J. P. W. 1992. Phylogenetic classification of nitrogen-fixing organisms. *In* Stacey,G.et al. [Eds.] *Biological Nitrogen Fixation*. Chapman & Hall. NY, pp. 43-86.

Yung, Y. L., & McElroy, M. B. 1979. Fixation of nitrogen in the prebiotic atmosphere. *Science*, 203:1002-1004.

Zehr, J. P., Mellon, M. T., & Hiorns, W. D. 1997. Phylogeny of cyanobacterial nifH genes: evolutionary implications and potential applications to natural assemblages. *Microbiology*, 143:1443-1450.

Zheng, L., Cash, V. L., Flint, D. H., & Dean, D. R. 1998. Assembly of iron-sulfur clusters: identification of an iscSUA-hscBA-fdx gene cluster from *Azotobacter vinelandii*. *Journal of Biological Chemistry*, 273:13264-13272.

CONSIDERAÇÕES FINAIS

Este trabalho teve como foco entender a capacidade de *Raphidiopsis raciborskii*, um organismo conhecido por sua capacidade de produzir toxinas e por sua capacidade de adaptação, de fixar nitrogênio atmosférico.

O sucesso ecológico dessa espécie é algo muito preocupante e uma das características que poderiam ser vantajosas para a sua dominância em ambientes aquáticos é a diazotrofia. Apesar dessa qualidade ser importante, existe um *gap* de conhecimento em relação à capacidade desta espécie de fixar nitrogênio, em geral é pouco estudada. Com isso em mente foram feitos três capítulos.

No capítulo 1: discutimos sobre a filogeografia da espécie *R. raciborskii*, incluindo a presença de cepas fixadoras e não fixadoras de nitrogênio. Observamos também que ambas as cepas podem co-ocorrer na natureza e sugerimos que populações sem heterócitos surgiram de populações com heterócitos por seleção natural em ambientes ricos em nitrogênio. Também, endossamos a proposta de Aguilera *et al*. (2018) que os gêneros *Raphidiopsis* e *Cylindrospermopsis* devem ser reunidos em um único gênero, ao adicionar mais cepas e mais regiões genômicas às análises filogenéticas das cepas. E, por fim, nossos experimentos também mostraram que heterócitos podem se desenvolver mesmo em um ambiente rico em nitrogênio.

No capítulo 2 mostramos que o efeito do nitrato na produção de heterócitos e na fixação de nitrogênio é variável dependendo da cepa de *R.raciborskii* estudada. Também, assim como no capítulo anterior observamos que o desenvolvimento de heterócitos pode ocorrer em meio rico em nitrogênio dissolvido, e ainda, pudemos perceber que a expressão de genes importantes para o funcionamento da nitrogenase também ocorre, ao contrário do que foi visto em outras cianobactérias diazotróficas.

No capítulo 3 fizemos uma breve revisão sobre a evolução da família gênica *nif*,

grande responsável pelo funcionamento da nitrogenase e fixação de nitrogênio. Também discutimos sobre os tipos de nitrogenase que existem (Nif, Anf e Vnf) e sobre como a nitrogenase pode ter surgido. As nitrogenases Anf e Vnf provavelmente surgiram da Nif nitrogenase que utiliza como cofator o molibdênio. Quando a Nif nitrogenase surgiu o molibdênio era escasso e esse ambiente pode ter causado uma pressão seletiva que deu origem a nitrogenases que utilizam outros metais.

SUPPLEMENTARY MATERIAL
CHAPTER ONE

1. Table S1

Species	Strain	Origin	Coordinates	Isolate date
R.raciborskii	UFMG34	Lagoa Santa Lake, MG, Brazil	19°38′S, 43°53′W	December/ 2003
R.raciborskii	UFMG36	Lagoa Santa Lake, MG, Brazil	19°38′S, 43°53′W	December/ 2003
R.raciborskii	UFMG37	Lagoa Santa Lake, MG, Brazil	19°38′S, 43°53′W	December/ 2003
R.raciborskii	UFMG40	Lagoa Santa Lake, MG, Brazil	19°38′S, 43°53′W	December/ 2003
R.raciborskii	UFMG42	Lagoa Santa Lake, MG, Brazil	19°38′S, 43°53′W	December/ 2003
R.raciborskii	UFMG44	Lagoa Santa Lake, MG, Brazil	19°38′S, 43°53′W	December/ 2003
Raphidiopsis	UFMG181	Pampulha Reservoir, MG, Brazil	19°51′15″, 43°58′43″W	April/ 2013
R.raciborskii	UFMG182	Pampulha Reservoir, MG, Brazil	19°51′15″, 43°58′43″W	March/ 2014
Raphidiopsis	UFMG183	Pampulha Reservoir, MG, Brazil	19°51′15″, 43°58′43″W	March/ 2014
R.raciborskii	UFMG185	Lagoa Santa Lake, MG, Brazil	19°38′S, 43°53′W	April/ 2014
R.raciborskii	UFMG186	Lagoa Santa Lake, MG, Brazil	19°38′S, 43°53′W	April/ 2014
R.raciborskii	UFMG187	Lagoa Santa Lake, MG, Brazil	19°38′S, 43°53′W	April/ 2014
R.raciborskii	UFMG188	Lagoa Santa Lake, MG, Brazil	19°38′S, 43°53′W	April/ 2014
R.raciborskii	UFMG232	Lagoa Santa Lake, MG, Brazil	19°38′S, 43°53′W	September/ 2016
Raphidiopsis	UFMG233	Pampulha Reservoir, MG, Brazil	19°51′15″, 43°58′43″W	September/ 2016
Raphidiopsis	UFMG234	Pampulha Reservoir, MG, Brazil	19°51′15″, 43°58′43″W	September/ 2016

1 2. Table S2

Species	Strain	Trichome form	Vegetative cells (µm)		Heterocyte
			Length	Width	
R.raciborskii	UFMG34	straight	3.42±0.69	11.82±4.71	Present
R.raciborskii	UFMG36	straight	2.94±0.84	9.13±4.04	Present
R.raciborskii	UFMG37	straight	2.39±0.31	10.13±2.81	Present
R.raciborskii	UFMG40	straight	2.50±0	9.81±5.71	Present
R.raciborskii	UFMG42	straight	2.50±0.41	8.38±3.06	Present
R.raciborskii	UFMG44	straight	2.88±0.59	10.56±5.04	Present
Raphidiopsis sp.	UFMG181	flexuous	2.94±0.61	9.63±5.15	Absent
R.raciborskii	UFMG182	straight	4.94±0.64	6.43±2.31	Present
Raphidiopsis sp.	UFMG183	flexuous	4.13±0.71	8.69±2.61	Absent
R.raciborskii	UFMG185	straight	3.50±0.96	6.88±2.80	Present
R.raciborskii	UFMG186	straight	5.06±1.25	8.50±2.74	Present
R.raciborskii	UFMG187	straight	4.81±0.61	9.94±3.45	Present
R.raciborskii	UFMG188	straight	4.38±0.95	5.56±1.54	Present
R.raciborskii	UFMG232	straight	2.75±2.51	8.62±3.67	Present
Raphidiopsis sp.	UFMG233	straight	2.50±0	7.19±4.90	Absent
Raphidiopsis sp.	UFMG234	straight	2.50±0	7.5±0	Absent

