

**UNIVERSIDADE DE SÃO PAULO
CENTRO DE ENERGIA NUCLEAR NA AGRICULTURA**

DANILLO OLIVEIRA DE ALVARENGA

**Análise genômica e funcional da cianobactéria *Nostoc* sp. CENA67
e caracterização da sua comunidade microbiana associada**

Piracicaba

2015

DANILLO OLIVEIRA DE ALVARENGA

**Análise genômica e funcional da cianobactéria *Nostoc* sp. CENA67
e caracterização da sua comunidade microbiana associada**

Tese apresentada ao Centro de Energia Nuclear na
Agricultura da Universidade de São Paulo para a
obtenção do título de Doutor em Ciências

Área de Concentração: Biologia na Agricultura e no
Ambiente

Orientadora: Profa. Dra. Marli de Fátima Fiore

Co-orientador: Prof. Dr. Artur Luiz da Costa da Silva

**Piracicaba
2015**

AUTORIZO A DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE

Dados Internacionais de Catalogação na Publicação (CIP)

Seção Técnica de Biblioteca - CENA/USP

Alvarenga, Danillo Oliveira de

Análise genômica e funcional da cianobactéria *Nostoc* sp. CENA67 e caracterização da sua comunidade microbiana associada / Danillo Oliveira de Alvarenga; orientadora Marli de Fátima Fiore; co-orientador Artur Luiz da Costa da Silva. - - Piracicaba, 2015.

115 f. : il.

Tese (Doutorado – Programa de Pós-Graduação em Ciências. Área de Concentração: Biologia na Agricultura e no Ambiente) – Centro de Energia Nuclear na Agricultura da Universidade de São Paulo.

1. Bactérias fixadoras de nitrogênio 2. Bioinformática 3. Ecologia microbiana 4. Genética bacteriana 5. Genomas 6. Metabolismo secundário 7. Microbiologia do solo 8. Terra preta antropogênica I. Título

CDU 579.262 : 575.111

AGRADECIMENTOS

À Profa. Dra. Marli de Fátima Fiore, pela orientação, suporte, conselhos e ensinamentos valiosos aos longos destes anos.

À Profa. Dra. Maria Paula Cruz Schneider e ao Prof. Dr. Artur Luiz da Costa da Silva, pelo apoio, pelas discussões e por compartilharem seu laboratório conosco.

Ao Prof. Dr. Robert Edwards, pela oportunidade de estagiar em seu laboratório.

Ao Prof. Dr. Alessandro de Mello Varani, por sua disponibilidade e ajuda indispensável nas análises genômicas.

Ao Prof. Dr. Augusto Etchegaray Jr., pelas discussões e colaborações.

À Dra. Janaina Rigonato, pelas lições, apoio e amizade.

Às(aos) colegas do Laboratório de Ecologia Molecular de Cianobactérias, Andresa Maíra da Fonseca, Ana Paula Dini Andreote, Bruno Costa Evangelista de Souza, Bruno Weiss, Caroline Hoff Riseti, Caroline Souza Pamplona da Silva, Elaine Crespim, Gabriela Machineski da Silva, Karina Heck da Silva, Maria Estela Stenico, Mariana Garcia Leal, Patricia Dayane Carvalho Schaker, Paula Cabral Tamasauskas, Stella Thomaz de Lima e Talita Schiavolin pela ajuda no laboratório, pelo companheirismo e pela amizade.

À Profa. Dra. Danielle Costa Carrara Couto, à Profa. Dra. Regiane Silva Kawasaki Francês e ao Prof. Dr. Rommel Tiago Jucá Ramos, pelo acolhimento e pelas lições de bioinformática.

A Genivaldo Gueiros Zacarias Silva e Fabyano Alvares Cardoso Lopes, pelo auxílio nas análises e pelas trocas de informações.

À Profa. Dra. Celia Regina Montes, pela ajuda na microscopia eletrônica de varredura.

A Marília Ribeiro Garcia Henyei, pela revisão e correção da formatação da tese.

Ao Programa de Pós-Graduação do Centro de Energia Nuclear na Agricultura e a toda a equipe deste Centro, por todo o auxílio fornecido.

À Fundação de Amparo à Pesquisa do Estado de São Paulo, pelo financiamento da pesquisa e pela bolsa de estudos.

À minha família, a maior responsável pela pessoa que hoje sou, que, mesmo à distância, nunca deixou de se mostrar presente.

A todas as pessoas que de alguma forma contribuíram direta ou indiretamente para a realização deste trabalho.

A visão da evolução como uma competição sangrenta crônica entre indivíduos e espécies, uma distorção popular da noção darwiniana de “sobrevivência do mais apto”, se dissolve perante uma nova visão de cooperação contínua, forte interação e interdependência entre as formas de vida. A vida não dominou o globo pelo combate, mas pela criação de redes. Formas de vida se multiplicaram e se tornaram complexas cooptando outras, não apenas matando-as.

Lynn Margulis

RESUMO

ALVARENGA, D. O. **Análise genômica e funcional da cianobactéria *Nostoc* sp. CENA67 e caracterização da sua comunidade microbiana associada.** 2015. 115 f. Tese (Doutorado) – Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Piracicaba, 2015.

Nostoc é um gênero cianobacteriano com distribuição ubíqua que tem importância em diversos ecossistemas. Contudo, poucos genomas estão atualmente disponíveis para esse gênero. Enquanto *Nostoc* spp. são as cianobactérias mais comumente relatadas em relações simbióticas com fungos, animais, plantas e outros organismos, associações com outros micro-organismos não receberam atenção similar. Como consequência das fortes interações entre cianobactérias e heterótrofos, culturas não axênicas são geralmente obtidas no isolamento dessas bactérias, o que proporciona uma oportunidade interessante para o desenvolvimento tanto de estudos genômicos quanto metagenômicos. Este trabalho teve como objetivo investigar as características genômicas e funcionais da linhagem *Nostoc* sp. CENA67, isolada de terra preta antropogênica, bem como estudar sua comunidade associada. Para esse fim, células de uma cultura não axênica de *Nostoc* sp. CENA67 foram sequenciadas com as plataformas MiSeq e Ion PGM e analisados com ferramentas genômicas e metagenômicas. A linhagem CENA67 de fato pertence à família Nostocaceae e possui algumas características em comum com cianobactérias do gênero *Nostoc*, porém diverge em certos aspectos morfológicos e filogenéticos do grupo típico de *Nostoc*, sugerindo que seja representante de um novo táxon. Além disso, seu genoma apresenta diferenças em relação aos genomas atualmente disponíveis para cianobactérias relacionadas ao gênero. A mineração desse genoma revelou 31 agrupamentos gênicos hipoteticamente relacionados à síntese de metabólitos secundários, a maioria dos quais não mostrou similaridade significativa com agrupamentos conhecidos. A análise de um agrupamento gênico de microviridina desvendou uma maior diversidade de genes para precursores dessa molécula do que se acreditava anteriormente, sugerindo que um número considerável de variantes ainda está a ser descoberta. A análise taxonômica da comunidade associada confirmou a dominância de cianobactérias na cultura, mas também revelou a presença de grande número de gêneros microbianos que normalmente são capazes de fixar nitrogênio atmosférico e estabelecer simbiose com plantas, incluindo *Mesorhizobium*, *Sinorhizobium* e *Starkeya*, entre outros. Rascunhos genômicos foram obtidos para *Bradyrhizobium diazoefficiens*, *Bradyrhizobium japonicum*, *Burkholderia lata* e *Hyphomicrobium nitrativorans*. Todavia, genes para fixação de nitrogênio não foram detectados nesses genomas, apesar de serem encontrados no genoma da cianobactéria e no metagenoma da comunidade, o que sugere que algumas populações podem estar sob pressão de seleção para a perda da capacidade de fixação de nitrogênio, provavelmente devido a este nutriente estar sendo fornecido pelo organismo mais abundante nesta comunidade, a cianobactéria. A análise funcional indicou vias exclusivas tanto à cianobactéria quanto à comunidade associada, e sugeriu a complementariedade de certos metabolismos. Os resultados possibilitam o aumento do conhecimento sobre a diversidade molecular e química do filo Cyanobacteria e levantam possíveis interações com micro-organismos simbiotes.

Palavras-chave: Consórcios microbianos. Microbiologia do solo. Interações microbianas. Metabolismo secundário. Microviridina.

ABSTRACT

ALVARENGA, D. O. **Genomic and functional analysis of the cyanobacterium *Nostoc* sp. CENA67 and characterization of its associated microbial community**. 2015. 115 f. Tese (Doutorado) – Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Piracicaba, 2015.

Nostoc is a cyanobacterial genus with ubiquitous distribution that is important in several ecosystems. However, few genomes are currently available for this genus. While *Nostoc* spp. are the most commonly reported cyanobacteria in symbiotic relationship with fungi, animals, plants, and other organisms, associations with other microorganisms have not received similar attention. As a consequence of tight interactions between cyanobacteria and heterotrophs, non-axenic cultures are usually achieved in the isolation of these bacteria, which provides an interesting opportunity for carrying out both genomic as metagenomic studies. This work aimed to investigate the genomic and functional characteristics of the strain *Nostoc* sp. CENA67, isolated from anthropogenic dark earth, and to study its associated community. For this purpose, cells from a non-axenic culture of *Nostoc* sp. CENA67 were sequenced with the platforms MiSeq and Ion PGM and analyzed with genomic and metagenomic tools. The strain CENA67 indeed belongs to the family Nostocaceae and presents some characteristics in common with cyanobacteria of the genus *Nostoc*, but diverges in certain morphological and phylogenetic aspects of the typical *Nostoc* group, suggesting that it is a representative of a new taxon. In addition, its genome presents differences in relation to the genomes currently available for cyanobacteria related to this genus. Genome mining revealed 31 gene clusters hypothetically related to the synthesis of secondary metabolites, most of which did not show significant similarity to known clusters. The analysis of a microviridin gene cluster unveiled a larger diversity of precursor genes for this molecule than was previously believed, suggesting that a considerable number of variants is still to be found. The taxonomic analysis of the associated community confirmed the dominance of cyanobacteria in the culture, but also revealed the presence of a great number of microbial genera that are usually capable of fixing atmospheric nitrogen and establishing symbiosis with plants, including *Mesorhizobium*, *Sinorhizobium*, and *Starkeya*, among others. Genomic drafts were obtained for *Bradyrhizobium diazoefficiens*, *Bradyrhizobium japonicum*, *Burkholderia lata*, and *Hyphomicrobium nitrativorans*. Nevertheless, genes for nitrogen fixation were not detected in these genomes, despite being found in the cyanobacterial genome and the community metagenome, suggesting that some populations might be under selection pressure for the loss of the ability to fix nitrogen, probably due to this nutrient being provided for the most abundant organism in this culture, the cyanobacterium. Functional analysis indicated pathways exclusive both to the cyanobacterium as to the associated community, and suggested the complementarity of certain metabolisms. These results allow the increase of the knowledge about the molecular and chemical diversity of the phylum Cyanobacteria and raise possible interactions with symbiotic microorganisms.

Keywords: Microbial consortia. Soil microbiology. Microbial interactions. Secondary metabolism. Microviridin.

SUMÁRIO

1. INTRODUÇÃO	13
2. REVISÃO DA LITERATURA	15
2.1 O gênero <i>Nostoc</i>	15
2.2 Associações ecológicas entre cianobactérias e outros organismos	18
2.3 Produtos naturais cianobacterianos	22
2.4 Genômica cianobacteriana	24
2.5 Terra preta antropogênica	29
3. HIPÓTESES	31
4. OBJETIVOS	32
4.1 Objetivo geral	32
4.2 Objetivos específicos	32
5. MATERIAL E MÉTODOS	33
5.1 Origem da amostra	33
5.2 Obtenção de biomassa	33
5.3 Caracterização morfológica	34
5.4 Extração de DNA	34
5.5 Reconstrução filogenética	35
5.6 Sequenciamento de DNA	37
5.7 Pré-processamento de sequências e montagem	37
5.8 Anotação e detecção de produtos naturais	38
5.9 Análises comparativas	38
5.10 Diversidade taxonômica e funcional	39
6. RESULTADOS E DISCUSSÃO	40
6.1 Caracterização morfológica e filogenética	40

6.2 Sequenciamento, montagem e anotação genômica	44
6.3 Detecção de metabólitos secundários	50
6.4 Análises genômicas comparativas	59
6.5 Diversidade taxonômica de micro-organismos associados	71
6.6 Caracterização funcional da comunidade associada	79
7. CONCLUSÕES	88
REFERÊNCIAS	89

1. INTRODUÇÃO

Novas tecnologias de sequenciamento de ácidos desoxirribonucleicos foram disponibilizadas para a comunidade científica a partir da década de 2000. Essas inovações tecnológicas se baseiam em metodologias que permitem a análise de um volume muito maior de nucleotídeos que o possível com o tradicional método de Sanger, que se manteve como o método-padrão de sequenciamento por mais de três décadas. Com as novas gerações de sequenciadores de DNA, a realização de projetos envolvendo genômica e metagenômica se tornou menos onerosa e mais rápida. Conseqüentemente, projetos desse gênero passaram a ter maior acessibilidade.

A bioinformática tem buscado acompanhar esse avanço, disponibilizando um número considerável de programas computacionais para a análise do crescente volume de informação que se torna possível pelas novas plataformas. Contudo, a despeito dos avanços na área, vários obstáculos ainda permanecem carentes de superação. A genômica cianobacteriana é um campo que ainda tem caminhado a passos relativamente lentos, consequência de uma série de peculiaridades inerentes ao estudo desses micro-organismos. Isso resulta em um número de genomas cianobacterianos ainda diminuto quando em comparação com a enorme variedade de sequências genômicas atualmente disponíveis para bactérias de outros filós.

Como produtores primários, os organismos do filo Cyanobacteria interagem ecologicamente tanto com eucariotos quanto com bactérias e arqueias, podendo estabelecer fortes associações com estes organismos. Em decorrência disso, o isolamento completo de linhagens cianobacterianas raramente é bem sucedido. As dificuldades de obtenção de culturas puras de cianobactérias e os desafios técnicos inerentes ao estudo de culturas mistas são fatores que contribuem fortemente para a complexidade desse tema.

Apesar disso, um número cada vez maior de grupos de pesquisa ao redor do mundo vêm buscando adentrar esse campo emergente. Uma tendência que pode ser claramente observada nos projetos que têm sido conduzidos nos anos mais recentes é o reconhecimento da necessidade de se estudar cianobactérias de grupos taxonômicos cuja diversidade ainda é pouco compreendida e/ou que provenham de ambientes pouco explorados. Os biomas brasileiros são uma ótima fonte de micro-organismos com essas características, já que estimativas apontam naqueles a existência de habitats com alta diversidade microbiana, cuja maior parte segue desconhecida.

Entre as cianobactérias de maior destaque, encontra-se *Nostoc*, um gênero de ampla distribuição mundial que apresenta importância ecológica, evolutiva, biogeoquímica, biotecnológica e ecotoxicológica em muitos ecossistemas terrestres e aquáticos. Os organismos desse gênero disponibilizam carbono e nitrogênio fixados em diversos habitats, e várias de suas espécies já foram relatadas em simbiose com outros organismos. Esse táxon apresenta grande diversidade metabólica, com várias moléculas bioativas tendo sido isoladas e caracterizadas a partir de extratos de linhagens pertencentes a esse gênero. Todavia, o gênero é polifilético e sua diversidade intergenérica é, no momento, bastante confusa, fato que também traz importância taxonômica ao seu estudo.