2

3

1 3. Table S3

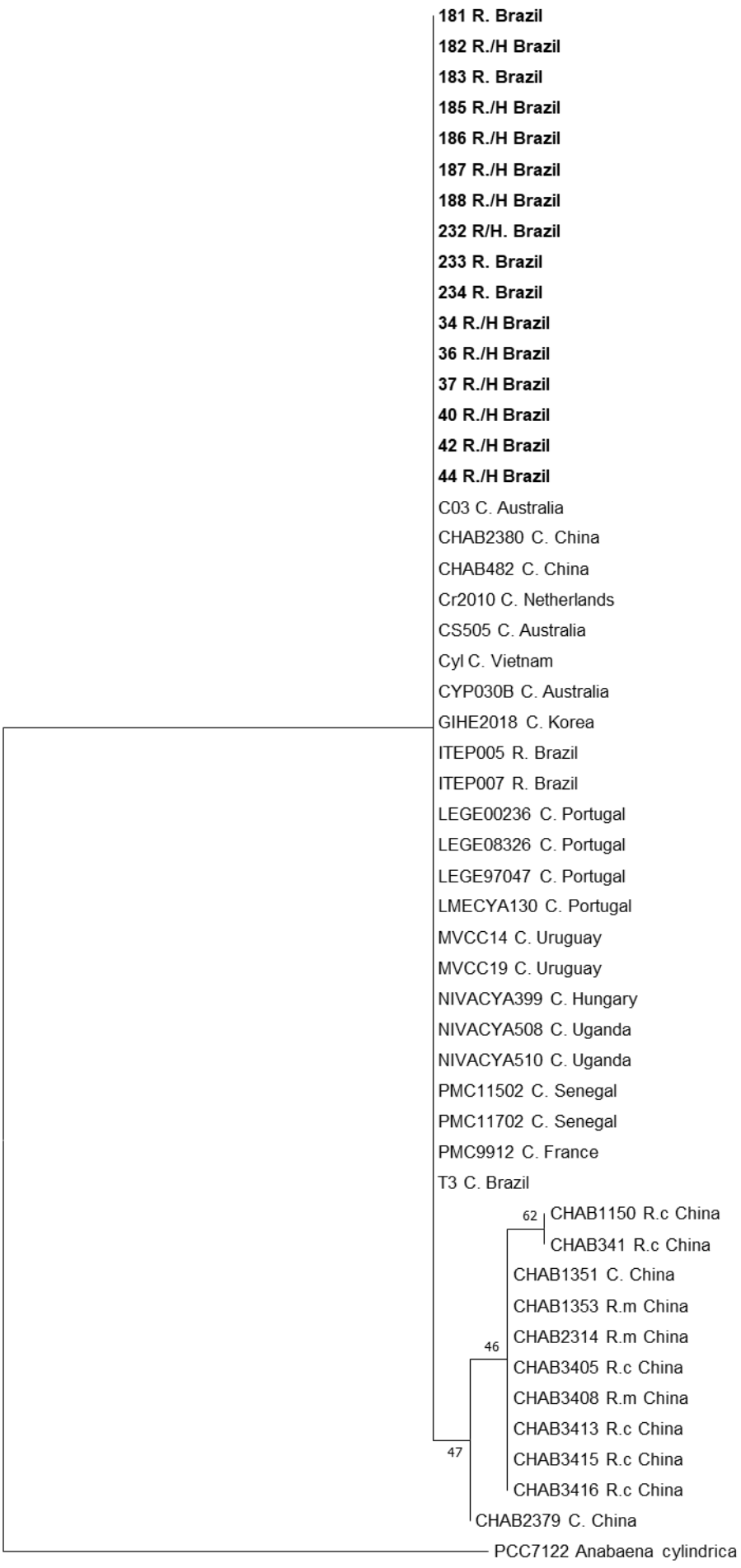
			Accession number			
Strain	Country	Species	16S rRNA	ITS	*rpoC1*	*nifH*
UFMG34	Brazil	*R. raciborskii*	MT129552	MT129544	MT231712	MT219501
UFMG36	Brazil	*R. raciborskii*	MT129553	MT129537	MT231713	MT219502
UFMG37	Brazil	*R. raciborskii*	MT129554	MT129545	MT231714	MT219503
UFMG40	Brazil	*R. raciborskii*	MT129555	MT129546	MT231715	MT219504
UFMG42	Brazil	*R. raciborskii*	MT129556	MT129547	MT231716	MT219505
UFMG44	Brazil	*R. raciborskii*	MT129557	MT129548	MT231717	MT219506
UFMG181	Brazil	*Raphidiopsis* sp.	MT129558	MT129536	MT231718	-
UFMG182	Brazil	*R. raciborskii*	MT129559	MT129538	MT231719	MT219507
UFMG183	Brazil	*Raphidiopsis* sp.	MT129560	MT129539	MT231720	-
UFMG185	Brazil	*R. raciborskii*	MT129561	MT129540	MT231721	MT219508
UFMG186	Brazil	*R. raciborskii*	MT129562	MT129541	MT231722	MT219509
UFMG187	Brazil	*R. raciborskii*	MT129563	MT129542	MT231723	MT219510
UFMG188	Brazil	*R. raciborskii*	MT129564	MT129543	MT231724	MT219511
UFMG232	Brazil	*R. raciborskii*	MT129567	MT129551	MT231727	MT219512
UFMG233	Brazil	*Raphidiopsis* sp.	MT129565	MT129549	MT231725	-
UFMG234	Brazil	*Raphidiopsis* sp.	MT129566	MT129550	MT231726	-

ITEP-005	Brazil	*Raphidiopsis* sp.	AJ582109	AJ582275	AJ582285	-
ITEP-007	Brazil	*Raphidiopsis* sp.	AJ582110	AJ582274	-	-
T3	Brazil	*C. raciborskii*	JQ707295	MK244670	MK241568	-
MVCC14	Uruguay	*C. raciborskii*	MBQY01000081	HQ112347	MBQY01000057	HQ112345
MVCC19	Uruguay	*C. raciborskii*	VIRO01000052	HQ112348	VIRO01000015	HQ112346
PMC99.08	Mexico	*C. raciborskii*	KF287099	KF287109	AJ582289	AJ582098
CHAB341	China	*R. curvata*	FJ890620	FJ890676	FJ890594	-
CHAB1150	China	*R. curvata*	JN873923	KU360063	FJ890595	-
CHAB1351	China	*C. raciborskii*	FJ890633	FJ890665	FJ890606	FJ890687
CHAB1353	China	*R. mediterranea*	FJ890619	KU360072	FJ890612	-
CHAB3405	China	*R. curvata*	KU360054	KU360065	KJ139732	-
CHAB3408	China	*R. mediterranea*	KU360060	KU360075	KX380687	-
CHAB3413	China	*R. curvata*	KU360057	KU360068	KJ139735	-
CHAB3415	China	*R. curvata*	KU360058	KU360069	KJ139736	-
CHAB3416	China	*R. curvata*	KU360059	KU360070	KJ139723	-
CHAB2314	China	*R. mediterranea*	FJ890660	FJ890615	FJ890613	-
CHAB2379	China	*C. raciborskii*	FJ890634	FJ890669	FJ890601	FJ890688
CHAB2380	China	*C. raciborskii*	FJ890623	FJ890670	FJ890605	FJ890677

CHAB482	China	*C. raciborskii*	FJ890627	FJ890661	FJ890596	FJ890678
GIHE2018	Korea	*C. raciborskii*	VHLJ01000002	VHLJ01000002	VHLJ01000002	VHLJ01000001
Cyl	Vietnam	*C. raciborskii*	HQ407329	HQ407357	HQ407339	
NIVA-CYA 508	Uganda	*C. raciborskii*	KF287096	AM502070	AM502052	AM502061
NIVA-CYA 510	Uganda	*C. raciborskii*	KF287097	AM502072	AM502063	AM502054
NIVA-CYA 511	Uganda	*C. raciborskii*	-	AM502073	AM502064	AM502055
PMC115.02	Senegal	*C. raciborskii*	KF287098	KF287108	AJ582093	AJ582101
PMC117.02	Senegal	*C. raciborskii*	GQ859598	AJ582271	AJ582092	AJ582100
PMC992	France	*C. raciborskii*	KF287091	AJ582273	AJ582287	AJ582095
LEGE 97047	Portugal	*C. raciborskii*	HQ407325	HQ407357	HQ407337	-
LEGE 00236	Portugal	*C. raciborskii*	KF287094	KF287102	KF287116	-
LEGE 08326	Portugal	*C. raciborskii*	HQ407324	HQ407357	HQ407335	-
LMECYA 130	Portugal	*C. raciborskii*	HQ407323	HQ407344	HQ407334	-
ZIE05CR	Germany	*C. raciborskii*	-	AM502074	AM502056	AM502065
ZIE11CR	Germany	*C. raciborskii*	-	AM502075	AM502057	AM502066
ZIE13CR	Germany	*C. raciborskii*	-	AM502076	AM502058	AM502067

Cr2010	Netherlands	*C. raciborskii*	PVMC01000089	PVMC01000089	PVMC01000069	PVMC01000055
NIVA-CYA 399	Hungary	*C. raciborskii*	KF287089	KF287102	KF287112	-
C03	Australia	*C. raciborskii*	NJHU01000086	NJHU01000336	NJHU01000164	NJHU01000043
CS-505	Australia	*C. raciborskii*	LYXA01000001	EU552055	LYXA01000001	LYXA01000001
CYP-030B	Australia	*C. raciborskii*	AJ582108	AJ582277	AJ582089	-
PCC 7122	United Kingdom	*Anabaena cylindrica*	NR_102457	KT290325	AB074793	NC_019771

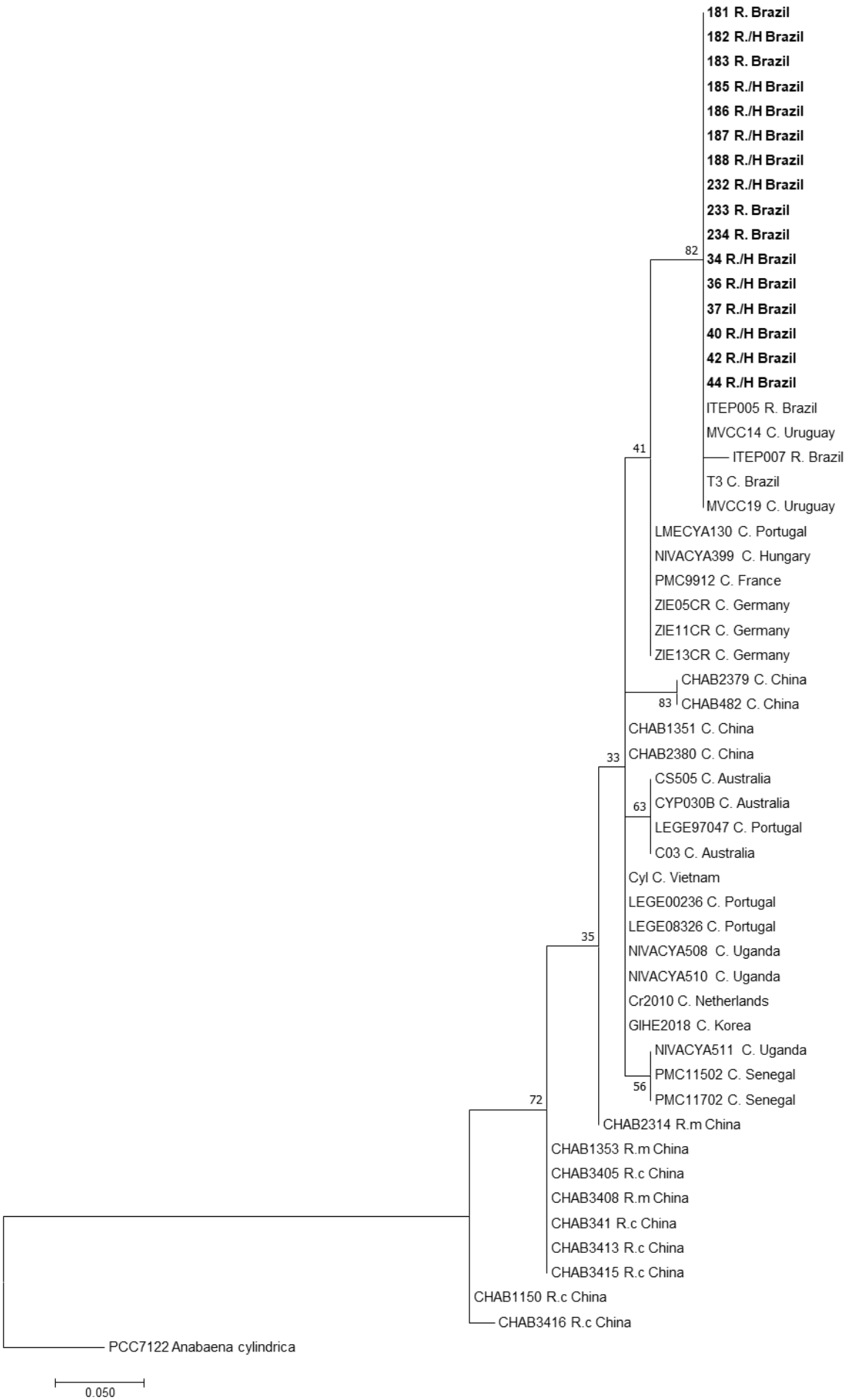
1


181 R. Brazil
182 R./H Brazil
183 R. Brazil
185 R./H Brazil
186 R./H Brazil
187 R./H Brazil
188 R./H Brazil
232 R/H. Brazil
233 R. Brazil
234 R. Brazil
34 R./H Brazil
36 R./H Brazil
37 R./H Brazil
40 R./H Brazil
42 R./H Brazil
44 R./H Brazil
C03 C. Australia
CHAB2380 C. China
CHAB482 C. China
Cr2010 C. Netherlands
CS505 C. Australia
Cyl C. Vietnam
CYP030B C. Australia
GIHE2018 C. Korea
ITEP005 R. Brazil
ITEP007 R. Brazil
LEGE00236 C. Portugal
LEGE08326 C. Portugal
LEGE97047 C. Portugal
LMECYA130 C. Portugal
MVCC14 C. Uruguay
MVCC19 C. Uruguay
NIVACYA399 C. Hungary
NIVACYA508 C. Uganda
NIVACYA510 C. Uganda
PMC11502 C. Senegal
PMC11702 C. Senegal
PMC9912 C. France
T3 C. Brazil
62
CHAB1150 R.c China
CHAB341 R.c China
CHAB1351 C. China
CHAB1353 R.m China
CHAB2314 R.m China
46
CHAB3405 R.c China
CHAB3408 R.m China
CHAB3413 R.c China
47
CHAB3415 R.c China
CHAB3416 R.c China
CHAB2379 C. China
PCC7122 Anabaena cylindrica
0.0100

1

4. Fig.S1

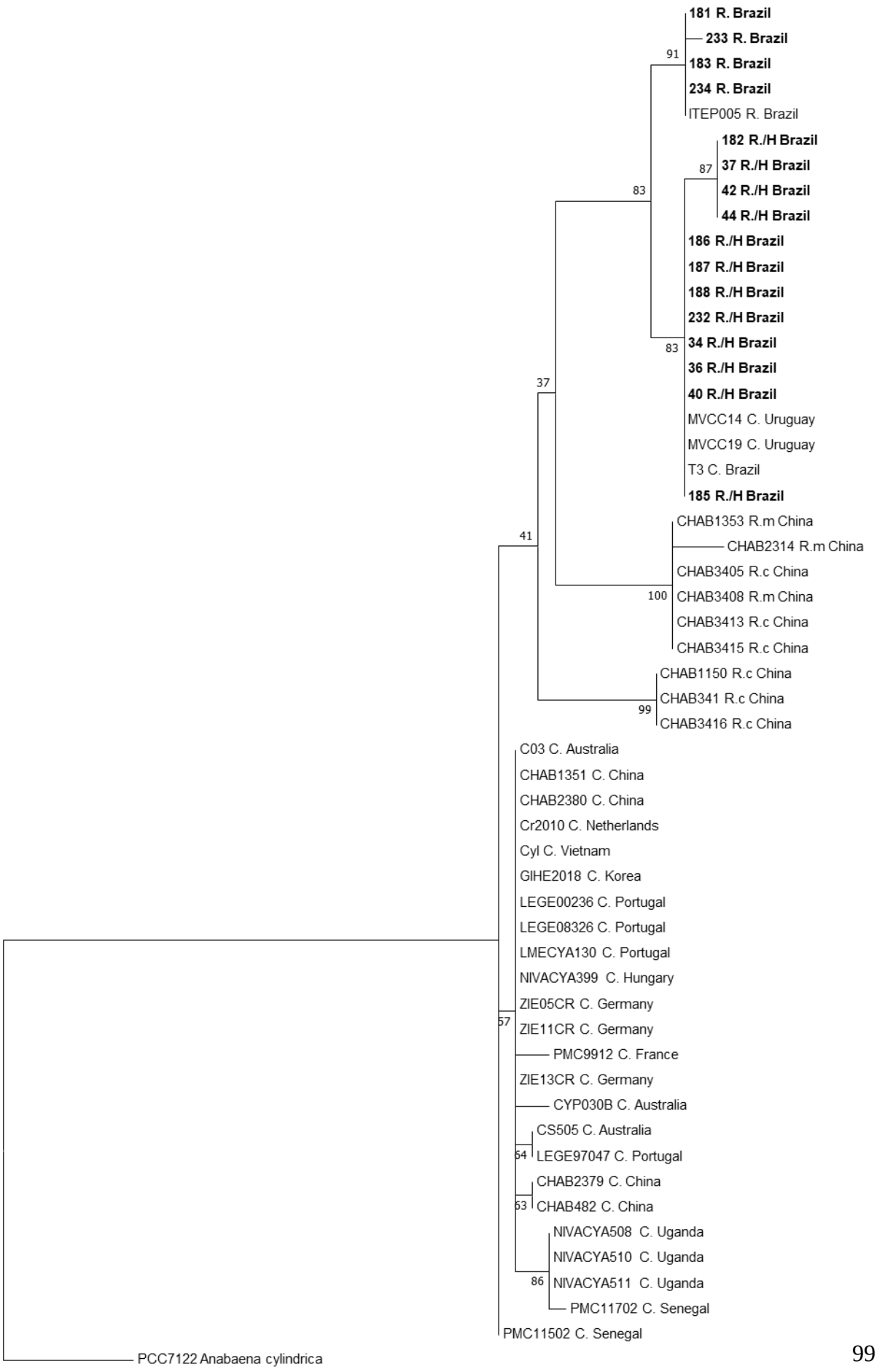
Fig.S1- Maximum likelihood phylogenetic analysis of *Raphidiopsis* / *Cylindrospermopsis* using 16S rRNA sequences. KT290325.1 *A.cylindrica* was selected as the outgroup taxon. Support values for the branches are indicated at the nodes. Strains sequenced in this study are indicated in bold.