CENA67 é uma linhagem cianobacteriana proveniente da região amazônica brasileira. Isolada a partir de terra preta antropogênica como um organismo de vida livre, CENA67 foi identificada por métodos clássicos no gênero *Nostoc*. Como isolados de *Nostoc* spp. obtidos a partir de solos da Amazônia costumam se agrupar em clados distintos em árvores filogenéticas, essa linhagem provavelmente apresenta diferenças significativas em relação às outras cianobactérias do gênero. Dessa forma, o estudo do genoma dessa cianobactéria possibilita o aumento do conhecimento sobre a diversidade do filo Cyanobacteria, bem como uma melhor compreensão a respeito de seu metabolismo secundário e de sua capacidade de estabelecer interações ecológicas.

2. REVISÃO DE LITERATURA

2.1 O gênero *Nostoc*

Nostoc Vaucher ex Born. et Flah. é o gênero-tipo da família Nostocaceae. Essa família agrupa cianobactérias filamentosas heterocitadas que são caracterizadas por tricomas isopolares obrigatoriamente unisseriados, com células que se dividem perpendicularmente ao eixo do tricoma, e que não apresentam ramificação ou zonas meristemáticas (KOMÁREK; ANAGNOSTIDIS, 1989; KOMÁREK, 2010a; KOMÁREK, 2013). Os gêneros dessa família são tradicionalmente distinguidos entre si pela morfologia de filamentos e de células vegetativas, posição de heterócitos e tipo de formação de acinetos, bem como pela forma da colônia, formato das células terminais, presença de bainha ou aerótopos e ciclo de vida (RAJANIEMI et al., 2005).

Membros do gênero *Nostoc* são tradicionalmente diferenciados do restante da família por sua morfologia metamérica, com formação apo-heterocítica de acinetos, e pela alta complexidade em seu ciclo de vida. Alguns membros desse gênero são capazes de produzir colônias gelatinosas macroscópicas que podem ser visualizadas como mantos finos na superfície de plantas aquáticas e de substratos submersos ou como esferas ocas de até vários centímetros, que podem protegê-los contra o ambiente (SAND-JENSEN, 2014). Seu complexo ciclo de vida, caracterizado por estágios iniciais com tricomas encerrados dentro de bainhas mucilaginosas esféricas, heterócitos polares, um tipo especial de germinação de acinetos e produção de hormogônios com motilidade, não é encontrado em nenhum dos demais gêneros da família (KOMÁREK; ANAGNOSTIDIS, 1989; KOMÁREK, 2013).

A ubiquidade do gênero é bastante marcante, tendo sido encontrados representantes do gênero *Nostoc* em todos os continentes. Além de observadas em diversos ecossistemas americanos, africanos, asiáticos, europeus e oceânicos, incluindo solos de deserto, pântanos, estuários, arrozais e fontes termais, colônias destas cianobactérias foram detectadas em regiões extremamente frias e secas de vales da Antártica e do Ártico. Nos anos 1960, o gênero também foi relatado a 1.000 m de profundidade no Oceano Índico e no Mar Mediterrâneo, porém este dado permanece questionável. Relatos de *Nostoc* spp. em ambiente aquático indicam quase exclusividade para água doce, geralmente com preferência por pH alcalino (POTTS, 2000).

Diversas colônias de *Nostoc* spp. podem adquirir proporções macroscópicas e exalar odores característicos de terra molhada. Consequentemente, elas têm despertado interesse humano ao longo dos séculos. Seu primeiro relato conhecido data de por volta de 1.600 AP, na China, durante a Dinastia Jin (séculos 4 e 5 EC). Durante essa época, o médico e alquimista Hong Ge introduziu colônias de *N. commune* na alimentação do Imperador Yuan, Rui Sima, que as batizou de *Ge-Xian-Mi* (“arroz do imortal Ge”). Colônias secas eram conhecidas popularmente na China da Dinastia Jin como *Fa-Cai* (“cabelo vegetal”). Na Idade Média, acreditava-se que estas colônias tinham origem sobrenatural ou que caíam do céu ao solo, e isto lhes rendeu em alguns locais da Alemanha, da França, do Reino Unido, da Escandinávia e das Filipinas denominações populares equivalentes a “flor do sol”, “estrela cadente”, “geleia das bruxas” e “manteiga das fadas”, entre outras. Outras superstições as associavam à mítica “pedra filosofal”. Esferas coloniais de *Nostoc* continuam a ser consumidas como alimento nos dias atuais, especialmente por algumas populações chinesas, equatorianas, fijianas, japonesas, javanesas, mexicanas, mongóis, peruanas, siberianas e tailandesas (POTTS, 1997; 2000).

Foi no século 15 EC, mais de 200 anos antes de Leeuwenhoek descrever “animálculos” em um microscópio rudimentar pela primeira vez, que o suíço Aureolus Philippus Theophrastus Bombastus von Hohenheim, conhecido como Paracelsus, cunhou o nome pelo qual esses organismos são ainda hoje conhecidos. *Nostoc* é o resultado de um jogo de palavras que une a palavra inglesa *Nosthryl* à alemã *Nasenloch* – ambas significando “narina”. Paracelsus optou por esse nome devido à semelhança visual que acreditava haver entre secreções nasais e o característico polissacarídeo extracelular frequentemente encontrado em colônias dessa cianobactéria (POTTS, 1997).

Apesar de árvores filogenéticas construídas com sequências de genes de RNAr 16S apontarem o agrupamento monofilético de grande parte dos representantes deste gênero, suas distâncias evolutivas sugerem que este clado contenha mais de um gênero (RAJANIEMI et al., 2005). A polifilia deste gênero não é uma descoberta recente, já que, em 1931, Elenkin já havia dividido o gênero *Nostoc* em quatro gêneros distintos: *Amorphonostoc*, *Nematonostoc*, *Sphaeronostoc* e *Stratonostoc*, tomando como critério principal a forma final das colônias; todavia, esta divisão não foi disseminada, e atualmente *Stratonostoc* é o único aceito dentre estes (KOMÁREK; ANAGNOSTIDIS, 1989; KOMÁREK, 2010b). Além disso, com exceção de um clado que consiste de uma mistura de isolados simbiotes e de vida livre, a maior parte das linhagens de cianobactérias atualmente classificadas no gênero *Nostoc* que vivem em associação simbiótica se organiza filogeneticamente em clados distintos dos membros de vida

livre (SVENNING; ERIKSSON; RASMUSSEN, 2005). Tentativas de dividir este táxon continuam a ser realizadas e novos gêneros vêm sendo descritos a partir de isolados inicialmente identificados morfológicamente como pertencentes a esse táxon (GENUÁRIO et al., 2015; HROUZEK et al., 2013; ŘEHÁKOVÁ et al., 2007), mas uma revisão profunda do gênero ainda permanece por ser feita.

Devido à sua versatilidade metabólica, que inclui não apenas a habilidade de fixação de carbono atmosférico, mas também de nitrogênio, cianobactérias do gênero *Nostoc* são importantes membros da comunidade microbiana em diversos habitats, incluindo solos. No solo, *Nostoc* spp. podem regular o pH, aumentar a concentração de carbono e nitrogênio, reduzir a quantidade de sódio em ambientes de alta salinidade, atuar como antagonistas de fitopatógenos, insetos e nematoides, degradar agrotóxicos e reduzir a contaminação com metais pesados como cobre e cromo (SINGH; PANDEY; SINGH, 2011).

Nostoc spp. isoladas de solo ou sedimento de áreas tropicais costumam estar relacionadas filogeneticamente, e há a possibilidade de que representem uma nova unidade taxonômica (PAPAEFTHIMIOU et al., 2008). Predominantemente terrestres, colônias desse gênero são dotadas de uma notável capacidade de resistência à dessecação e de rápida retomada do metabolismo após a reidratação, características essenciais para organismos que não possuem acesso a reservas de umidade do subsolo e/ou que podem estar sujeitos a resfriamento (DODDS; GUDDER; MOLLENHAUER, 1995).

A exemplo do observado a nível de gênero, a diversidade intergenérica neste táxon é bastante heterogênea e continua sem ser resolvida, especialmente quando se considera as espécies que são classificadas dentro do gênero, mas que não formam colônias macroscópicas. Apesar de as espécies exclusivamente microscópicas desempenharem diferentes papéis de importância ecológica na biosfera, há poucos estudos sobre elas, cujo número pode ser maior que o que está atualmente descrito (MATEO et al., 2011). Mais de 200 espécies já foram descritas dentro do gênero *Nostoc*; porém, com o conhecimento atual, apenas cerca de 40 espécies são reconhecidas, as quais provavelmente constituem pelo menos dois subgêneros (KOMÁREK; HAUER, 2013). Logo, diversidade taxonômica, genômica e química maior do que é atualmente conhecida é provável de ser encontrada em cianobactérias relacionadas a este gênero.

2.2 Associações ecológicas entre cianobactérias e outros organismos

O conjunto de capacidades anabólicas das cianobactérias as leva a desempenhar importantes papéis ecológicos em diversos ambientes e a estabelecer interações mutualistas com um grande número de organismos. Simbioses epibióticas ou endobióticas entre cianobactérias e eucariotos como animais (ascídios, vermes equiurídeos, larvas de nematóceros, esponjas), cromalveolados (diatomáceas, dinoflagelados), fungos (líquens, *Geosiphon*) e plantas (antocerotófitas, cicadófitas, hepáticas, musgos, *Azolla*, *Gunnera*) são bem documentadas (ADAMS, 2000; ADAMS et al., 2013).

Além de conter cianobactérias de vida livre, *Nostoc* é o gênero cianobacteriano mais frequentemente relatado em associações ecológicas, normalmente contribuindo com nitrogênio biodisponível em interações simbióticas (BERGMAN; RAI; RASMUSSEN, 2007). Em amostras isoladas de solos e sedimentos de várzea e igapó na bacia amazônica brasileira, espécies conhecidas e potencialmente novas de *Nostoc* foram apontadas como sendo as mais abundantes cianobactérias fixadoras de nitrogênio (FIORE et al., 2005). Além da maior abundância, alguns desses isolados também apresentavam as maiores taxas de fixação de nitrogênio dentre os isolados obtidos no trabalho citado.

Diversas linhagens de *Nostoc* spp. estabelecem associações simbióticas com plantas, tais como as briófitas *Blasia* spp. e *Anthoceros* spp., a angiosperma *Gunnera* spp. e cicadófitas (CHAPMAN; MARGULIS, 1998). Devido às numerosas vantagens que essas associações apresentam, a criação de simbioses artificiais entre cianobactérias e plantas que não interagem naturalmente é um assunto que tem recebido atenção (AHMED; STAL; HASNAIN, 2010; GUSEV et al., 2003). *Nostoc* spp. são também os mais comuns simbiontes cianobacterianos em líquens (RIKKINEN, 2009). Em solo, uma espécie de *Nostoc* foi encontrada habitando dentro de uma “bexiga” multinucleada unicelular nas pontas das hifas de *Geosiphon pyriformis*, uma micorriza arbuscular, possivelmente fornecendo carbono e nitrogênio fixados ao hospedeiro (SCHÜBLER, 2002). Colônias de algumas espécies podem, ainda, estabelecer mutualismo com larvas de dípteros do gênero *Cricotopus* (SABATER; MUÑOZ, 2000). Um dos fatores que podem estar relacionados com a capacidade de simbiose nesse táxon é uma intensa produção de hormogônios (PAPAEFTHIMIOU et al., 2008).

Enquanto associações entre cianobactérias e eucariotos foram bastante estudadas, associações entre cianobactérias e outras bactérias ou arqueias, por outro lado, não receberam atenção similar. Não obstante, diversos micro-organismos heterotróficos podem se beneficiar de associações com bactérias capazes de fotossíntese oxigênica, de fixação de nitrogênio e,

ocasionalmente, de produzir determinados metabólitos secundários, características encontradas em um número considerável de cianobactérias. De fato, cianobactérias estabelecem associações simbióticas com micro-organismos heterotróficos, incluindo alguns em firme contato com seu envelope celular ou mesmo dentro de seu glicocálice (ZHUBANOVA et al., 2013) (Figura 1). Biofilmes dominados por cianobactérias possuem alta qualidade nutricional e podem sustentar grande biomassa de consumidores primários (NAGARKAR et al., 2004; YAMAMURO, 1999).

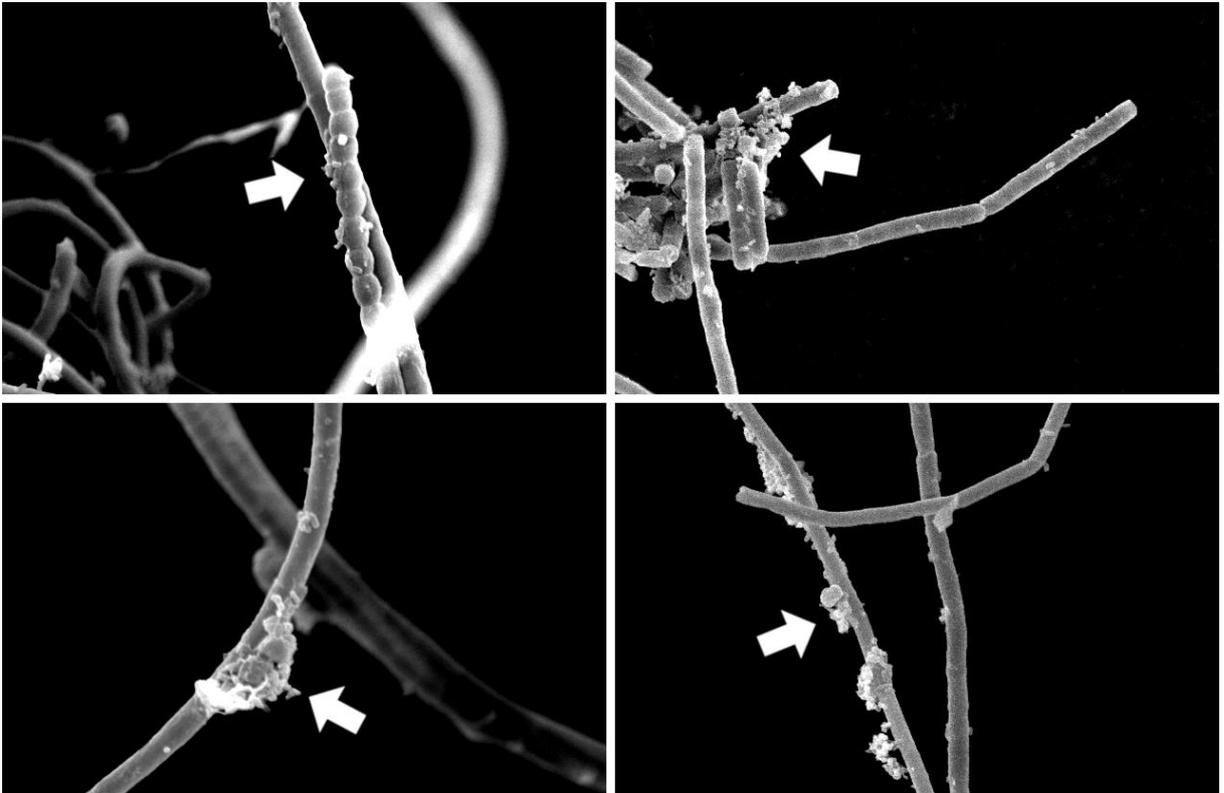


Figura 1 – Micrografias de uma cultura unicianobacteriana não axênica obtidas por microscopia eletrônica de varredura, onde células de outros micro-organismos são observadas em filamentos da cianobactéria cultivada. As setas apontam exemplos de agregados de células microbianas aderidas aos filamentos

Associações cianobacterianas com outros micro-organismos foram estimadas entre as mais antigas formas de vida conhecidas, fornecendo importantes interações ecológicas por bilhões de anos. Exopolissacarídeos secretados por cianobactérias são constantemente colonizados por micro-organismos, e podem dar origem a biofilmes e mantos microbianos complexos e com alta riqueza e abundância, onde autótrofos e heterótrofos encontram amplo espaço para interagir (COLE et al., 2014; PAERL; PINCKNEY; STEPPE, 2000).