181 R. Brazil
182 R./H Brazil
183 R. Brazil
185 R./H Brazil
186 R./H Brazil
187 R./H Brazil
188 R./H Brazil
232 R./H Brazil
233 R. Brazil
234 R. Brazil
82
34 R./H Brazil
36 R./H Brazil
37 R./H Brazil
40 R./H Brazil
42 R./H Brazil
44 R./H Brazil
ITEP005 R. Brazil
MVCC14 C. Uruguay
41
ITEP007 R. Brazil
T3 C. Brazil
MVCC19 C. Uruguay
LMECYA130 C. Portugal
NIVACYA399 C. Hungary
PMC9912 C. France
ZIE05CR C. Germany
ZIE11CR C. Germany
ZIE13CR C. Germany
CHAB2379 C. China
83
CHAB482 C. China
CHAB1351 C. China
33
CHAB2380 C. China
CS505 C. Australia
63
CYP030B C. Australia
LEGE97047 C. Portugal
C03 C. Australia
Cyl C. Vietnam
LEGE00236 C. Portugal
LEGE08326 C. Portugal
35
NIVACYA508 C. Uganda
NIVACYA510 C. Uganda
Cr2010 C. Netherlands
GIHE2018 C. Korea
NIVACYA511 C. Uganda
PMC11502 C. Senegal
56
PMC11702 C. Senegal
72
CHAB2314 R.m China
CHAB1353 R.m China
CHAB3405 R.c China
CHAB3408 R.m China
CHAB341 R.c China
CHAB3413 R.c China
CHAB3415 R.c China
CHAB1150 R.c China
CHAB3416 R.c China
PCC7122 Anabaena cylindrica
0.050

1

5. Fig.S2

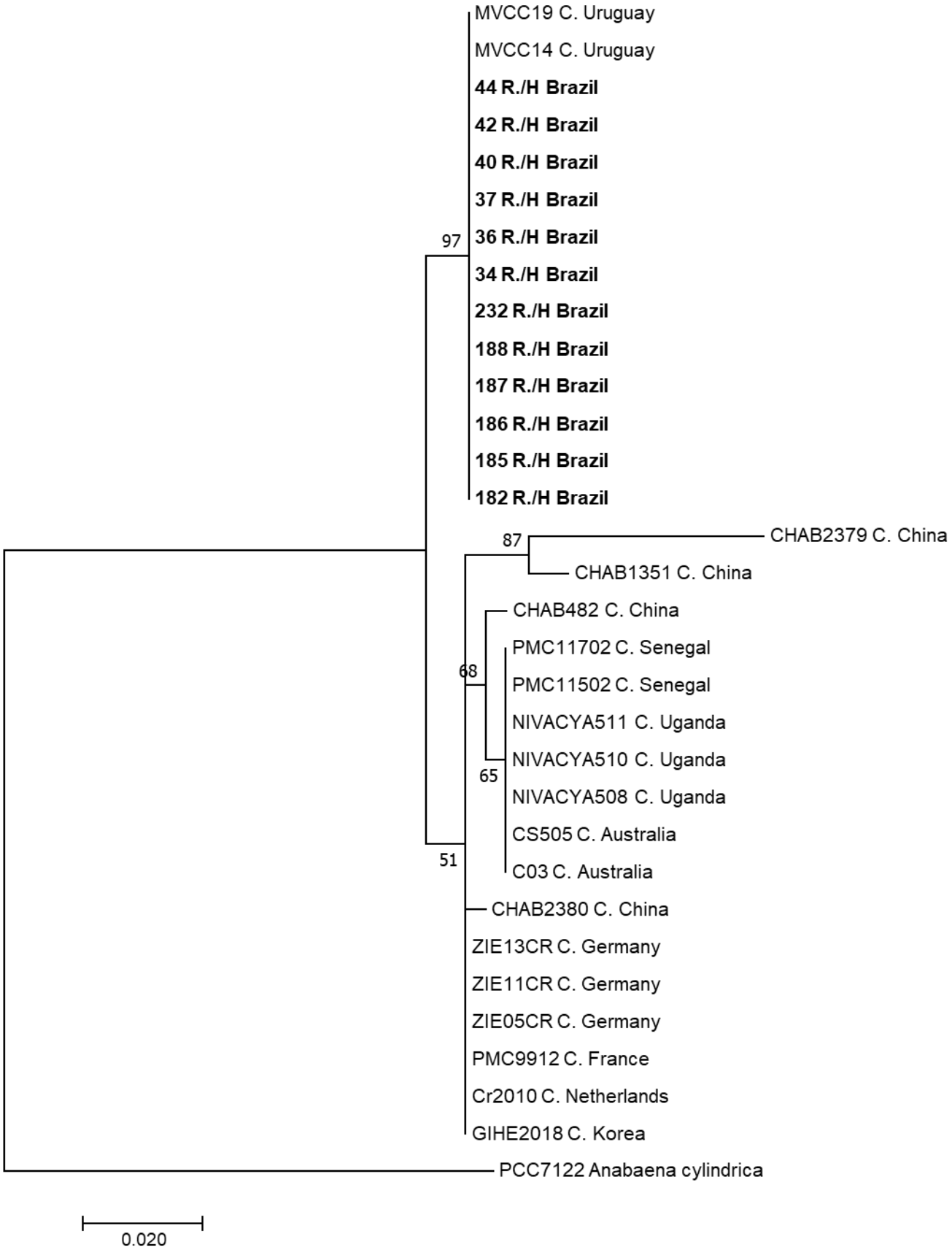
Fig.S2- Maximum likelihood phylogenetic analysis of *Raphidiopsis* / *Cylindrospermopsis* using ITS 16S-23S sequences. KT290325.1 *A.cylindrica* was selected as the outgroup taxon. Support values for the branches are indicated at the nodes. Strains sequenced in this study are indicated in bold.


181 R. Brazil
233 R. Brazil
183 R. Brazil
234 R. Brazil
ITEP005 R. Brazil
182 R./H Brazil
37 R./H Brazil
42 R./H Brazil
44 R./H Brazil
186 R./H Brazil
187 R./H Brazil
188 R./H Brazil
232 R./H Brazil
34 R./H Brazil
36 R./H Brazil
40 R./H Brazil
MVCC14 C. Uruguay
MVCC19 C. Uruguay
T3 C. Brazil
185 R./H Brazil
CHAB1353 R.m China
CHAB2314 R.m China
CHAB3405 R.c China
CHAB3408 R.m China
CHAB3413 R.c China
CHAB3415 R.c China
CHAB1150 R.c China
CHAB341 R.c China
CHAB3416 R.c China
C03 C. Australia
CHAB1351 C. China
CHAB2380 C. China
Cr2010 C. Netherlands
Cyl C. Vietnam
GIHE2018 C. Korea
LEGE00236 C. Portugal
LEGE08326 C. Portugal
LMECYA130 C. Portugal
NIVACYA399 C. Hungary
ZIE05CR C. Germany
ZIE11CR C. Germany
PMC9912 C. France
ZIE13CR C. Germany
CYP030B C. Australia
CS505 C. Australia
LEGE97047 C. Portugal
CHAB2379 C. China
CHAB482 C. China
NIVACYA508 C. Uganda
NIVACYA510 C. Uganda
NIVACYA511 C. Uganda
PMC11702 C. Senegal
PMC11502 C. Senegal
PCC7122 Anabaena cylindrica
91
83
87
83
37
41
100
99
57
64
63
86
0.020

1

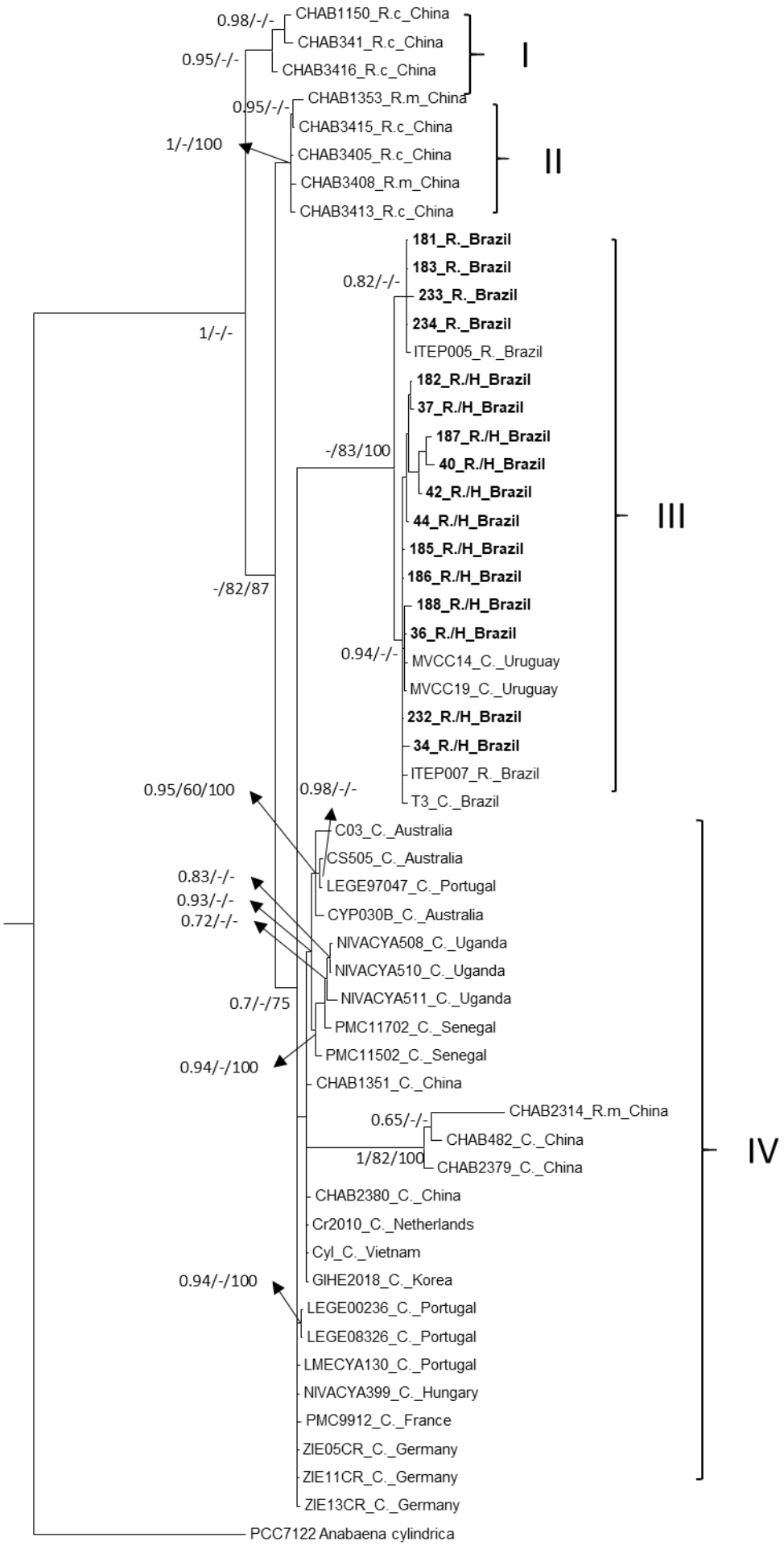
6. Fig.S3

Fig.S3- Maximum likelihood phylogenetic analysis of *Raphidiopsis* / *Cylindrospermopsis* using *rpoC1* sequences. KT290325.1 *A.cylindrica* was selected as the outgroup taxon. Support values for the branches are indicated at the nodes. Strains sequenced in this study are indicated in bold.


MVCC19 C. Uruguay
MVCC14 C. Uruguay
44 R./H Brazil
42 R./H Brazil
40 R./H Brazil
37 R./H Brazil
97
36 R./H Brazil
34 R./H Brazil
232 R./H Brazil
188 R./H Brazil
187 R./H Brazil
186 R./H Brazil
185 R./H Brazil
182 R./H Brazil
87
CHAB2379 C. China
CHAB1351 C. China
CHAB482 C. China
PMC11702 C. Senegal
68
PMC11502 C. Senegal
NIVACYA511 C. Uganda
NIVACYA510 C. Uganda
65
NIVACYA508 C. Uganda
CS505 C. Australia
51
C03 C. Australia
CHAB2380 C. China
ZIE13CR C. Germany
ZIE11CR C. Germany
ZIE05CR C. Germany
PMC9912 C. France
Cr2010 C. Netherlands
GIHE2018 C. Korea
PCC7122 Anabaena cylindrica
0.020

7. Fig.S4

Fig.S4- Maximum likelihood phylogenetic analysis of *Raphidiopsis* / *Cylindrospermopsis* using *nifH* sequences. KT290325.1 *A.cylindrica* was selected as the outgroup taxon. Support values for the branches are indicated at the nodes. Strains sequenced in this study are indicated in bold.


0.98/-/-
CHAB1150_R.c_China
CHAB341_R.c_China
0.95/-/-
CHAB3416_R.c_China
I
0.95/-/-
CHAB1353_R.m_China
CHAB3415_R.c_China
1/-/100
CHAB3405_R.c_China
CHAB3408_R.m_China
CHAB3413_R.c_China
II
181_R._Brazil
183_R._Brazil
0.82/-/-
233_R._Brazil
234_R._Brazil
1/-/-
ITEP005_R._Brazil
182_R./H_Brazil
37_R./H_Brazil
187_R./H_Brazil
-/83/100
40_R./H_Brazil
42_R./H_Brazil
44_R./H_Brazil
III
185_R./H_Brazil
186_R./H_Brazil
-/82/87
188_R./H_Brazil
36_R./H_Brazil
0.94/-/-
MVCC14_C._Uruguay
MVCC19_C._Uruguay
232_R./H_Brazil
34_R./H_Brazil
ITEP007_R._Brazil
0.95/60/100
0.98/-/-
T3_C._Brazil
C03_C._Australia
CS505_C._Australia
0.83/-/-
LEGE97047_C._Portugal
0.93/-/-
CYP030B_C._Australia
0.72/-/-
NIVACYA508_C._Uganda
NIVACYA510_C._Uganda
0.7/-/75
NIVACYA511_C._Uganda
PMC11702_C._Senegal
0.94/-/100
PMC11502_C._Senegal
CHAB1351_C._China
0.65/-/-
CHAB2314_R.m_China
CHAB482_C._China
1/82/100
CHAB2379_C._China
IV
CHAB2380_C._China
Cr2010_C._Netherlands
Cyl_C._Vietnam
0.94/-/100
GIHE2018_C._Korea
LEGE00236_C._Portugal
LEGE08326_C._Portugal
LMECYA130_C._Portugal
NIVACYA399_C._Hungary
PMC9912_C._France
ZIE05CR_C._Germany
ZIE11CR_C._Germany
ZIE13CR_C._Germany
PCC7122 Anabaena cylindrica
0.0100

1

8. Fig.S5

Fig.S5- Phylogenetic analysis of *Raphidiopsis* / *Cylindrospermopsis* strains derived from concatenated data using 16S-rRNA, ITS 16S-23S and *rpoC1* sequences using topology given by Bayesian analysis. KT290325.1 *A.cylindrica* was selected as the outgroup taxon. Support values (BI/ML/MP) for the branches are indicated at the nodes of the tree when bootstrap values are above 60% and BI is above 0.8. Strains sequenced in this study are indicated in bold.