Como observado em estromatólitos, mantos dominados por cianobactérias em diversos ecossistemas podem suportar comunidades microbianas altamente desenvolvidas e com interações bastante complexas (COHEN; GUREVITZ, 2006). De maneira semelhante, a reprodução cianobacteriana massiva que comumente surge em resposta à eutrofia em corpos d'água – um fenômeno conhecido como “floração cianobacteriana” – pode também ser acompanhada por associações com alta diversidade de bactérias heterotróficas, com várias destas sendo capazes de estimular o crescimento cianobacteriano (BERG et al., 2009).

Em decorrência das interações entre cianobactérias e outros micro-organismos, bactérias heterotróficas podem rapidamente se sobrepor às células da cianobactéria-alvo durante seu processo de isolamento; conseqüentemente, a purificação se torna um processo bastante complexo e dispendioso (WATERBURY, 2006). Apesar de diversas técnicas de purificação de cianobactérias compreendendo tanto separação física quanto tratamentos químicos terem sido publicadas, culturas cianobacterianas axênicas continuam sendo muito difíceis de ser alcançadas, já que a maioria dos métodos é específica a uma determinada linhagem e possui baixa taxa de sucesso (CHOI et al., 2008; SENA et al., 2008). Portanto, micro-organismos heterotróficos recalcitrantes são uma preocupação constante durante a purificação de culturas cianobacterianas, uma situação que é potencialmente mais complicada para linhagens de ambientes de alta diversidade microbiana, tais como solos.

Em alguns casos, uma cultura axênica é considerada virtualmente impossível de ser obtida devido às fortes associações entre cianobactérias e micro-organismos simbiotes. Algumas técnicas, tais como lavagem e centrifugação de células, são capazes de reduzir o número de micro-organismos não fixados que sejam menores ou maiores que as células-alvo (SENA et al., 2008; VÁZQUEZ-MARTÍNEZ et al., 2004), mas não podem remover micro-organismos com fortes conexões às bainhas cianobacterianas. Além disso, a substância fibrosa formada por carboidratos encontrada em mucilagens de células cianobacterianas pode estabelecer firmes agregados com outros micro-organismos, resultando em fortes ligações, capazes até mesmo de protegê-los da ação de antibióticos (VÁZQUEZ-MARTÍNEZ, 2004).

Conseqüentemente, a maioria dos esforços convencionais de isolamento cianobacteriano termina com culturas não axênicas, consistindo de consórcios microbianos compostos primariamente por cianobactérias e alguns micro-organismos proximamente associados. Apesar de essa condição poder ser vista como uma desvantagem, micro-organismos em consórcio podem agir sinergicamente e se tornar mais eficientes que culturas axênicas em alguns processos, ou podem desempenhar funções complexas, com passos múltiplos, que se tornam possíveis somente com o cocultivo de populações distintas

(BRENNER; YOU; ARNOLD, 2008). O estudo de genomas cianobacterianos através do sequenciamento de culturas não axênicas também é potencialmente capaz de otimizar o tempo de pesquisa, já que, ao contornar a etapa de purificação, o tempo necessário para a obtenção de resultados é drasticamente reduzido.

Como mutualismo, competição, predação/parasitismo, comensalismo, amensalismo e neutralismo desempenham um papel central na modelagem da estabilidade e da dinâmica de comunidades em consórcio, o cocultivo permite a descoberta de complexas redes de interação estabelecidas dentro de tais sistemas, as quais geralmente são mediadas por trocas de moléculas que dependem ou independem de contato (FAUST; RAES, 2012; SONG et al., 2014). Consórcios também tornam possível a condução de estudos sobre interação ecológica e coevolução (BRENNER; YOU; ARNOLD, 2008). Em um consórcio microbiano, cianobactérias podem ajudar a criar melhores condições para a comunidade desempenhar uma determinada tarefa (ZHUBANOVA et al., 2013). É possível, ainda, que cianobactérias facilitem o desenvolvimento de bactérias não previamente cultivadas, como observado em culturas não axênicas de outros micro-organismos fotossintetizantes, tais como clorófitas do gênero *Chlorella* (OTSUKA et al., 2008).

Em contraposição às abordagens convencionais de genômica microbiana, que pressupõem dados provenientes de amostras axênicas, a metagenômica trabalha com dados resultantes de amostras mistas, compostas por misturas de diferentes populações de micro-organismos (MCHARDY; RIGOUTSOS, 2010). Abordagens metagenômicas se mostraram úteis para o desvelamento da composição, da genética e do metabolismo de comunidades microbianas naturais e artificiais, e são, portanto, apropriadas para a avaliação de consórcios microbianos (SONG et al., 2014).

Avaliações de culturas cianobacterianas não axênicas por sequenciamento em larga escala são similares a análises de sequências obtidas de estratégias metagenômicas de enriquecimento, focadas na investigação de consórcios microbianos em meios líquidos que estimulam o crescimento de um micro-organismo em específico (SCHLOSS; HANDELSMAN, 2004). Tal abordagem tem algumas vantagens sobre a análise de amostras ambientais, já que a montagem de metagenomas é mais eficiente sob condições de menor riqueza e maior coerência genômica, onde enfrenta menor interferência de sequências irrelevantes (TEELING; GLÖCKNER, 2012). Dessa forma, culturas não axênicas de cianobactérias são mais adequadamente analisadas sob a ótica da metagenômica, e oferecem a oportunidade de realizar tanto a avaliação genômica da cianobactéria-alvo quanto o estudo metagenômico da comunidade microbiana associada à mesma.

2.3 Produtos naturais cianobacterianos

Uma ampla gama de substâncias químicas é sintetizada por cianobactérias, tanto sob condições axênicas quanto simbióticas, em sistemas artificiais ou naturais. Espécies do gênero *Nostoc* foram descritas como produtoras de variantes já conhecidas ou inéditas de hidrocarbonetos, arsenolipídeos, ácidos graxos, terpenoides, carotenoides, macróciclos, sacarídeos, peptídeos, lipopeptídeos e derivados de aminoácidos, com muitas dessas moléculas sendo empregadas em interações ecológicas, tais como a competição (DEMBITSKY; ŘEZANKA, 2005) ou a simbiose (LIAIMER et al., 2015).

Diversas moléculas com atividades antitumorais, antibacterianas, antiprotozoárias, imunomodulatórias e inibidoras de protease, entre outras, foram descritas, apontando para as cianobactérias como uma fonte prolífica para a produção de moléculas bioativas (SINGH et al., 2011). Alguns autores consideram que o potencial para descoberta de produtos naturais desses organismos é somente superado pelo potencial de mixobactérias e do gênero actinobacteriano *Streptomyces* (NUNNERY; MEVERS; GUNNERY, 2010). Apesar do papel ecológico da maioria desses metabólitos ainda não ser conhecida, a ecologia química deste filo é um tópico de crescente exploração (LEÃO et al., 2012a).

Nostoc se encontra entre os gêneros cianobacterianos que apresentam potencial para a geração de compostos medicinais exclusivos (DODDS et al., 1995). Em espécies de *Nostoc*, podem ser encontrados diversos metabólitos secundários bioativos, como microcistina (inibidora de fosfatases), nostoficina (com atividade citotóxica), criptoficina (inibidora da polimerização de tubulinas), nostociclamida (inibidora do crescimento de algas e bactérias), nostocarbolina (inibidora de acetilcolinesterase), nostocina A (inibidora do crescimento de algas e plantas), muscorida A (com atividade antibacteriana) e cianobacterina (que inibe o crescimento de *Synechococcus*) (KURMAYER, 2011).

A produção de metabólitos secundários bioativos por cianobactérias é bastante conhecida. Esses micro-organismos comumente produzem uma grande diversidade de peptídeos, alcaloides e lipopolissacarídeos bioativos (BÖRNER; DITTMANN, 2005; DITTMANN; NEILAN; BÖRNER, 2001), muitos dos quais com modos de ação desconhecidos ou pouco explorados. Algumas das substâncias conhecidas têm efeito anti-inflamatório, anti-infeccioso, antitumoral ou ação neurológica (NUNNERY; MEVERS; GERWICK, 2010). Contudo, as moléculas que tornam o filo Cyanobacteria notório são as cianotoxinas, moléculas que geralmente apresentam ações citotóxicas, hepatotóxicas ou neurotóxicas sobre animais, ou ação contra alguns organismos do zooplâncton e fitoplâncton e

até mesmo contra outras cianobactérias (KAEBERNICK; NEILAN, 2001). Atualmente, a maior parte dos estudos sobre cianotoxinas é voltada aos seus efeitos à saúde de seres humanos (KOREIVIENĖ et al., 2014; ZANCHETT; OLIVEIRA-FILHO, 2013).

O papel ecológico e evolutivo das cianotoxinas ainda não é compreendido, porém diversas hipóteses já foram levantadas. Elas podem ser agrupadas em estratégias para auxílio fisiológico, focadas no aprimoramento da homeostasia, da fotossíntese e do crescimento, e estratégias para vantagem competitiva, voltadas para defesas celulares em resposta à predação e à disputa por recursos (HOLLAND; KINNEAR, 2013). Mesmo cianobactérias em simbiose podem produzir cianotoxinas, mas suas funções e consequências para os outros organismos da associação também ainda são desconhecidas (GEHRINGER et al., 2012; KAASALAINEN et al., 2013a; 2013b).

O filo Cyanobacteria não é considerado uma fonte tradicional de produtos naturais para a indústria, já que o estudo de moléculas bioativas nestes micro-organismos se restringiu primariamente aos mecanismos de produção de cianotoxinas. Não obstante, pesquisas desenvolvidas nas últimas décadas demonstraram em diversas cianobactérias a síntese de uma fascinante variedade de enzimas envolvidas em diversos processos bioquímicos, incluindo uma grande quantidade de moléculas exclusivas ao grupo ou raramente encontradas em outros organismos (KEHR; PICCHI; DITTMANN, 2011). Contudo, a tradicional subestimação da biodiversidade cianobacteriana, fruto das limitações de seus sistemas clássicos de classificação, provoca a falsa percepção de que apenas poucos gêneros são responsáveis por uma rica produção de metabólitos secundários (ENGINE et al., 2011). Apesar disso, produtos naturais cianobacterianos continuam a ser negligenciados, com quedas contínuas nos relatos de novos metabólitos deste filo desde 2012 (BLUNT et al., 2015).

Abordagens genômicas se mostram cada vez mais poderosas para a descoberta de novas fontes de produtos naturais à medida que o número de sequências de genomas disponíveis aumenta. Por meio de estratégias que permitem partir da sequência gênica e chegar à predição da estrutura final da molécula, é possível agora averiguar mais facilmente a diversidade metabólica de micro-organismos cuja relevância como fontes de metabólitos era antes desconhecida ou subestimada devido às restrições metodológicas e à dispendiosidade dos protocolos clássicos, baseados na fermentação e na posterior detecção de substâncias de interesse por bioensaios ou análises químicas em extratos brutos, que impedem a exploração do verdadeiro potencial biossintético microbiano (WINTER; BEHNKEN; HERTWECK, 2011). A mineração de genomas tem levado à descoberta de novos produtos naturais e de novos mecanismos de biossíntese a partir da caracterização de sequências gênicas que

codificam peptídeo sintetases não ribossômicas, policetídeo sintases e terpeno sintases, entre outras (CORRE; CHALLIS, 2009). Através da mineração genômica, mesmo vias biossintéticas que permaneceram desconhecidas por 70 anos têm sido elucidadas (CRAWFORD; CLARDY, 2012). Contudo, a maioria dos agrupamentos gênicos levantados por mineração genômica ainda possui produtos desconhecidos (CALTEAU et al., 2014).

As pesquisas em genômica de micro-organismos nas últimas décadas se voltaram amplamente para o estudo de linhagens bacterianas patogênicas a animais ou a plantas com o objetivo de estabelecer inventários de genes que promovessem o estabelecimento de alvos para novos antibióticos, a disponibilização de novos candidatos para a produção de vacinas e avanços na compreensão da biologia desses patógenos, com a descoberta de genes envolvidos na síntese de produtos naturais sendo, portanto, um resultado meramente incidental (DONADIO; MONCIARDINI; SOSIO, 2007). Nos últimos anos, contudo, o aumento do número de projetos genômicos voltados para recursos com relevância ambiental passou a favorecer o descobrimento de novos metabólitos secundários (WINTER; BEHNKEN; HERTWECK, 2011).

2.4 Genômica cianobacteriana

Em 1996, foi finalizado o sequenciamento e a anotação do primeiro genoma de cianobactéria, *Synechocystis* sp. PCC 6803 (KANEKO et al., 1996). Atualmente, pouco mais de 200 genomas cianobacterianos completos estão disponíveis em bancos de dados públicos, um número ínfimo perto dos mais de 30.000 genomas pertencentes a organismos de 50 filos de bactérias e de 11 filos de arqueias já sequenciados (LAND et al., 2015). A maioria desses genomas cianobacterianos é proveniente de linhagens de água marinha ou doce, depositadas na coleção de culturas do Instituto Pasteur, em Paris, na França, constituindo, portanto, uma mera amostra da enorme diversidade deste filo.

Poucos genomas identificados no gênero *Nostoc* foram disponibilizados até o momento. Além disso, em decorrência da polifilia desse táxon, apenas o genoma de uma espécie pertencente ao grupo do qual a espécie-tipo faz parte, *N. punctiforme* PCC 73102, foi sequenciado. O posicionamento taxonômico de *Nostoc* (*Anabaena*) PCC 7120 ainda está sob discussão, mas provavelmente esta linhagem representa um gênero distinto de ambos (SVENNING; ERIKSSON; RASMUSSEN, 2005). Assim como PCC 7120, a linhagem 0708, um simbiote obrigatório, foi inicialmente classificada como *Anabaena* (RAN et al., 2010). Neste momento, ela pode ser encontrada identificada como *Trichormus azollae* em alguns

bancos de dados, cujo gênero está representado por outro genoma, depositado como *Anabaena variabilis* ATCC 29413. As linhagens *Nostoc* sp. PCC 7107 e *Nostoc* sp. PCC 7524 também se mostram relacionadas ao gênero *Trichormus*. Como verificado na Tabela 1, outros sete projetos de sequenciamento genômico de organismos identificados no gênero *Nostoc* estão registrados no banco de dados GenBank (disponível em <http://www.ncbi.nlm.nih.gov>).