CHAPTER THREE

1. Table S1

Protein Tree Name	Acession number
chlL_Microcystis	WP_002759152
SufS_Microcystis	WP_002796790
SufS_Nodularia1	WP_006194433
NifN_Nodularia	WP_006194547
NifE_Nodularia	WP_006194548
NifK_Nodularia	WP_006194550
NifD_Nodularia	WP_006194555
NifS_Nodularia1	WP_006194563
NifB_Nodularia	WP_006194566
NifS_Nodularia2	WP_006195150
NifH_Nodularia	WP_006195291
IscS_Nodularia	WP_006195727
NifS_Nodularia3	WP_006195758
SufS_Nodularia2	WP_006197168
CsdA_Nodularia	WP_006198655
CsdA_Cylindrospermopsis	WP_006276239
NifS_Cylindrospermopsis1	WP_006276272
SufS_Cylindrospermopsis	WP_006277484
chlL_Cylindrospermopsis	WP_006277886
NifS_Cylindrospermopsis3	WP_006278222
NifN_Cylindrospermopsis	WP_006278574
NifE_Cylindrospermopsis	WP_006278575
NifK_Cylindrospermopsis	WP_006278576
NifD_Cylindrospermopsis	WP_006278577
NifH_Cylindrospermopsis	WP_006278578
NifU_Cylindrospermopsis	WP_006278579
NifB_Cylindrospermopsis	WP_006278582
chlL_Arthrospira	WP_006616040
SufS_Arthrospira	WP_006617208
NifS_Arthrospira1	WP_006617890
IscS_Arthrospira	WP_006619733

SufS-like_Arthrospira	WP_006619904
SufS_Crocosphaera	WP_007303368
chlL_Crocosphaera	WP_007305348
NifB_Crocosphaera	WP_007305797
NifS_Crocosphaera1	WP_007305799
NifH_Crocosphaera	WP_007305800
NifD_Crocosphaera	WP_007305801
NifK_Crocosphaera	WP_007305802
NifE_Crocosphaera	WP_007305804
NifN_Crocosphaera	WP_007305805
CsdA_Crocosphaera	WP_007306176
NifS_Crocosphaera2	WP_007307242
NifS_Crocosphaera3	WP_007307470
NifS_Richelia1	WP_008231943
NifD_Richelia	WP_008231967
NifK_Richelia	WP_008231969
NifE_Richelia	WP_008231972
NifN_Richelia	WP_008231974
CsdA_Richelia	WP_008232764
NifS_Richelia2	WP_008233357
NifS_Richelia3	WP_008235065
NifS_Raphidiopsis1	WP_009342443
CsdA_Raphidiopsis	WP_009342482
chlL_Raphidiopsis	WP_009342543
SufS_Raphidiopsis	WP_009343202
CsdA_Fischerella1	WP_009454895
NifS_Fischerella1	WP_009455941
CsdA_Fischerella2	WP_009458960
NifN_Fischerella	WP_009459211
NifE_Fischerella	WP_009459212
NifK_Fischerella	WP_009459214
NifD_Fischerella	WP_009459215
NifH_Fischerella	WP_009459216
NifU_Fischerella	WP_009459217
NifS_Fischerella2	WP_009459218
chlL_Fischerella	WP_009460143
chlL_Pseudanabaena	WP_009625694
SufS_Pseudanabaena	WP_009628598
NifS_Pseudanabaena1	WP_009628609
NifS_Fischerella3	WP_009756654
NifN_Rhodopseudomonas1	WP_011156935
NifE_Rhodopseudomonas1	WP_011156936
NifH_Rhodopseudomonas1	WP_011156939
NifB_Rhodopseudomonas1	WP_011156940
VnfD_Rhodopseudomonas	WP_011156941
VnfK_Rhodopseudomonas	WP_011156943

SufS-like_Rhodopseudomonas1	WP_011156987
AnfK_Rhodopseudomonas	WP_011156998
AnfD_Rhodopseudomonas	WP_011157000
NifH_Rhodopseudomonas2	WP_011157001
BchX_Rhodopseudomonas	WP_011157083
chlL_Rhodopseudomonas	WP_011157104
MoaA_Rhodopseudomonas	WP_011157533
VnfN_Rhodopseudomonas1	WP_011157900
VnfE_Rhodopseudomonas1	WP_011157901
NifH_Rhodopseudomonas3	WP_011157906
NifB_Rhodopseudomonas2	WP_011157907
VnfE_Rhodopseudomonas2	WP_011157916
VnfN_Rhodopseudomonas2	WP_011157917
NifS_Rhodopseudomonas1	WP_011158013
SufS-like_Rhodopseudomonas2	WP_011158017
NifH_Rhodopseudomonas4	WP_011158165
NifE_Rhodopseudomonas2	WP_011158166
NifB_Rhodopseudomonas3	WP_011158184
NifH_Rhodopseudomonas5	WP_011158185
VnfE_Rhodopseudomonas3	WP_011158186
VnfN_Rhodopseudomonas3	WP_011158187
NifS_Rhodopseudomonas2	WP_011160140
NifU_Rhodopseudomonas	WP_011160141
NifN_Rhodopseudomonas2	WP_011160148
NifE_Rhodopseudomonas3	WP_011160149
NifK_Rhodopseudomonas	WP_011160150
NifD_Rhodopseudomonas	WP_011160151
NifH_Rhodopseudomonas6	WP_011160152
NifB_Rhodopseudomonas4	WP_011160162
chlL_Synechococcus	WP_011242451
CsdA_Synechococcus	WP_011242610
NifS_Synechococcus1	WP_011243862
NifS_Synechococcus2	WP_011244475
SufS_Synechococcus	WP_011244663
MoaA_Microcystis	WP_012264115
SufS-like_Microcystis	WP_012264804
NifS_Microcystis1	WP_012264870
CsdA_Microcystis1	WP_012267325
CsdA_Microcystis2	WP_012267653
NifS_Microcystis2	WP_012268040
NifN_Nostoc1	WP_012407195
NifE_Nostoc1	WP_012407196
NifK_Nostoc	WP_012407198
NifH_Nostoc1	WP_012407220

NifU_Nostoc	WP_012407222
NifS_Nostoc1	WP_012407223
NifB_Nostoc	WP_012407225
NifH_Nostoc2	WP_012407784
CsdA_Nostoc	WP_012408094
NifS_Nostoc2	WP_012409112
chlL_Nostoc	WP_012410753
SufS_Nostoc	WP_012411128
CsdA_Nostoc	WP_012411245
NifH_Nostoc3	WP_012412139
NifE_Nostoc2	WP_012412140
NifN_Nostoc2	WP_012412141
NifS_Nostoc3	WP_012412674
CsdA_Cyanothece	WP_012626206
MoaA_Cyanothece	WP_012626365
NifB_Cyanothece	WP_012627794
NifS_Cyanothece4	WP_012627795
NifU_Cyanothece1	WP_012627796
NifH_Cyanothece	WP_012627797
NifD_Cyanothece	WP_012627798
NifK_Cyanothece	WP_012627799
CsdA_Cyanothece2	WP_012628734
SufS_Cyanothece	WP_012628827
chlL_Cyanothece	WP_012629307
NifE_Cyanothece	WP_012630142
NifN_Cyanothece	WP_012630143
NifS_Cyanothece5	WP_012630361
NifH_Azobacter1	WP_012698831
NifD_Azobacter1	WP_012698832
NifK_Azobacter	WP_012698833
NifE_Azobacter1	WP_012698838
NifN_Azobacter1	WP_012698839
NifU_Azobacter	WP_012698853
NifS_Azobacter	WP_012698854
VnfK_Azobacter	WP_012698948
VnfD_Azobacter	WP_012698950
NifH_Azobacter2	WP_012698955
NifN_Azobacter2	WP_012698964
NifE_Azobacter2	WP_012698965
CsdA_Azobacter	WP_012700134
SufS-like_Azobacter1	WP_012700292
SufS-like_Azobacter2	WP_012700705
MinD_Azobacter	WP_012701597
SufS_Azobacter	WP_012702425
NifD_Azobacter2	WP_012702501
IscS_Azobacter	WP_012702551

AnfK_Azobacter	WP_012703359
AnfD_Azobacter	WP_012703361
NifH_Azobacter3	WP_012703362
NifB_Azobacter	WP_012703541
NifS_Arthrospira2	WP_014276333
CsdA_Arthrospira	WP_014277404
CsdA_Anabaena	WP_015078166
NifS_Anabaena1	WP_015079312
chlL_Anabaena	WP_015079916
IscS_Anabaena	WP_015080463
MoaA_Anabaena	WP_015080852
SufS_Anabaena	WP_015080909
NifB_Anabaena	WP_015081223
NifS_Anabaena2	WP_015081225
NifU_Anabaena	WP_015081226
NifK_Anabaena	WP_015081327
NifE_Anabaena	WP_015081328
CsdA_Pleurocapsa1	WP_015142986
IscS_Pleurocapsa	WP_015143132
chlL_Pleurocapsa	WP_015144620
CsdA_Pleurocapsa	WP_015144649
NifS_Pleurocapsa1	WP_015144696
NifB_Pleurocapsa1	WP_015145027
NifS_Pleurocapsa2	WP_015145029
NifU_Pleurocapsa1	WP_015145030
NifH_Pleurocapsa1	WP_015145031
NifD_Pleurocapsa1	WP_015145032
NifK_Pleurocapsa1	WP_015145033
NifE_Pleurocapsa1	WP_015145035
NifN_Pleurocapsa1	WP_015145036
NifB_Pleurocapsa2	WP_015145423
NifE_Pleurocapsa2	WP_015145424
NifD_Pleurocapsa2	WP_015145427
NifH_Pleurocapsa2	WP_015145430
NifU_Pleurocapsa2	WP_015145431
NifS_Pleurocapsa3	WP_015145432
CsdA_Pleurocapsa2	WP_015145697
SufS_Pleurocapsa	WP_015145952
CsdA_Oscillatoria1	WP_015148208
IscS_Oscillatoria	WP_015149768
NifS_Oscillatoria1	WP_015150747
SufS_Oscillatoria	WP_015151524
NifN_Calothrix	WP_015195966
NifK_Calothrix	WP_015195987
NifD_Calothrix	WP_015195999
NifU_Calothrix	WP_015196130

NifB_Calothrix	WP_015196133
CsdA_Calothrix1	WP_015196350
NifS_Calothrix1	WP_015196359
CsdA_Calothrix2	WP_015197228
chlL_Calothrix	WP_015199079
CsdA_Calothrix3	WP_015199516
NifS_Calothrix2	WP_015200226
SufS_Calothrix	WP_015200445
SufS_Crinalium	WP_015202049
CsdA_Crinalium	WP_015202101
NifS_Crinalium	WP_015203676
chlL_Crinalium	WP_015204277
CsdA_Fischerella	WP_016870876
SufS_Fischerella	WP_016871167
NifB_Fischerella	WP_016871834
SufS_Rhizobium	WP_025394708
NifS_Rhizobium1	WP_025394713
NifN_Rhizobium	WP_025397927
NifK_Rhizobium	WP_025397928
NifD_Rhizobium	WP_025397929
NifH_Rhizobium	WP_025397930
NifB_Rhizobium	WP_025397935
NifS_Rhizobium2	WP_025398073
IscS_Planktothrix	WP_026788021
chlL_Aphanizomenon	WP_027400713
IscS_Aphanizomenon	WP_027400972
SufS_Aphanizomenon	WP_027401844
CsdA_Aphanizomenon	WP_027402365
CsdA_Aphanizomenon2	WP_027403363
NifS_Aphanizomenon1	WP_027404062
NifN_Aphanizomenon	WP_027404273
NifE_Aphanizomenon	WP_027404274
NifK_Aphanizomenon	WP_027404275
NifU_Aphanizomenon	WP_027404310
NifS_Aphanizomenon2	WP_027404311
NifB_Aphanizomenon	WP_027404313
NifD_Aphanizomenon	WP_035083244
NifS_Cylindrospermopsis2	WP_040009945
NifS_Raphidiopsis2	WP_040554000
NifS_Pseudanabaena2	WP_040689494
IscS_Crinalium	WP_041226702
IscS_Cyanothece	WP_041237125
NifK_Pleurocapsa2	WP_041392380
NifN_Anabaena1	WP_041458126
NifD_Nostoc1	WP_041565131
NifS_Calothrix3	WP_041740145

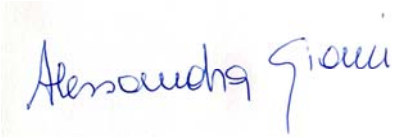
NifS_Planktothrix	WP_042152977
CsdA_Planktothrix1	WP_042153795
CsdA_Planktothrix2	WP_042154113
NifU_Nodularia	WP_042201636
VnfN_Rhodopseudomonas4	WP_042441092
NifS_Oscillatoria2	WP_044196545
NifB_Richelia	WP_044304347
NifU_Richelia	WP_044304348
CsdA_Scytonema1	WP_048866554
NifS_Scytonema1	WP_048867133
SufS_Scytonema1	WP_048867898
CsdA_Scytonema2	WP_048869496
NifB_Scytonema	WP_048869715
NifU_Scytonema	WP_048869717
NifH_Scytonema	WP_048869719
NifD_Scytonema	WP_048869720
NifK_Scytonema	WP_048869721
NifN_Scytonema	WP_048869737
NifS_Scytonema2	WP_048870147
NifE_Scytonema	WP_048870149
IscS_Scytonema	WP_048871366
SufS_Scytonema2	WP_048873723
NifD_Anabaena	WP_051155216
NifE_Rhizobium	WP_051449438
IscS_Phormidesmis	WP_068815317
CsdA_Phormidesmis1	WP_068816503
CsdA_Phormidesmis2	WP_068817012
chlL_Phormidesmis	WP_068817061
NifS_Phormidesmis1	WP_068817507
SufS_Phormidesmis	WP_068819288
NifS_Phormidesmis2	WP_068820037
chlL_Scytonema1	WP_071654239
chlL_Scytonema2	WP_071654348
chlL_Nodularia	WP_071839309
chlL_Oscillatoria	WP_071884374
CsdA_Phormidium	WP_073608222
NifS_Phormidium1	WP_073608572
NifS_Phormidium2	WP_073608600
SufS_Phormidium	WP_073609226
chlL_Phormidium	WP_073611316
chlL_Richelia	WP_081606214
NifN_Pleurocapsa2	WP_083888315
CsdA_Oscillatoria2	WP_083896854
CsdA_Anabaena	WP_085929767
CsdA_Pseudanabaena	WP_100228984
NifS_Thermosynechococcus1	NP_680906

NifS_Thermosynechococcus2	NP_681703
SufS_Thermosynechococcus	NP_682858
CsdA_Thermosynechococcus	NP_683117
chlL_Thermosynechococcus	NP_683137
SufS_Prochlorococcus	NP_874477
NifS_Prochlorococcus1	NP_874589
CsdA_Prochlorococcus	NP_874807
chlL_Prochlorococcus	NP_874937
chlL-B_Prochlorococcus	NP_874938
NifS_Prochlorococcus2	NP_875341
SufS-like_Gloeobacter	NP_924319
chlL_Gloeobacter	NP_925316
CsdA_Gloeobacter1	NP_925547
CsdA_Gloeobacter2	NP_926896
IscS_Gloeobacter	NP_927330

2. Fig.S1

Fig. S1- Phylogenetic relationships of Bacterial Nif. The expanded tree containing 320 protein sequences.

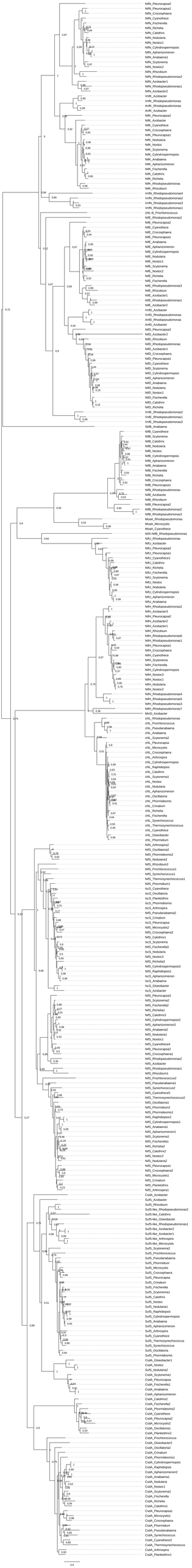
Dissertação defendida por Laísa Corrêa Braga Marques e aprovada, em 17 de dezembro de 2020, pela banca examinadora constituída pelos professores:



Dra. Alessandra Giani (orientadora)