Tabela 1 – Projetos de sequenciamento genômico de linhagens do gênero *Nostoc* sendo desenvolvidos atualmente

Linhagem	Número de Acesso	Centro Responsável	Hábitat
<i>Nostoc piscinale</i> CENA21	PRJNA288507	Universidade Federal do Pará / University of California in San Diego	sedimentos, Rio Solimões
<i>Nostoc</i> sp. 'Peltigera membranacea cyanobiont N6'	PRJNA279350	University of Iceland	não especificado
<i>Nostoc</i> sp. 'Lobaria pulmonaria (5183) cyanobiont'	PRJNA275880	University of Iceland	não especificado
<i>Nostoc linckia</i>	PRJNA274736	University of Hong Kong	não especificado
<i>Nostoc</i> sp. CCAP 1453/38	PRJNA266493	Biology Centre ASCR	não especificado
<i>Nostoc</i> sp. sp2	PRJNA262403	DOE Joint Genome Institute	não especificado
<i>Nostoc</i> sp. sp3	PRJNA262218	DOE Joint Genome Institute	associada a musgo, floresta boreal

De acordo com sua história evolutiva, as cianobactérias, responsáveis pela oxigenação da atmosfera terrestre, podem ser consideradas como um elo de ligação entre organismos anaeróbicos e organismos obrigatoriamente aeróbicos (HAREL et al., 2015). O ancestral mais recente das cianobactérias foi estimado como tendo aproximadamente 4,5 Mpb e de 1.678 a 3.291 genes, dos quais apenas entre 4 e 6 % permanece exclusivo aos genomas das cianobactérias modernas, os quais apresentam como inovações sequências para formação de filamentos, diferenciação de heterócitos, diazotrofia e competência simbiótica (LARSSON; NYLANDER; BERKMAN, 2011).

A alta diversidade genética nesses micro-organismos é evidenciada por ampla distribuição e hipervariabilidade de elementos transponíveis, os quais chegam a constituir até 10,95 % do genoma total de algumas linhagens (LIN et al., 2010). Alto percentual desses elementos pode ser detectado em genomas de cianobactérias que habitam ambientes considerados extremos, o que pode estar relacionado à sua capacidade de adaptação às

condições encontradas nesses ambientes (KIDWELL; LISCH, 2001; LIN et al., 2010). Além disso, a transferência horizontal de genes também é responsável por gerar alta diversidade molecular nesses micro-organismos (ZHAXYBAYEVA et al., 2006).

O material genético das cianobactérias geralmente é composto, assim como o de outras bactérias, por uma ou mais moléculas de plasmídeo e por DNA cromossômico, o qual, em média, pode variar em tamanho entre 1,4 e 8,2 Mpb (MEEKS et al., 2001; ZEHR et al., 2008). Análises do número de cópias de cromossomos em células de cianobactérias indicam que neste filo pode-se observar a ocorrência de monoploidia, oligoploidia ou poliploidia, de modo que algumas cianobactérias podem apresentar até 218 cópias durante sua fase exponencial (GRIESE; LANGE; SOPPA, 2011), o que poderia favorecer a recombinação. Durante situações de estresse, estratégias de sobrevivência tornam esse número ainda maior. Acinetos podem conter até 450 cromossomos, um número explicado pela necessidade de uma rápida retomada da atividade metabólica e da divisão celular após a dormência da célula (SUKENIK et al., 2012).

A divergência genômica geralmente reflete adaptações fisiológicas aos diferentes nichos que as cianobactérias podem ocupar. Duas estratégias de adaptação podem ser inferidas pelas análises de genomas do filo Cyanobacteria: a expansão genômica, que confere amplo potencial adaptativo pelo aumento de famílias gênicas, e a redução genômica, que resulta na eliminação de genes dispensáveis à adaptação a um nicho específico (LARSSON; NYLANDER; BERKMAN, 2011). Pressões de seleção provocariam mudanças em fatores como tamanho do genoma, percentual G-C, número de genes e taxa de evolução. Enquanto cianobactérias podem desenvolver estratégias individuais de interação com o ambiente, outros de seus sistemas são conservados globalmente (SIMM et al., 2015). Assim como ocorre em outros organismos, há um conjunto de genes essenciais em cianobactérias que apresenta taxas de conservação bastante elevadas e maior resistência à transferência horizontal, consistindo de sequências geralmente envolvidas na produção de estruturas proteicas complexas e em vias bioquímicas indispensáveis (LARSSON; NYLANDER; BERKMAN, 2011; SHI; FALKOWSKI, 2008). É esse conjunto de genes, que encerra grande parte da história evolutiva do organismo, o alvo de vários trabalhos recentes.

Disponibilizadas a partir de 2005, as novas gerações das tecnologias de sequenciamento de DNA foram, em conjunto com as novas tecnologias computacionais, as principais impulsionadoras da pesquisa em genômica e metagenômica nas últimas décadas. Elas tornaram possível que fosse gerado um volume de dados muito maior e em menor tempo que os obtidos com a antiga tecnologia de sequenciamento de Sanger et al. (1977), que havia

se tornado o método-padrão de sequenciamento de DNA. As plataformas de sequenciamento de DNA de alta performance levaram a reduções drásticas tanto no custo econômico quanto no esforço necessário para a caracterização de genomas, o que levou ao crescimento exponencial dos artigos científicos descrevendo pesquisas genômicas. Essas pesquisas saturaram o meio editorial, de modo que, apesar de terem fornecido dados fundamentais para o avanço da genética e da evolução nos últimos 20 anos, apenas genomas com características muito singulares e inovadoras ou análises comparativas de múltiplos genomas são considerados para publicação atualmente (SMITH, 2013).

Através da genômica, têm se tornado possíveis os seguintes avanços: compreender a trajetória dos genes de cianotoxinas, assim como o modo como são adquiridos, e levantar informações para o entendimento de sua função ecofisiológica e evolução (MOUSTAFA et al., 2009); prever funções bioquímicas de genes de atuação ainda desconhecida (HUANG et al., 2006); conhecer a história evolutiva do filo e da vida do planeta (MULKIDJANIAN et al., 2006); estudar possíveis alterações geológicas e biogeoquímicas ocorridas em conjunto com alterações na biota (BATTISTUZZI; FEIJÃO; HEDGES, 2004); e, por meio da filogenômica, o aprimoramento da taxonomia (EISEN; FRASER, 2003; THOMPSON et al., 2015). Pesquisas metagenômicas também têm realizado descobertas importantes, e revelaram o novo filo Melainobacteria, que foi identificado como proximamente relacionado ao filo Cyanobacteria (DI RIENZI et al., 2013; SOO et al., 2014).

Novas técnicas foram introduzidas à classificação taxonômica das cianobactérias para questionar a tradicional ênfase em características morfológicas, levando à revisão de seus critérios. O relativamente novo campo da genômica comparativa se destaca e pode levar a uma nova revolução dentro da sistemática cianobacteriana. A filogenômica e a verificação da presença de sequências de proteínas encontradas apenas em determinados táxons, assim como a análise de inserções ou deleções conservadas em sequências de proteínas estruturais, se mostram promissoras para o esclarecimento de graus de ancestralidade não passíveis de serem inferidos por análises filogenéticas comuns. A união da filogenômica com as análises de sinapomorfia e de proteínas exclusivas vem redefinindo a compreensão da evolução das cianobactérias e tem dado maior robustez à classificação cianobacteriana moderna (GUPTA, 2009; GUPTA; MATHEWS, 2010; KOMÁREK et al., 2014).

Devido ao número relativamente baixo de genomas cianobacterianos atualmente disponíveis, a aplicação de abordagens filogenômicas a análises em níveis taxonômicos mais baixos (como gênero ou espécie) ainda não foi possível de ser realizada em larga escala com sucesso. Contudo, ela ainda permanece como a ferramenta mais promissora para trazer para o

filo Cyanobacteria uma classificação mais natural, capaz de corrigir os erros produzidos pela taxonomia tradicional, largamente baseada em morfologia, e esclarecer as complicadas relações evolutivas dos diversos táxons polifiléticos que persistem na classificação cianobacteriana, como, por exemplo, o gênero *Nostoc* e suas espécies. Potencialmente, mesmo erros causados por transferência lateral de genes ou outros processos que obscurecem o sinal filogenético são possíveis de serem resolvidos com a filogenômica (KAUFF; BÜDEL, 2011). Se neste momento há questões técnicas que dificultam o amplo uso taxonômico da genômica para cianobactérias, o sequenciamento de genomas tende a se tornar cada vez mais rápido e menos custoso, levando a um número crescente de genomas disponíveis.

Filogeneticamente, há um desequilíbrio na disponibilidade de sequências de cianobactérias, que acarreta uma má representação da genômica do filo. Isso prejudica o conhecimento da genética do filo, já que sequências de táxons negligenciados podem ajudar a trazer respostas importantes (RICHARDS, 2015). O sequenciamento de genomas de cianobactérias relacionadas a táxons pouco estudados possibilita não apenas o aumento do conhecimento da biologia do filo, mas também da evolução e da diversidade de aspectos mais gerais como morfologia, fotossíntese, metabolismo secundário e endossimbiose (SHI et al., 2013). O alto número de sequências repetitivas encontradas no genoma de grande parte das cianobactérias é um fator limitante para a montagem dos genomas destes organismos mesmo com o aumento da cobertura do sequenciamento, fator que contribui para que a grande maioria dos genomas atualmente disponíveis mantenha-se incompleta, no estado conhecido como “rascunho”, com a omissão de alto número de sequências de nucleotídeos em regiões significativas (WANG et al., 2012). Aproximadamente 90 % dos genomas sequenciados até o presente momento não está completo (LAND et al., 2015).

Mesmo para linhagens cianobacterianas provenientes de culturas axênicas, os sequenciamentos genômicos com alta cobertura e realizados por diferentes estratégias podem se mostrar insuficientes para a montagem genômica em uma sequência contígua única, um fato que ilustra o desafio proposto pelo alto número de repetições contido nos genomas de organismos desse filo. Todavia, isso não prejudica a obtenção de dados que satisfaçam os requerimentos necessários para a realização de análises genômicas comparativas (HUMBERT et al., 2013).

2.5 Terra preta antropogênica

De acordo com evidências arqueológicas, algumas paisagens amazônicas foram profundamente transformadas por antigas atividades humanas ao redor delas (KERN et al., 2009). Em contraste com solos comumente encontrados na região amazônica, que são normalmente ácidos e pobres em nutrientes, a terra preta antropogênica é um solo amazônico rico em matéria orgânica e com alta fertilidade química criado por ocupações humanas nativas estabelecidas mais de um milênio antes da colonização europeia do continente americano (COSTA et al., 2009; WOODS et al., 2009). Solos de terra preta amazônica são caracterizados por cor escura e altos níveis de Ca, Mg, Zn, Mn, P e C, que são resultantes principalmente de carvão residual de fogueiras para cozimento e da deposição de cinzas, cascas, restos de caça e coleta e dejetos humanos no local (KERN et al., 2009). Atividades humanas tendem a ter impacto negativo sobre cianobactérias de solo, especialmente durante processos de urbanização, mas poucas informações sobre como cianobactérias são impactadas em ambientes seminaturais e em solos de cultivo estão disponíveis (SHARMA, 2015).

Como resultado dessas características, os antrossolos podem em alguns casos apresentar comunidades microbianas distintas quando comparados com solos adjacentes de mesma mineralogia, porém não modificados. Por outro lado, essas comunidades microbianas podem ser comparativamente semelhantes às comunidades de outros antrossolos, mesmo quando estes foram formados a partir de diferentes tipos de solos e estão atualmente submetidos a práticas e coberturas consideravelmente dissimilares. Esse fato indica que, nesse ambiente, o histórico do solo pode ter maior influência sobre sua estrutura microbiana do que a composição do mesmo (GROSSMAN et al., 2010).

Antrossolos de características semelhantes tanto em termos de formação como de ecologia são hoje encontrados em sítios arqueológicos no norte da Alemanha, podendo ser considerados “terra preta nórdica” (WIEDNER et al., 2015). De forma similar ao ocorrido em terras amazônicas, esses solos foram formados como produto do sistema de subsistência viquingue/eslávica há pelo menos 1.000 anos AP, após a deposição contínua e prolongada de dejetos, tais como esterco humano e não humano, restos de caça e matéria orgânica carbonizada. Conseqüentemente, apresentam cor escura e são densos e enriquecidos com elementos como C, N, P, Ca, Mg, K, Na, Fe, Cu, Zn, Mn e Ba.

Além da caracterização de filos bacterianos negligenciados, tais como o filo Cyanobacteria, o conhecimento da diversidade microbiana geral também pode se beneficiar da caracterização de ambientes quimicamente díspares (SCHLOSS; HANDELSMAN, 2004). Estima-se que um número considerável de espécies de cianobactérias é ainda desconhecido, e acredita-se que há muito conhecimento a ser adquirido da caracterização da biodiversidade deste filo em regiões geográficas inexploradas (NABOUT et al., 2013). Cianobactérias também podem contribuir direta ou indiretamente para o aumento da concentração de nitrogênio, fósforo, potássio, ferro e outros minerais no solo e favorecer a eficiência de seu uso por plantas (KUMAR et al., 2015). Esses argumentos indicam a importância da caracterização de linhagens cianobacterianas e das comunidades com quem interagem de ambientes pouco explorados, como a terra preta antropogênica.

3. HIPÓTESES

Nostoc sp. CENA67 é uma cianobactéria que se diferencia do restante dos micro-organismos atualmente classificados em seu gênero, o que pode ser verificado por métodos clássicos e genômicos. Consequentemente, essa linhagem apresenta agrupamentos gênicos envolvidos na síntese de metabólitos secundários ainda desconhecidos. Culturas não axênicas dessa cianobactéria apresentam micro-organismos em associação com a mesma, que são passíveis de estudo por ferramentas metagenômicas.

4. OBJETIVOS

4.1 Objetivo geral

Este trabalho teve como objetivo geral a avaliação dos aspectos morfológicos, moleculares, genômicos e metabólicos da linhagem cianobacteriana *Nostoc* sp. CENA67, isolada a partir de solo de terra preta antropogênica da floresta amazônica brasileira, bem como a análise de características dos micro-organismos associados a essa cianobactéria em uma cultura não axênica.

4.2 Objetivos específicos

Entre os objetivos específicos deste projeto, encontram-se:

- Caracterizar a cianobactéria morfolologicamente;
- Verificar suas relações filogenéticas com outras cianobactérias;
- Caracterizar o genoma da cianobactéria-alvo;
- Comparar o genoma obtido com genomas de linhagens proximamente relacionadas;
- Avaliar agrupamentos gênicos envolvidos na síntese de metabólitos secundários;
- Analisar a diversidade taxonômica e funcional da comunidade microbiana associada.

5. MATERIAL E MÉTODOS

5.1 Origem da amostra

A linhagem *Nostoc* sp. CENA67 foi isolada a partir de uma amostra de terra preta antropogênica coletada em 2004 nos primeiros 20 cm de solo do sítio arqueológico Hatahara (3°16'29"S e 60°12'16"O), localizado no município de Iranduba, a 30 km a sudoeste de Manaus-AM. O sítio Hatahara encontra-se na margem esquerda do Rio Solimões, próximo à confluência com o Rio Negro. Essa linhagem é mantida em cultivo no Laboratório de Biologia Celular e Molecular do CENA/USP em meio de cultura líquido AA/4 (ALLEN; ARNON, 1955), composto pelos seguintes reagentes e concentrações: K_2HPO_4 132,68 mg·L⁻¹; $MgSO_4 \cdot 7H_2O$ 62,5 mg·L⁻¹; NaCl 62,5 mg·L⁻¹; $CaCl_2 \cdot 2H_2O$ 18,75 mg·L⁻¹; $FeSO_4 \cdot 7H_2O$ 4,982 mg·L⁻¹; KOH 1,891 mg·L⁻¹; H_3BO_3 715 µg·L⁻¹; $MnCl_2 \cdot 4H_2O$ 450 µg·L⁻¹; $Na_2MoO_4 \cdot 2H_2O$ 76,375 µg·L⁻¹; $ZnSO_4 \cdot 7H_2O$ 55 µg·L⁻¹; $CuSO_4 \cdot 5H_2O$ 19,75 µg·L⁻¹; $CoCl_2 \cdot 6H_2O$ 10 µg·L⁻¹; e NH_4VO_3 5,75 µg·L⁻¹. As células foram acondicionadas a 25±1 °C sob iluminação fluorescente de 40 µmol fótons·m⁻²·s⁻¹ e com fotoperíodo de 14 horas de luz e 10 horas de escuro. Essa é uma cultura não axênica, sendo composta, dessa maneira, tanto por células de *Nostoc* sp. quanto de alguns micro-organismos não cianobacterianos associados a elas.