Dra. Maria Kolman

Dr. Álvaro Cantini Nunes


NtN_Pleurocapsa2
NtN_Pleurocapsa1
NtN_Crocosphaera
NtN_Cyanothece
NtN_Fischerella
NtN_Richelia
NtN_Calothrix
NtN_Nodularia
NtN_Nostoc1
NtN_Cylindrospermopsis
NtN_Aphanizomenon
NtN_Anabaena1
NtN_Scytonema
NtN_Nostoc2
NtN_Rhizobium
NtN_Rhodopseudomonas2
NtN_Azobacter1
NtN_Rhodopseudomonas1
NtN_Azobacter2
VntN_Azobacter
VntN_Rhodopseudomonas
AntN_Rhodopseudomonas
AntN_Azobacter
NtP_Pleurocapsa2
NtN_Azobacter
NtN_Cyanothece
NtN_Crocosphaera
NtP_Pleurocapsa1
NtN_Nodularia
NtN_Nostoc
NtN_Scytonema
NtN_Cylindrospermopsis
NtN_Anabaena
NtN_Aphanizomenon
NtN_Fischerella
NtN_Calothrix
NtN_Richelia
NtN_Rhodopseudomonas
NtN_Rhizobium
VntN_Rhodopseudomonas4
VntN_Rhodopseudomonas2
VntN_Rhodopseudomonas3
VntN_Rhodopseudomonas1
chN_6_Prochlorococcus
NtE_Pleurocapsa2
NtE_Cyanothece
NtE_Crocosphaera
NtE_Pleurocapsa1
NtE_Anabaena
NtE_Aphanizomenon
NtE_Cylindrospermopsis
NtE_Nodularia
NtE_Nostoc1
NtE_Scytonema
NtE_Nostoc2
NtE_Richelia
NtE_Fischerella
NtE_Rhodopseudomonas3
NtE_Rhizobium
NtE_Azobacter1
NtE_Rhodopseudomonas1
NtE_Azobacter2
VntD_Azobacter
VntD_Rhodopseudomonas
AntD_Rhodopseudomonas
AntD_Azobacter
NtD_Pleurocapsa2
NtD_Azobacter2
NtD_Rhizobium
NtD_Rhodopseudomonas
NtD_Azobacter1
NtD_Crocosphaera
NtD_Pleurocapsa1
NtD_Cyanothece
NtD_Scytonema
NtD_Cylindrospermopsis
NtD_Aphanizomenon
NtD_Anabaena
NtD_Nodularia
NtD_Nostoc1
NtD_Fischerella
NtD_Calothrix
NtD_Richelia
VntE_Rhodopseudomonas2
VntE_Rhodopseudomonas4
VntE_Rhodopseudomonas3
SkN_Anabaena
NtB_Cyanothece
NtB_Scytonema
NtB_Calothrix
NtB_Nodularia
NtB_Nostoc
NtB_Cylindrospermopsis
NtB_Aphanizomenon
NtB_Anabaena
NtB_Fischerella
NtB_Richelia
NtB_Crocosphaera
NtB_Pleurocapsa1
NtB_Rhodopseudomonas
NtB_Azobacter
NtB_Rhizobium
NtB_Pleurocapsa2
NtB_Rhodopseudomonas2
NtB_Rhodopseudomonas3
MoaA_Rhodopseudomonas
MoaA_Microcystis
MoaA_Cyanothece
NtH,NtB_Rhodopseudomonas
NtU_Rhodopseudomonas
NtU_Azobacter
NtU_Pleurocapsa2
NtU_Pleurocapsa1
NtU_Cyanothece1
NtU_Calothrix
NtU_Richelia
NtU_Fischerella
NtU_Scytonema
NtU_Nostoc
NtU_Nodularia
NtU_Cylindrospermopsis
NtU_Aphanizomenon
NtU_Anabaena
NtH_Rhodopseudomonas2
NtH_Azobacter3
NtH_Pleurocapsa2
NtH_Azobacter2
NtH_Azobacter1
NtH_Rhizobium
NtH_Rhodopseudomonas6
NtH_Rhodopseudomonas1
NtH_Pleurocapsa3
NtH_Crocosphaera
NtH_Cyanothece
NtH_Scytonema
NtH_Fischerella
NtH_Cylindrospermopsis
NtH_Nostoc3
NtH_Nostoc1
NtH_Nodularia
NtH_Nostoc2
NtH_Rhodopseudomonas4
NtH_Rhodopseudomonas5
NtH_Rhodopseudomonas
NtH_Rhodopseudomonas7
MoO_Azobacter
chL_Rhodopseudomonas
chL_Prochlorococcus
chL_Pseudanabaena
chL_Anabaena
chL_Scytonema2
chL_Pleurocapsa
chL_Microcystis
chL_Crocosphaera
chL_Arthrospira
chL_Cylindrospermopsis
chL_Raphidiopsis
chL_Calothrix
chL_Scytonema1
chL_Nostoc
chL_Nodularia
chL_Aphanizomenon
chL_Oscillatoria
chL_Phormidesmis
chL_Cylindrium
chL_Richelia
chL_Fischerella
chL_Synechococcus
chL_Thermosynechococcus
chL_Cyanothece
chL_Gloeobacter
chL_Phormidium
NtS_Arthrospira2
NtS_Oscillatoria2
NtS_Phormidesmis2
NtS_Nodularia3
NtS_Rhizobium2
NtS_Prochlorococcus1
NtS_Synechococcus1
NtS_Thermosynechococcus1
NtS_Phormidium1
tscS_Cyanothece
tscS_Pleurocapsa
tscS_Oscillatoria
tscS_Plankothrix
tscS_Phormidesmis
tscS_Arthrospira
NtS_Pseudanabaena2
tscS_Cylindrium
tscS_Pleurocapsa
NtS_Microcystis2
NtS_Crocosphaera3
NtS_Calothrix1
tscS_Scytonema
NtS_Fischerella3
tscS_Nodularia
NtS_Nostoc3
NtS_Richelia3
NtS_Cylindrospermopsis3
NtS_Raphidiopsis1
tscS_Aphanizomenon
tscS_Anabaena
tscS_Gloeobacter
tscS_Azobacter
NtS_Pleurocapsa3
NtS_Scytonema2
NtS_Fischerella2
NtS_Richelia1
NtS_Calothrix3
NtS_Cylindrospermopsis2
NtS_Aphanizomenon2
NtS_Anabaena2
NtS_Nodularia1
NtS_Nostoc1
NtS_Cyanothece4
NtS_Pleurocapsa2
NtS_Crocosphaera1
NtS_Rhodopseudomonas2
NtS_Azobacter
NtS_Rhodopseudomonas3
NtS_Rhizobium1
NtS_Prochlorococcus2
NtS_Pseudanabaena1
NtS_Synechococcus2
NtS_Cyanothece5
NtS_Thermosynechococcus2
NtS_Oscillatoria1
NtS_Phormidium2
NtS_Phormidesmis1
NtS_Raphidiopsis2
NtS_Cylindrospermopsis1
NtS_Anabaena3
NtS_Aphanizomenon1
NtS_Scytonema3
NtS_Richelia2
NtS_Calothrix2
NtS_Nostoc2
NtS_Nodularia2
NtS_Pleurocapsa3
NtS_Crocosphaera2
NtS_Microcystis1
NtS_Cylindrium
NtS_Plankothrix
NtS_Arthrospira1
tscA_Azobacter
tscS_Azobacter
tscS_Rhizobium
tscS-like_Rhodopseudomonas2
tscS-like_Calothrix
tscS-like_Gloeobacter
tscS-like_Rhodopseudomonas1
tscS-like_Azobacter2
tscS-like_Azobacter1
tscS-like_Arthrospira
tscS-like_Microcystis
tscS_Scytonema2
tscS_Prochlorococcus
tscS_Pseudanabaena
tscS_Phormidium
tscS_Microcystis
tscS_Crocosphaera
tscS_Pleurocapsa
tscS_Cylindrium
tscS_Fischerella
tscS_Scytonema1
tscS_Calothrix
tscS_Nostoc
tscS_Nodularia2
tscS_Raphidiopsis
tscS_Cylindrospermopsis
tscS_Anabaena
tscS_Aphanizomenon
tscS_Arthrospira
tscS_Cyanothece
tscS_Thermosynechococcus
tscS_Synechococcus
tscS_Oscillatoria
tscS_Phormidesmis
tscA_Gloeobacter1
tscA_Nostoc
tscS_Nodularia2
tscA_Scytonema3
tscA_Pleurocapsa
tscA_Fischerella2
tscA_Anabaena
tscA_Aphanizomenon
tscA_Calothrix2
tscA_Fischerella1
tscA_Phormidesmis2
tscA_Cyanothece
tscA_Pleurocapsa2
tscA_Microcystis2
tscA_Oscillatoria1
tscA_Plankothrix2
tscA_Prochlorococcus
tscA_Gloeobacter2
tscA_Oscillatoria2
tscA_Cylindrium
tscA_Phormidesmis1
tscA_Cylindrospermopsis
tscA_Raphidiopsis
tscA_Aphanizomenon2
tscA_Anabaena1
tscA_Nodularia
tscA_Nostoc1
tscA_Scytonema2
tscA_Fischerella
tscA_Richelia
tscA_Calothrix1
tscA_Pleurocapsa1
tscA_Microcystis1
tscA_Crocosphaera
tscA_Phormidium
tscA_Pseudanabaena
tscA_Synechococcus
tscA_Cyanothece2
tscA_Thermosynechococcus
tscA_Arthrospira
tscA_Plankothrix1
0.9