5.2 Obtenção de biomassa

Para obtenção de biomassa, culturas de CENA67 foram incubadas em frascos Erlenmeyer de 1 L contendo 500 mL de meio líquido AA/4. Após três semanas de crescimento, as células foram concentradas em tubos de 50 mL por centrifugação a 7690 ×g durante 10 minutos em uma centrífuga Hettich Universal Mikro320R (Hettich, Tuttlingen, Alemanha). Em seguida ao descarte dos sobrenadantes, foram feitas três lavagens das células através da adição de 50 mL de água ultrapura seguida de homogeneização em vórtex, centrifugação a 7690 ×g durante 10 minutos e descarte dos sobrenadantes. As células concentradas foram submetidas a uma nova lavagem com a adição de uma solução de extran 0,05 % seguida de homogeneização, centrifugação a 2370 ×g durante 10 minutos e descarte dos sobrenadantes. Este procedimento foi repetido duas vezes com solução de lavagem (NaCl 50 mM, Tris-HCl 10 mM pH 7,5, EDTA 2,5 mM pH 8,0 e etanol 50 %) a 4650 ×g e uma vez com solução salina 0,85 % a 7690 ×g. As células lavadas e concentradas foram armazenadas a -20 °C até a extração de DNA.

5.3 Caracterização morfológica

Para examinar as características morfológicas da linhagem CENA67, amostras de culturas líquidas com 14 dias e 45 dias de crescimento foram coletadas e observadas sob um microscópio óptico Zeiss Axiostar Plus (Carl Zeiss, Jena, Alemanha) de acordo com os critérios de Papaefthimiou et al. (2008). Medidas morfométricas foram tomadas das amostras com o auxílio de um micrômetro ocular pela observação das células com um aumento de 1000 \times . O comprimento e a largura de vinte e cinco células vegetativas e heterócitos intercalares e terminais, selecionados aleatoriamente, foram medidos. Para tentar estimular a produção de acinetos pela linhagem, a cultura foi mantida sob escuro absoluto ou a luz constante e a temperaturas de 25 \pm 1 °C ou 10 \pm 1 °C, sendo constantemente monitorada para verificação da produção destas células. Além das dimensões celulares, seu formato (cônico, arredondado, abarrilado ou oval) também foi verificado. Os filamentos foram classificados em relação ao seu formato (em espiral, ondulados ou retos) e seu tamanho (curtos, se formados por menos de 50 células, e longos, se filamentos de mais de 50 células estivessem presentes). Por fim, as bainhas foram observadas após a adição de nanquim às lâminas e classificadas como difusivas (se não havia bordas nítidas na microcolônia) ou como envelopes (se uma bainha delicada e não envelopada se expandia por 2-5 μ m ao redor do filamento).

5.4 Extração de DNA

Para as extrações de DNA, foram coletadas células na fase exponencial de crescimento pela concentração em centrífuga durante 10 minutos a 7392 \times g. A extração foi realizada a partir de modificações no protocolo descrito por Lin et al. (2010), reproduzidas a seguir. As células de CENA67 coletadas foram inicialmente ressuspensas em 500 μ L de tampão de lise (composto por Tris-HCl 100 mM pH 8,0, EDTA 1 mM pH 9,0 e NaCl 100 mM) e digeridas com 5 μ L de uma solução de lisozima 100 mg \cdot mL⁻¹ a 37 °C por 30 minutos. Seguiu-se a adição de 10 μ L de proteinase K 20 mg \cdot mL⁻¹ e 20 μ L de SDS 10 % aos microtubos e a incubação a 55 °C durante 1 hora. Em seguida, 600 μ L de uma solução de fenol/clorofórmio/álcool isoamílico (25:24:1) foram adicionados aos tubos, os quais foram então homogeneizados e centrifugados a 7392 \times g por 5 minutos. Posteriormente, os sobrenadantes foram transferidos para novos tubos, aos quais foram adicionados 600 μ L de clorofórmio/álcool isoamílico (24:1), sendo novamente homogeneizados e centrifugados a 7392 \times g por 5 minutos. Os sobrenadantes foram mais uma vez transferidos para novos tubos,

acrescidos de 4 μL de RNase A 10 $\text{mg}\cdot\text{mL}^{-1}$ e incubados a 37 °C por 20 minutos. O DNA extraído foi então precipitado com 100 μL de acetato de sódio 3 M pH 4,8 e 600 μL de isopropanol a -80 °C durante 1 hora e centrifugado por 10 minutos a 7392 $\times g$. Por fim, o DNA foi lavado com 500 μL de etanol 70 % e centrifugado a 7392 $\times g$ por 5 minutos. Após secagem a 37 °C, o DNA foi ressuscitado em água ultrapura estéril. A integridade das amostras foi verificada em gel de agarose 1 % após corrida eletroforética em tampão TBE 0,5 X (Tris-borato 22 mM e EDTA 0,5 mM pH 8,0) com o corante SYBR Green (Molecular Probes, Eugene, OR, EUA). O registro do gel foi realizado por meio do fotodocumentador Gel Logic 212 Imaging System (Molecular Imaging System Carestream Health, Inc, Rochester, NY, EUA). A quantificação do DNA se deu com o fluorômetro Qubit, utilizando o reagente Quant-iT dsDNA BR Assay Kit (Invitrogen) de acordo com as instruções do fabricante.

5.5 Reconstrução filogenética

O DNA genômico extraído foi submetido à amplificação por reação em cadeia da polimerase (PCR) do DNAr 16S com os oligonucleotídeos iniciadores 27F1 e 1494Rc (NEILAN et al., 1997). A amplificação foi feita em solução contendo tampão PCR 1X (Tris HCl 20 mM pH 8,4; KCl 50 mM), 0,2 mM de cada dNTP, MgCl_2 3 mM, 1 U de Platinum Taq DNA Polimerase (Invitrogen), 10 ng de DNA, 5 μmol de cada iniciador e água ultrapura (Millipore) esterilizada. A reação foi realizada em um termociclador Techne TC-412 (Bibby Scientific Limited, Stone, Staffordshire, Inglaterra) sob as seguintes condições: desnaturação a 95 °C durante 3 minutos; 30 ciclos a 94 °C por 10 segundos, 50 °C por 20 segundos e 72 °C por 1 minuto; e extensão final a 72 °C durante 7 minutos. A integridade da amplificação foi analisada por eletroforese, conforme descrito anteriormente, por comparação com o marcador de massa molecular Low Mass DNA Ladder (Invitrogen). Os produtos da amplificação foram clonados e inseridos em células de *Escherichia coli* DH5 α com pGEM-T Easy Vector Systems (Promega, Madison, WI, EUA) por choque térmico (SAMBROOK; RUSSEL, 2001; ZHOU; CLARK; GOMEZ-SANCHEZ, 1995). A presença dos insertos foi checada através de PCR com os iniciadores CYA359F, CYA781Ra e CYA781Rb, sob condições descritas anteriormente (NÜBEL et al., 1997). Colônias positivas foram conduzidas ao sequenciamento pelo método de Sanger et al. (1977) no aparelho ABI PRISM 3100 Genetic Analyzer (Life Technologies, Carlsbad, CA, EUA) com DYEnamic ET Dye Terminator Cycle Sequencing Kit (GE Healthcare, Little Chalfont, Buckinghamshire, Inglaterra) e os iniciadores M13F, M13R, 341-357F, 357-341R, 685-704F, 704-685R,

1099-1114F e 1114-1099R. As sequências geradas foram analisadas com o pacote de programas Phred/Phrap/Consed (EWING et al., 1998; EWING; GREEN, 1998; GORDON et al., 1998). As sequências dos iniciadores utilizados neste trabalho encontram-se na Tabela 2.

Tabela 2 – Oligonucleotídeos iniciadores utilizados para análise filogenética

Iniciador	Sequência	Referência
27F1	5'-AGAGTTTGATCTGCTCAG-3'	NEILAN et al., 1997
1494Rc	5'-TACGGCTACCTTGTTACGAC-3'	NEILAN et al., 1997
CYA359F	5'-GGGGAATYTTCCGCAATGGG-3'	NÜBEL et al., 1997
CYA781Ra	5'-GACTACTGGGGTATCTAATCCCATT-3'	NÜBEL et al., 1997
CYA781Rb	5'-GACTACAGGGGTATCTAATCCCTTT-3'	NÜBEL et al., 1997
M13F	5'-GCCAGGGTTTTCCAGTCACGA-3'	MESSING, 1983
M13R	5'-GAGCGGATAACAATTCACACAGG-3'	MESSING, 1983
341-357F	5'-CCTACGGGAGGCAGCAG-3'	LANE, 1991 (modificado)
357-341R	5'-CTGCTGCCTCCCGTAGG-3'	LANE, 1991 (modificado)
685-704F	5'-GTAGSGGTGAAATSCGTAGA-3'	LANE, 1991 (modificado)
704-685R	5'-TCTACGSATTTACCCSCTAC-3'	LANE, 1991 (modificado)
1099-1114F	5'-GCAACGAGCGMRACCC-3'	LANE, 1991 (modificado)
1114-1099R	5'-GGGTYKCGCTCGTTGC-3'	LANE, 1991 (modificado)

Visando compreender melhor as relações evolutivas entre CENA67 e outras linhagens relacionadas ao gênero *Nostoc*, a sequência resultante foi comparada com outras sequências depositadas no banco de dados GenBank, com a ferramenta BLAST (ALTSCHUL et al., 1990). A sequência obtida foi alinhada junto a sequências proximamente relacionadas com MUSCLE 3.8.31 (EDGAR, 2004) e o melhor modelo evolutivo para o alinhamento foi estimado com jModelTest 2.1.7 (DARRIBA et al., 2012; GUINDON; GASCUEL, 2003). Árvores filogenéticas foram reconstruídas pelos métodos de máxima verossimilhança (FELSENSTEIN, 1981), utilizando o programa RAxML 8.2.3 (STAMATAKIS, 2006), e inferência bayesiana (MAU; NEWTON; LARGET, 1999), realizada com o programa MrBayes 3.2.5 (RONQUIST; HUELSENBECK, 2003). A árvore final foi visualizada com Figtree 1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree>) e editada com Inkscape 0.48.5 (<http://inkscape.org>).

5.6 Sequenciamento de DNA

O DNA genômico foi utilizado para o preparo de bibliotecas genômicas com Nextera XT DNA Sample Prep Kit (Illumina) de acordo com as instruções do fornecedor, resumidas a seguir. O DNA genômico foi inicialmente fragmentado por transposases e recebeu sequências adaptadoras às suas extremidades após incubação em termociclador a 55 °C por 5 minutos. As enzimas foram neutralizadas com o tampão Neutralize Tagment (Illumina) à temperatura ambiente por 5 min e, em seguida, ocorreu a amplificação do DNA com os iniciadores i7 e i5, específicos aos adaptadores adicionados, com a seguinte ciclagem térmica: 72 °C por 3 min; 95 °C por 30 s; 12 ciclos de 95 °C por 10 s, 55 °C por 30 s e 72 °C por 30 s; e 72 °C por 5 min. Os produtos da amplificação foram purificados com 25 µL de microesferas magnéticas AMPure XP 0,5 X e etanol 80 %. Após purificada, a biblioteca foi normalizada para o sequenciamento. Após os produtos da biblioteca serem desnaturados e diluídos, eles foram transferidos para um cartucho de reagentes. Uma célula de fluxo foi limpa com água ultrapura estéril e inserida no sequenciador, seguida dos reagentes para sequenciamento e do cartucho contendo as sequências da biblioteca. O sequenciamento foi realizado com MiSeq Reagent Kit (Illumina) no Centro de Genômica Funcional Aplicada à Agropecuária e à Agroenergia, da Universidade de São Paulo, em Piracicaba-SP, segundo instruções do fabricante.

Um sequenciamento com Ion PGM (Life Technologies) também foi realizado no Laboratório de Polimorfismo de DNA, na Universidade Federal do Pará, em Belém-PA. Para esse sequenciamento, 5 µg do DNA extraído anteriormente foram utilizados na construção de bibliotecas de pares com 3 Kpb de distância entre si com Ion PGM Template OT2 400 Kit (Life Technologies) seguindo instruções do fabricante e procedimentos descritos anteriormente (RAMOS et al., 2013). Os produtos das bibliotecas foram depositados em um Ion 318 Chip v2 (Life Technologies) e o sequenciamento ocorreu com Ion PGM Sequencing 400 Kit (Life Technologies) segundo instruções do fabricante.

5.7 Pré-processamento de sequências e montagem

As qualidades das leituras brutas obtidas foram verificadas com o programa FastQC 0.10.1 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). PEAR 0.9.6 (ZHANG et al., 2014) foi utilizado para unir pares com extremidades que se sobrepunham. Bases com índice de qualidade menor que Phred 30, bases ambíguas e sequências menores que 50 pb

foram filtradas com Seqclean 1.3.12 (<http://cores.ibest.uidaho.edu/software/seqclean>) e PRINSEQ 0.20.4 (SCHMIEDER; EDWARDS, 2011a). Sequências com similaridade a contaminantes metagenômicos conhecidos foram removidas com DeconSeq 0.4.3 (SCHMIEDER; EDWARDS, 2011b). O conjunto de dados resultante do pré-processamento foi montado com SPAdes 3.1.1 (BANKEVICH et al., 2014) e QUASt 2.3 (GUREVICH et al., 2013) foi utilizado para avaliar os resultados das montagens. O tamanho genômico total foi estimado com o programa Kmer Spectrum Analyzer (WILLIAMS et al., 2013).

5.8 Anotação e detecção de produtos naturais

O arquivo gerado ao final dos procedimentos de montagem foi anotado automaticamente com Prokka 1.8 (SEEMANN, 2014) e com o servidor RAST (AZIZ et al., 2008; OVERBEEK et al., 2014). Para a detecção de genes provenientes de agrupamentos gênicos potencialmente envolvidos na síntese de metabólitos secundários, utilizou-se o programa antiSMASH 3.0.3 (BLIN et al., 2013; MEDEMA et al., 2011; WEBER et al., 2015). Outros genes de interesse foram identificados a partir de alinhamentos locais com bancos de dados customizados com BLAST+ 2.2.29 (ALTSCHUL et al., 1990; CAMACHO et al., 2009). A anotação automática de regiões de interesse foi averiguada e, quando necessário, corrigida por meio de curadoria manual com Artemis 16.0.0 (RUTHERFORD et al., 2000) e BLAST (ALTSCHUL et al., 1990). Quando necessário, árvores foram construídas conforme descrito anteriormente.

5.9 Análises comparativas

Comparações genômicas da linhagem *Nostoc* sp. CENA67 com genomas de linhagens disponíveis em bancos de dados públicos foram realizadas com o servidor RAST para verificação da proximidade com outros genomas e sintonia, enquanto GET_HOMOLOGUES 11032014 (CONTRERAS-MOREIRA; VINUESA, 2013) foi utilizado para a comparação de sequências proteicas em larga escala e estimativas de tamanho do pangenoma e do genoma essencial dos genomas das cianobactérias já classificadas dentro do gênero *Nostoc*. Após a detecção do agrupamento gênico de DNAr dentre as sequências montadas, MUSCLE 3.8.31 foi utilizado para o alinhamento do espaço intergênico 16S-23S de CENA67 junto a sequências de espaços intergênicos provenientes de táxons relacionados depositadas em bancos de dados públicos. A identificação de regiões conservadas foi realizada de acordo com

Iteman et al. (2000). As regiões D1-D1' e Box B foram selecionadas para a previsão de estruturas secundárias com Mfold (ZUKER; STIEGLER, 1984; ZUKER, 2003), que foram visualizadas com RnaViz 2 (DE RIJK et al., 2003).

5.10 Diversidade taxonômica e funcional

Sequências contíguas do metagenoma montado foram analisadas e identificadas com Kraken 0.10.4 (WOOD; SALZBERG, 2014). Anotações do metagenoma da comunidade associada e dos genomas obtidos foram fornecidos, respectivamente, pelos servidores MG-RAST (MEYER et al., 2008) e RAST (AZIZ et al., 2008; BRETTIN et al., 2015; OVERBEEK et al., 2014). Quando necessário, a curadoria manual das anotações foi realizada com Artemis 16.0.0 (RUTHERFORD et al., 2000) e BLAST (ALTSCHUL et al., 1990). A afiliação taxonômica dos micro-organismos associados a *Nostoc* sp. CENA67 foi explorada inicialmente em dados não montados com FOCUS 0.26 (SILVA et al., 2014) para também avaliar a abundância relativa dos táxons. Alternativamente, o DNAr 16S das leituras sequenciadas foi recuperado com SortMeRNA 2.0 (KOPYLOVA; NOÉ; TOUZET, 2012) e analisado com QIIME 1.9 (CAPORASO et al., 2010a) usando uclust 2.21 (EDGAR, 2010), PyNAST 1.2.2 (CAPORASO et al., 2010b) e o banco de dados Greengenes 13_8 (DESANTIS et al., 2006). Além disso, o DNAr 16S foi montado com EMIRGE 0.60 (MILLER et al., 2011). O resultado das identificações foi utilizado para investigar as diferenças funcionais entre a cianobactéria e a comunidade associada. Vias metabólicas foram comparadas no servidor MG-RAST com o banco de dados KEGG. Real Time Metagenomics (EDWARDS et al., 2012) foi utilizado para comparações de subsistemas.

6. RESULTADOS E DISCUSSÃO

6.1 Caracterização morfológica e filogenética

CENA67 é uma cianobactéria filamentosa com características típicas da família Nostocaceae, apresentando tricomas isopolares unisseriados, divisão celular perpendicular ao eixo do tricoma, ausência de ramificação e de zonas meristemáticas, e diferenciação de heterócitos. Os filamentos de CENA67 são longos e ondulados e se reproduzem por hormogônios igualmente ondulados, porém curtos, e imóveis. Ambos são envoltos por bainhas do tipo envelope. As células vegetativas desses filamentos se diferenciam em heterócitos intercalares, que podem ter formato arredondado ou abarrilado, ou em heterócitos terminais arredondados. As características morfológicas de CENA67 podem ser observadas na Figura 2.

Em média, as células vegetativas dessa linhagem possuem 7 μm de comprimento e 3,7 μm de largura, o equivalente a uma proporção comprimento/largura de 1,9. Os heterócitos intercalares possuem em média 8 μm de comprimento e 6,9 μm de largura, enquanto os heterócitos terminais possuem média de 5,3 μm de comprimento e 5 μm de largura, com proporções comprimento/largura de aproximadamente 1,2 e 1, respectivamente. Mesmo sob as diferentes condições de estresse a que foi submetida (que incluíram longos períodos de cultivo, baixas temperaturas, luminosidade constante e ausência de luminosidade), não foi observada nas culturas de CENA67 a formação de acinetos, essenciais para o início da formação colonial classicamente observada no gênero.

A ausência de acinetos impede nessa cianobactéria um ciclo de vida que se aproxime do ciclo típico do gênero *Nostoc*. Por consequência, colônias macroscópicas também não são encontradas na linhagem estudada. É pouco provável que o observado seja um artefato do cultivo, pois ao passo que o formato dos tricomas e as dimensões celulares nesse táxon variam em isolados de acordo com o meio de cultura utilizado, as características do ciclo de vida permanecem estáveis sob diferentes condições de cultura, motivo pelo qual são constantemente utilizadas para a distinção de espécies (MATEO et al., 2011). Dessa forma, apesar de morfológicamente poder ser identificada como pertencente ao gênero *Nostoc* de acordo com critérios clássicos, CENA67 não possui algumas das características típicas descritas para este gênero.

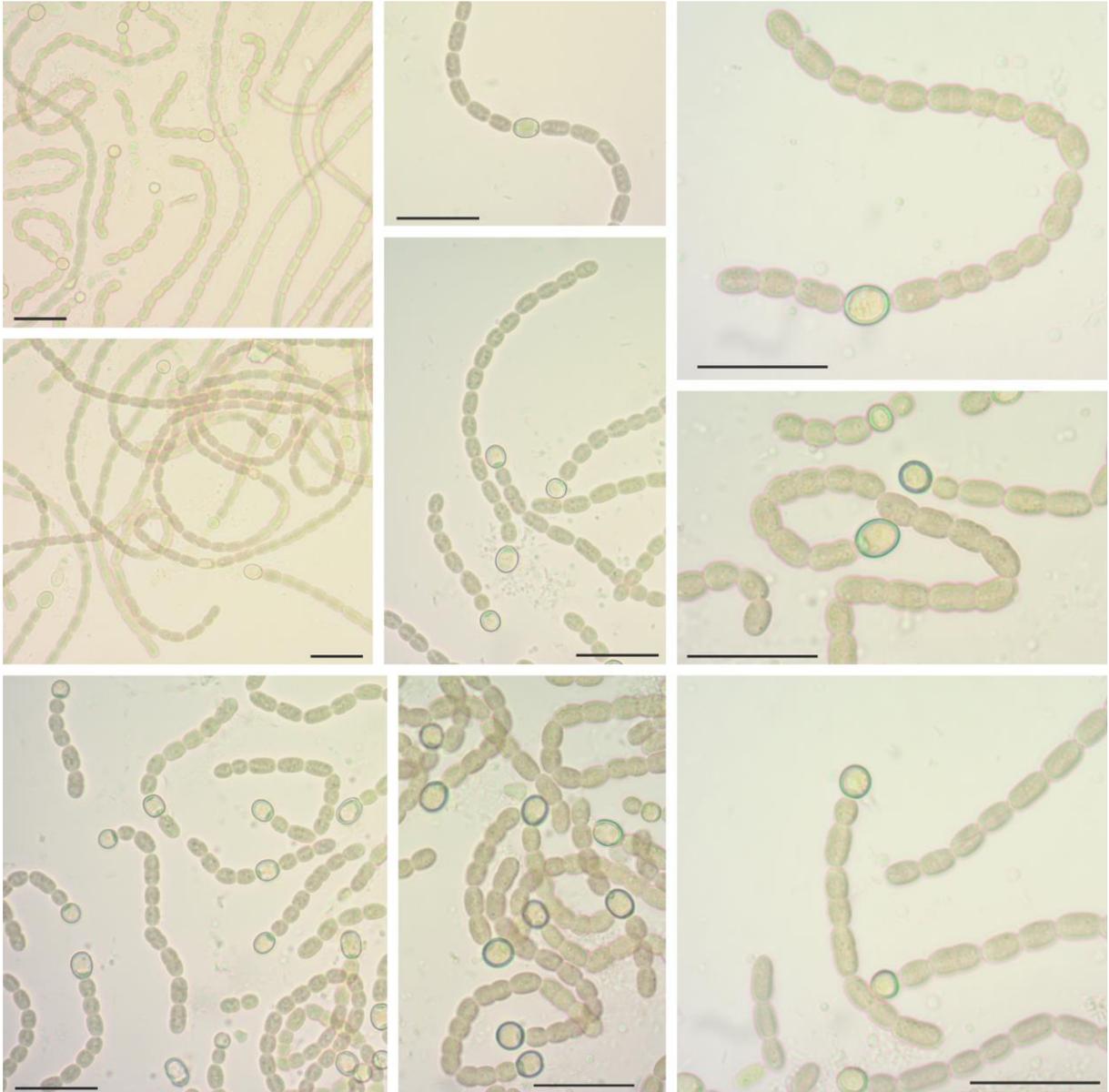


Figura 2 – Fotomicrografias ópticas evidenciando a morfologia de CENA67. Escala: 20 μm

O sequenciamento do DNAr 16S de *Nostoc* sp. CENA67 gerou uma sequência de 1411 pb, a qual apresenta maior proximidade com *Nostoc ellipsoforum* V, linhagem com a qual compartilha 97 % de identidade. Todavia, apesar da relativa alta identidade, a análise filogenética desse gene também apontou haver divergências entre a linhagem CENA67 e o restante das cianobactérias atualmente classificadas dentro do gênero *Nostoc* (Figura 3). Como pode ser observado, CENA67 se comporta filogeneticamente como um grupo externo a um grande número de linhagens atualmente classificadas no gênero *Nostoc* e em alguns gêneros relacionados, como *Desmonostoc*, *Halotia*, *Nodularia*, *Roholtiella* e *Trichormus*, bem como a algumas linhagens de *Anabaena*.

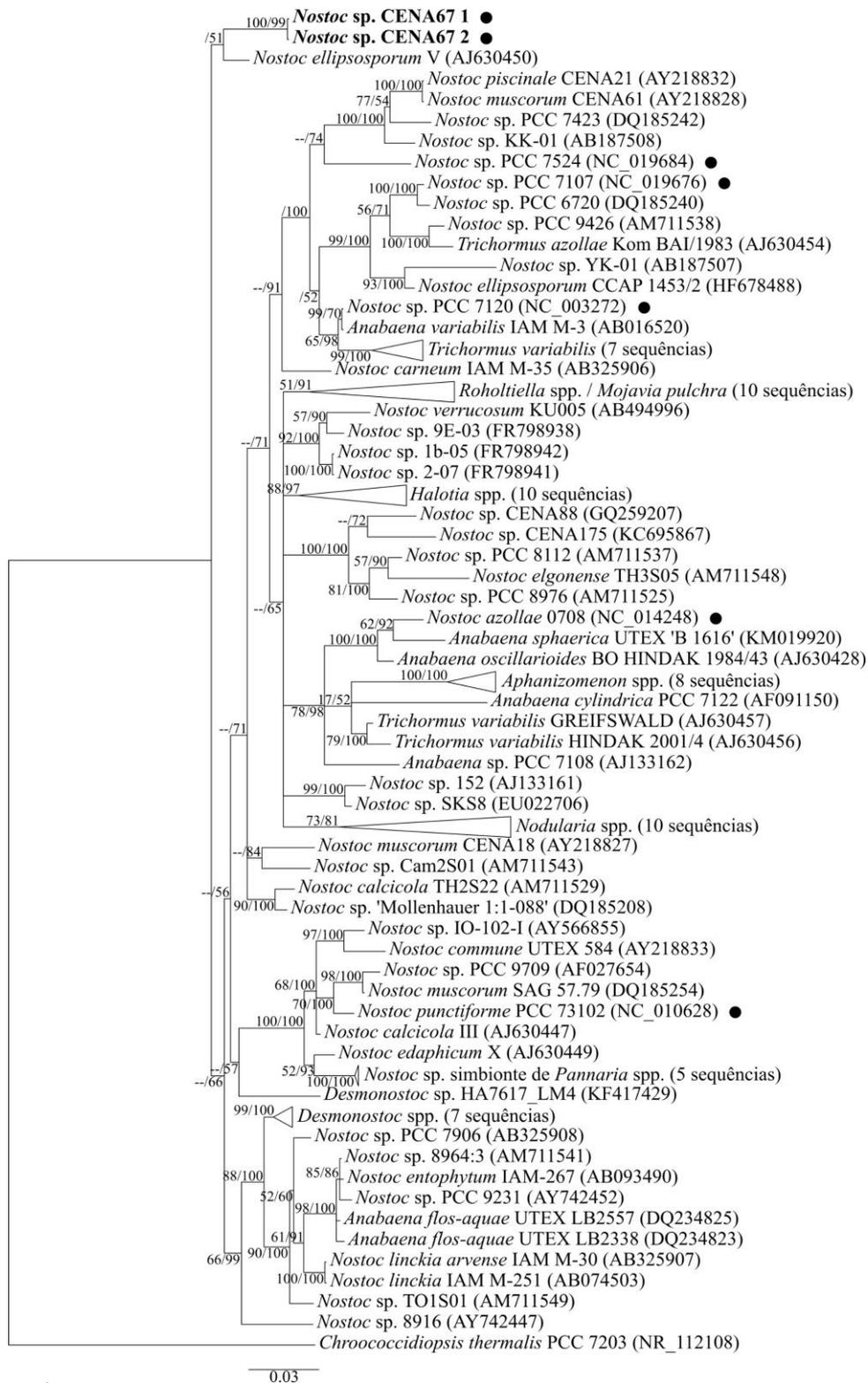


Figura 3 – Árvore filogenética do DNAr 16S de *Nostoc* sp. CENA67 gerada por inferência bayesiana. As sequências obtidas no presente trabalho se encontram destacadas em negrito. Linhagens já classificadas no gênero *Nostoc* com genomas disponíveis se encontram destacadas com pontos. Nos nós dos ramos encontram-se as probabilidades posteriores bayesianas precedidas por percentuais de reamostragem maiores que 50 % em ramos também observados por máxima verossimilhança

Percentuais de identidade de DNAr 16S entre gêneros diferentes na ordem Nostocales podem chegar a 98 %, maior que o observado entre gêneros de outras ordens. Por consequência, sequências providas de cianobactérias da ordem Nostocales frequentemente possuem posições instáveis em árvores filogenéticas e são atraídas por sequências não relacionadas, tornando difícil a separação de gêneros pela utilização exclusiva de sequências de DNAr 16S. Isso provavelmente ocorre devido ao ancestral dessa ordem ter divergido muito cedo das outras cianobactérias, o que deixou em seus descendentes alta distância evolutiva em relação aos seus táxons mais próximos e menor distância entre si (CASAMATTA; GOMEZ; JOHANSEN, 2006; KORELUSOVÁ, 2008). Nesses casos, sequências de genes funcionais podem fornecer um sinal filogenético mais claro, permitindo distinguir entre cianobactérias de alta similaridade morfológica e ribossômica (HENSON et al., 2002).

Marcada por uma história bastante conturbada, a sistemática de cianobactérias foi alvo de diversas tentativas de conciliar morfologia, filogenética e outros fatores, principalmente na última década (HOFFMANN; KOMÁREK; KAŠTOVSKÝ, 2005; KOMÁREK, 2010b; KOMÁREK, 2014). Diversas presunções baseadas em critérios morfológicos têm sido derrubadas à medida que aumenta o conhecimento da diversidade desses micro-organismos. Uma das mais profundas reavaliações ocorreu após a percepção de que a ocorrência de filamentos ramificados, anteriormente considerada monofilética, surgiu independentemente em um grande número de táxons do filo (GUGGER; HOFFMANN, 2004; SINGH et al., 2015a). Mesmo ao longo dessas reavaliações, a definição da família Nostocaceae permaneceu estável, e nesse momento ainda inclui exclusivamente gêneros com filamentos não ramificados (KOMÁREK et al., 2014). Contudo, novos gêneros com ramificação vêm sendo identificados nessa família por evidências filogenéticas (BOHUNICKÁ et al., 2015; KAŠTOVSKÝ et al., 2014).

Esses fatos demonstram mais uma vez que, quando sozinhos, os caracteres morfológicos, mesmo os mais tradicionais e os tidos como estáveis, não são capazes de guiar com segurança a inferência de padrões evolutivos no filo Cyanobacteria. De maneira semelhante, a filogenética baseada exclusivamente em sequências do DNAr 16S tem se mostrado limitada para alguns grupos taxonômicos, como observado entre gêneros da ordem Nostocales. A taxonomia polifásica, baseada no consenso de diferentes análises, foi adotada pela sistemática cianobacteriana para tentar contornar essa limitação (KOMÁREK, 2005; VANDAMME, 1996). Formas de aprimorar esse método continuam a ser propostas (MISHRA et al., 2015; PALINSKA; SUROSZ, 2014). Atualmente, o método mais promissor para a resolução dos problemas da sistemática bacteriana é a inclusão de análises de genomas

(CHUN; RAINEY, 2014; THOMPSON et al., 2015; VANDAMME; PEETERS, 2014). Contudo, o baixo número de genomas cianobacterianos disponíveis inviabiliza a implementação desse método em larga escala para este filo.

Não obstante, as diferenças morfológicas, ecológicas e filogenéticas entre CENA67 e outras cianobactérias identificadas no gênero *Nostoc* observadas neste trabalho são indicativos da possibilidade de a linhagem estudada representar um novo táxon, como uma nova espécie de *Nostoc* ou mesmo um novo gênero da família Nostocaceae. Os dados levantados, contudo, ainda não são suficientes para determinar com exatidão o reposicionamento taxonômico, visto que um baixo número de organismos relacionados a essa linhagem está disponível para análise no momento.

6.2 Sequenciamento, montagem e anotação genômica

De maneira geral, os dados sequenciados com a tecnologia Illumina apresentaram valores de qualidade Phred consideravelmente mais elevados do que os obtidos com a tecnologia Ion Torrent (Figura 4). O sequenciamento com Ion PGM gerou um total de 1.641.285.796 bases, provenientes de 9.365.430 leituras, cujos comprimentos variaram entre valores de 50 e 572 pb após o pré-processamento, onde a maioria das leituras remanescentes continha entre 201 e 250 pb de comprimento (Figura 5A). O sequenciamento com MiSeq gerou 39.254.914 leituras, cuja sequência mais curta continha 50 pb de comprimento enquanto a sequência mais longa possuía 592 pb de comprimento após o pré-processamento. O processamento das sequências obtidas com MiSeq resultou em um conjunto de dados de aproximadamente 11 GB contendo 16.666.240 sequências, totalizando 4.961.016.101 bases. O comprimento da maioria das sequências no conjunto de dados final esteve entre 251 e 300 pb (Figura 5B).

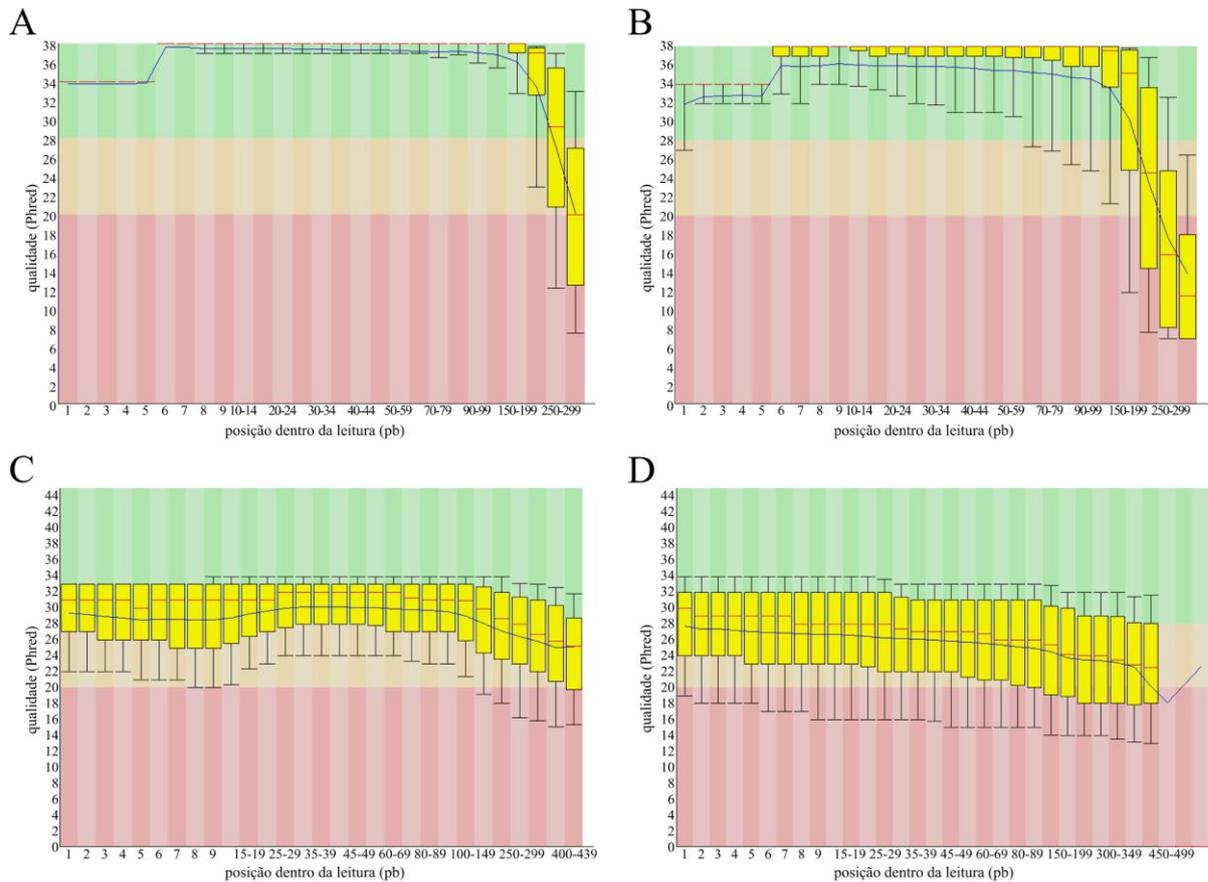
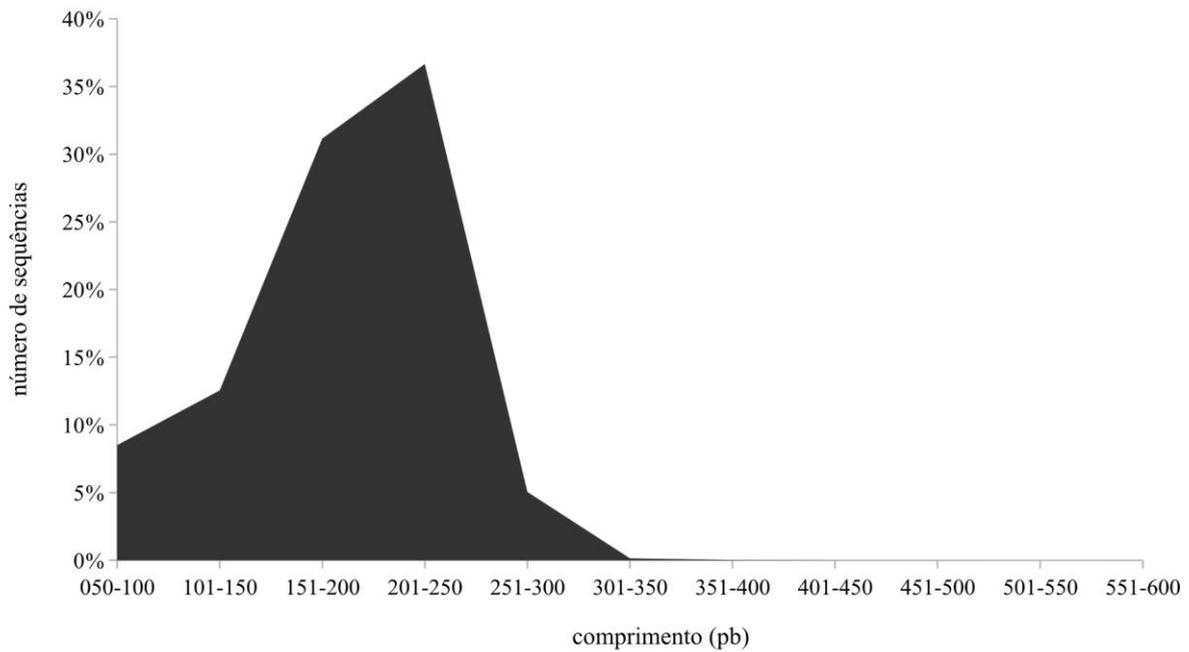


Figura 4 – Qualidade de leituras obtidas em sequenciamentos de uma cultura de *Nostoc* sp. CENA67. A e B: Conjunto de pares R1 e R2 (respectivamente) gerados por sequenciamento com MiSeq; C e D: Pares F e R (respectivamente) de leituras sequenciadas com Ion PGM

Apesar de o sequenciamento realizado com a tecnologia Ion Torrent ter sido capaz de gerar algumas leituras que atingiram comprimentos próximos de 600 pb, a grande maioria das sequências não ultrapassou 350 pb após os procedimentos de filtragem, ao passo que as filtragens de sequências de MiSeq resultaram em uma melhor distribuição do comprimento de suas leituras. Alguns trabalhos afirmam que o comprimento das leituras tem menor importância para a montagem que o método utilizado (CHAISSON; BRINZA; PEVZNER, 2009). Todavia, é inegável que sequências mais longas e de melhor qualidade beneficiam as montagens e aumentam sua confiabilidade, especialmente em casos de amostras com grande número de micro-organismos, como em metagenomas, ou de genomas com alto número de repetições, como os encontrados em cianobactérias.

A



B

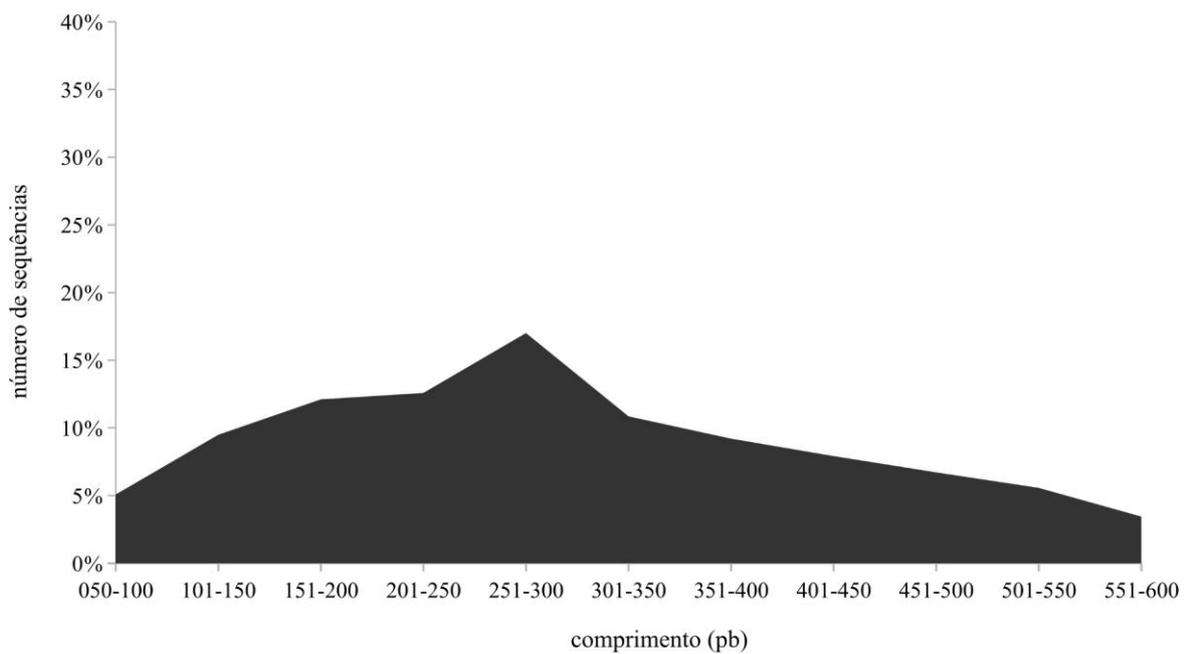


Figura 5 – Distribuição dos comprimentos das sequências remanescentes após o pré-processamento dos dados obtidos com os sequenciamentos. A: Sequências obtidas com Ion PGM; B: Sequências obtidas com MiSeq

Apesar disso, mesmo sequências menores podem ajudar nos processos de montagem. Nas novas gerações de tecnologias de sequenciamento de DNA, o comprimento e o grau de confiabilidade das leituras é dependente da relação sinal/ruído, cujas fontes diferem em cada tecnologia. Por consequência, as sequências produzidas por cada tecnologia também estão suscetíveis a diferentes tipos de erros. Os erros mais comumente encontrados na tecnologia Illumina são de substituição, onde uma base é confundida com outra, enquanto os erros mais frequentes na tecnologia Ion Torrent são ocasionados por homopolímeros, onde repetições sequenciais de uma mesma base são subestimadas ou superestimadas, resultando em inserções ou deleções artificiais (MARDIS, 2013). Logo, uma combinação de ambas as tecnologia de sequenciamento pode ser benéfica, já que é possível que os enganos de um sequenciador sejam corrigidos com os dados obtidos com outro, levando menos erros e ambiguidades aos programas montadores.

A montagem gerou 34.429 sequências contíguas, que totalizaram aproximadamente 182 Mpb, com um N50 de 44.962 e $36 \times$ de cobertura média. A predição de genes apontou 197.635 genes únicos na montagem. Dentre essas sequências, foram detectadas 267 sequências cianobacterianas, totalizando aproximadamente 8,2 Mpb, com N50 de 58.528 e uma cobertura média de $224 \times$. A montagem do genoma de um genótipo ou uma espécie individual a partir de dados metagenômicos pode ser realizada com precisão quando a cobertura deste genoma dentro do metagenoma é de pelo menos $20 \times$; caso contrário, corre-se maior risco de geração de quimeras dos genomas de linhagens próximas (LUO et al., 2011). Dessa maneira, o genoma cianobacteriano montado apresenta baixo risco de quimeras.

O programa Kmer Spectrum Analyzer estimou o tamanho do genoma cianobacteriano em 7,8 Mpb, próximo ao total das sequências detectadas na montagem. Além disso, o mais próximo genoma cianobacteriano completo, pertencente à cianobactéria *Nostoc* sp. PCC 73102, contém um total de 8,23 Mpb. Estimativas de tamanho genômico por métodos bioinformáticos apresentam congruência com estimativas por eletroforese em gel de campo pulsado (WILLIAMS et al., 2013). Devido à dificuldade do corte por enzimas de restrição em decorrência do alto percentual de metilação no genoma de cianobactérias, também é possível que o tamanho genômico tenha maior precisão na estimativa com bioinformática que com análises baseadas em restrição e eletroforese. Em vista desses fatos e apesar do número de sequências contíguas obtidas, o genoma montado possivelmente está próximo de completo.

Não obstante a alta cobertura obtida com o sequenciamento e a montagem do genoma da cianobactéria-alvo, algumas regiões genômicas não puderam ser contempladas com este procedimento, deixando-o em estado de rascunho. Um dos motivos mais prováveis para tal

resultado é um número elevado de regiões genômicas de repetição. Situação semelhante é comumente encontrada em montagens de genomas de outras cianobactérias (WANG et al., 2012). Em outras bactérias, repetições genômicas e elementos transponíveis também atuam consistentemente na interrupção de montagens, podendo ser geralmente observada uma correlação entre as regiões mais fragmentadas do genoma e a predição de ilhas genômicas (RICKER et al., 2012).

Uma estratégia que vem sendo proposta para contornar os problemas inerentes à montagem de regiões genômicas complexas é a utilização de leituras com até dezenas de Kpb de comprimento (KOREN; PHILLIPPY, 2015). Sequências de comprimentos maiores que os possíveis de serem obtidos atualmente têm se tornado viáveis com as tecnologias SMRT, comercializada pela Pacific BioSciences (CHIN et al., 2013; HUDDLESTON et al., 2014; ZHANG et al., 2012), e MinION, da Oxford Nanopore (QUICK; QUINLAN; LOMAN, 2014; SCHNEIDER; DEKKER, 2012), porém com a desvantagem de terem menor qualidade que as sequências obtidas com outras tecnologias atuais, o que requer maior cobertura. Essa desvantagem busca ser contornada por um novo método de preparo de bibliotecas disponibilizado para a tecnologia Illumina, o qual utiliza leituras marcadas de alta qualidade que são reconstituídas em leituras longas sintéticas, que atingem até 18 Kpb de comprimento (MCCOY et al., 2014).

Sequências obtidas com a tecnologia SMRT, sozinhas ou combinadas com outras tecnologias, vem sendo utilizadas com sucesso na montagem dos genomas de algumas cianobactérias em culturas axênicas (BILLER et al., 2014; KOPF et al., 2014; YAMAGUCHI et al., 2015). A eficácia de tal método para a montagem de genomas provenientes de culturas mistas, contudo, ainda permanece por ser testada, o que não permite afirmar se ele beneficiaria a finalização do genoma estudado. A despeito disso, os sequenciamentos com leituras longas se mostram promissores para esse fim.

A anotação detectou 7.190 genes no rascunho genômico de *Nostoc* sp. CENA67, os quais foram identificados em 412 subsistemas (Figura 6). A maior parte dos genes encontrados (73 %) está classificada fora de subsistemas, ou seja, pertence a grupos de famílias proteicas ligadas a processos biológicos ainda desconhecidos. Desses genes, aproximadamente 70 % codifica proteínas hipotéticas. Isso significa que há uma alta diversidade genética inexplorada no genoma estudado, fornecendo amplo material para a descoberta de vias potencialmente envolvidas na biossíntese de produtos naturais ainda desconhecidos (MICALLEF et al., 2015; SHIH et al., 2013).



Figura 6 – Distribuição de subsistemas do genoma da linhagem cianobacteriana CENA67

Certos micro-organismos fotossintéticos também são atraentes para a recuperação de áreas contaminadas devido às suas capacidades de biodegradação de xenobióticos e de biossorção de metais pesados, a despeito de sua baixa adoção neste campo (DE PHILIPPIS et al., 2011; IDI et al., 2015; KAPLAN, 2013). Algumas anotações genômicas observadas sugerem que a cianobactéria estudada poderia ser útil em processos de biorremediação de solos contaminados com xenobióticos, já que genes que codificam enzimas para as vias de degradação de benzoato, fluorobenzoato, cloro-ciclo-hexano, clorobenzeno e tolueno foram detectados. Genes para resistência a alguns metais pesados como cobre, mercúrio, cobalto, zinco, cádmio e arsênio também foram observados entre as sequências genômicas dessa linhagem cianobacteriana.

Genes para a biossíntese de auxinas foram anotados no genoma, corroborando resultados obtidos anteriormente em ensaios bioquímicos (dados não exibidos). Ácido indol acético tem um papel importante na colonização de raízes e na promoção de crescimento de plantas por cianobactérias do gênero *Nostoc* (HUSSAIN et al., 2015). *Nostoc* spp. também podem beneficiar plantas através do antagonismo a fitopatógenos e insetos herbívoros por intermédio de metabólitos secundários (SINGH et al., 2015b). Micro-organismos de solo, tanto abundantes quanto raros, possuem grande influência sobre o desenvolvimento de plantas, incluindo aspectos como taxa de crescimento, resistência a doenças e tolerância a estresse, de modo que a seleção sobre comunidades microbianas de solo pode resultar na alteração de algumas funções vegetais (PANKE-BUISSE et al., 2015). Tendo em vista que a cianobactéria estudada é proveniente de solo com traços interessantes para o cultivo de plantas, é possível que seja capaz de produzir outras substâncias que também tenham papel importante no estabelecimento de interações ecológicas nesse hábitat.

6.3 Detecção de metabólitos secundários

Apesar de ter sido detectado anteriormente por PCR e sequenciamento, não foi detectado dentre as sequências montadas o gene *sxtI*, supostamente envolvido na síntese de saxitoxina. Em contrapartida, foram detectados regiões contendo identidade com os genes *sxtV* (86 %) e *sxtW* (88 %) de *Cylindrospermopsis raciborskii* T3, ambas contidas na porção inicial de uma sequência contígua com 16.804 pb de comprimento. Análises de outros genes dessa sequência contígua não apontaram outras regiões com identidade com genes previstos como participantes da via biossintética das saxitoxinas. Os genes *sxtV* e *sxtW*, que codificam, respectivamente, ferredoxina 4-Fe e 4-S e um homólogo de fumarato redutase/succinato desidrogenase (KELLMANN et al., 2008), ainda não foram encontrados em genomas de outros táxons. Esse par estaria envolvido na transferência de elétrons para SxtH e SxtT, com SxtV produzindo fumarato pela extração de um par de elétrons do succinato e os transferindo para SxtW (PEARSON et al., 2010).

A suposta via biossintética da saxitoxina é altamente complexa, com mais de 35.000 pb de comprimento e, em média, genes para 26 proteínas, com 30 funções catalíticas (KELLMANN et al., 2008). Em algumas linhagens de *Lyngbya wollei*, um agrupamento gênico de saxitoxina codificado por aproximadamente 36.000 pb pode ser encontrado com diferenças bastante significativas quando comparado aos agrupamentos gênicos das outras linhagens, tais como alguns genes adicionais exclusivos e a inativação do gene que codifica uma *O*-carbamoiltransferase, *sxtI* (MIHALI et al., 2011). Dessa forma, o gene *sxtI* não é essencial para a produção de saxitoxina nessas linhagens. Em contrapartida, o gene *sxtA*, que codifica uma policetideo sintase envolvida nos primeiros estágios da síntese de saxitoxina, pode ser detectado também em linhagens não produtoras (CIRÉS et al., 2014). Logo, é possível que, apesar da aparente conservação da via biossintética prevista para saxitoxinas, variações notáveis possam ser encontradas em genomas de diferentes táxons. É também possível que o gene *sxtI*, que foi encontrado anteriormente em CENA67, não seja essencial para a produção de saxitoxinas e possa estar envolvido em outros processos celulares, assim como os genes *sxtV* e *sxtW*. Análises químicas recentes não apontam a produção de saxitoxina em culturas da linhagem estudada (dados não exibidos).

Foi demonstrado que uma linhagem de *Nostoc* sp. produz maiores quantidades de microcistina e nostoficina em condições de estresse fisiológico, fato que mostra que a regulação da produção de cianotoxinas e outros metabólitos secundários pode ter alguma relação com certas condições ambientais, tais como temperatura, irradiância e disponibilidade

de nitrogênio e fósforo, variando entre táxons que ocupam diferentes nichos ecológicos (KURMAYER, 2011). Dessa maneira, é possível que, ao longo do processo de cultivo, tenha sido selecionada uma subpopulação não produtora de saxitoxina que, contudo, continha alguns genes remanescentes do agrupamento, os quais foram recrutados para outras funções celulares ou permaneceram inativados por regulação gênica.

A previsão de metabólitos secundários com antiSMASH com o algoritmo-padrão resultou na predição de 31 agrupamentos gênicos, que incluíam os seguintes: seis agrupamentos potencialmente envolvidos na produção de terpenos; cinco agrupamentos contendo tanto genes que codificam peptídeo sintetases não ribossômicas (NRPS) quanto genes codificando policetídeo sintases (PKS) tipo I; três contendo NRPS; dois que podem ser responsáveis pela produção de microviridinas; dois possivelmente envolvidos na síntese de bacteriocinas, bem como dois na síntese de lantipeptídeos; dois que contêm tanto genes que codificam PKS tipo I quanto genes que codificam PKS ainda não classificadas; um agrupamento com PKS tipo I; um agrupamento com PKS não classificada; e agrupamentos que poderiam estar envolvidos com a síntese de laço-peptídeo, laderano e cianobactina. Também foram encontrados quatro agrupamentos classificados como “outro”, um termo utilizado para aquelas sequências que, apesar de terem seu envolvimento na síntese de metabólitos secundários previsto com sucesso, ainda não podem ser enquadradas em uma categoria específica (Tabela 3).

A detecção com o algoritmo ClusterFind apontou 66 agrupamentos adicionais no rascunho genômico dessa linhagem, incluindo os seguintes: quarenta e nove agrupamentos potencialmente envolvidos na produção de sacarídeos; três agrupamentos que possivelmente resultam em moléculas híbridas sacarídeo-terpeno e um, em sacarídeo-bacteriocina; dois potencialmente envolvidos na síntese de ácidos graxos; e um agrupamento que poderia gerar um sacarídeo modificado por PKS de outro tipo. Esse algoritmo também detectou dezessete agrupamentos hipotéticos que não foram passíveis de categorização. Treze dentre os agrupamentos preditos com o algoritmo-padrão apresentaram similaridade com genes envolvidos na produção de metabólitos secundários já caracterizados. Todavia, apenas um desses agrupamentos hipotéticos apresentou 100 % de genes similares com algum agrupamento conhecido, o do sesquiterpeno geosmina, enquanto o percentual dos outros agrupamentos preditos variou entre 6 e 75 % (Tabela 3). Apenas quatro dentre os agrupamentos gênicos preditos adicionalmente com o algoritmo ClusterFind apresentaram similaridade com agrupamentos conhecidos, que não passou de 14 %.

Tabela 3 – Agrupamentos gênicos preditos no genoma de *Nostoc* sp. CENA67

Agrupamento	Tamanho (pb)	Categoria	Agrupamento Mais Similar	Genes Similares (%)
1	22.731	laço-peptídeo	não identificado	–
2	58.765	PKS tipo I	crocacina	47
3	50.025	NRPS / PKS tipo I	nostopeptolídeo	50
4	55.670	NRPS	aeruginosida	11
5	20.254	microviridina	não identificado	–
6	19.410	bacteriocina	não identificado	–
7	22.506	terpeno	não identificado	–
8	10.656	bacteriocina	não identificado	–
9	32.055	NRPS / PKS tipo I	criptoficina	37
10	23.231	outra PKS	ambiguina	6
11	19.931	terpeno	não identificado	–
12	51.142	NRPS	nodularina	50
13	20.932	terpeno	não identificado	–
14	26.956	terpeno	geosmina	100
15	32.951	outro	não identificado	–
16	24.056	outro	não identificado	–
17	59.994	NRPS / PKS tipo I	nostoficina	27
18	19.335	bacteriocina-lantipeptídeo	não identificado	–
19	41.753	PKS tipo I / outra PKS	glicolipídeo (heterócitos)	42
20	26.330	cianobactina	não identificado	–
21	29.836	PKS tipo I / outra PKS	glicolipídeo (heterócitos)	71
22	13.587	microviridina	microviridina K	50
23	35.823	NRPS	cianopectolina	42
24	23.452	bacteriocina-lantipeptídeo	não identificado	–
25	25.765	NRPS / PKS tipo I	cianopectina	75
26	17.926	NRPS / PKS tipo I	não identificado	–
27	15.483	laderano	não identificado	–
28	14.309	terpeno	não identificado	–
29	13.169	outro	não identificado	–
30	10.912	terpeno	não identificado	–
31	3.090	outro	não identificado	–

Diversos dos metabólitos secundários conhecidamente sintetizados por cianobactérias são produzidos por meio de vias não ribossômicas com a participação de enzimas como peptídeo sintetases não ribossômicas e/ou policetídeo sintases, e atuam em importantes processos celulares, incluindo captação de ferro e fixação de nitrogênio (CALTEAU et al.,

2014). Vias que contêm ambos os tipos de enzimas, NRPS e PKS, são frequentemente encontradas na biossíntese de produtos naturais cianobacterianos, assim como enzimas híbridas que combinam domínios tanto de NRPS quanto de PKS em uma mesma cadeia polipeptídica (KEHR; PICCHI; DITTMANN, 2011). Aproximadamente 70 % das cianobactérias apresenta essas enzimas, cujos genes totalizam por volta de 5 % dos genomas que as contêm, com média de cinco agrupamentos de NRPS e/ou PKS por genoma (SHIH et al., 2013). No genoma estudado, o dobro desse valor, ou dez agrupamentos, foram preditos e estimados com um total de 481.985 pb, ou aproximadamente 5,86 % do genoma montado, percentual que, entretanto, pode ter sido superestimado pelo algoritmo de descoberta adotado.

Metabólitos secundários peptídicos podem ser gerados por modificações pós-traducionais de proteínas ribossômicas, com a catálise de macrociclizações do peptídeo e de alterações nas cadeias laterais de seus aminoácidos, ou pela ação de peptídeo sintetases não ribossômicas (NRPS), enzimas multimodulares que contêm domínios responsáveis pela incorporação de aminoácidos à cadeia polipeptídica e pela modificação dos mesmos. Enquanto a síntese ribossômica está geralmente restrita a uma amostra de 20 aminoácidos, as NRPS podem utilizar cerca de 300 substratos proteinogênicos e não proteinogênicos, gerando, assim, maior diversidade molecular e funcional e se responsabilizando pela produção da maior parte dos peptídeos atualmente conhecidos em cianobactérias. Os domínios de condensação, de adenilação e carreador de peptidil são essenciais em NRPS e podem ser complementados pelos domínios de metiltransferase, cetorredutase e epimerase (KEHR; PICCHI; DITTMANN, 2011).

De maneira similar às NRPS, policetídeo sintases (PKS) são enzimas multimodulares que atuam na síntese não ribossômica de metabólitos secundários, porém utilizam ácidos carboxílicos como substrato para a síntese de policetídeos. Os domínios de cetossintase, aciltransferase e carreador de acila são geralmente encontrados nessas moléculas e podem ser acompanhados pelos domínios de cetorredutase, desidratase e enoil redutase, os quais podem levar a diferentes estados de redução dos grupos funcionais ceto do produto (KEHR; PICCHI; DITTMANN, 2011). As PKS são tradicionalmente divididas em três tipos: I. enzimas multifuncionais compostas por diversos módulos, que atuam por síntese linear; II. complexos multienzimáticos com um único sítio ativo cada, com síntese iterativa; e III. enzimas com vários módulos, com síntese iterativa; contudo, a diversidade de mecanismos e estruturas nessa família de proteínas é significativamente subestimada por essa divisão (SHEN, 2003). Algumas PKS incomuns que não possuem características claras de nenhum dos três tipos permanecem por ser classificadas adequadamente.

Entre os peptídeos modificados pós-traducionalmente, encontram-se as bacteriocinas, metabólitos secundários que possuem atividade antimicrobiana, cujos genes são amplamente distribuídos em genomas bacterianos (COTTER; ROSS; HILL, 2013). Esses peptídios de baixa massa molecular geralmente se originam com a clivagem de proteínas ribossômicas precursoras seguida de modificações pós-traducionais, tais como formação de lantionina, macrociclicização, desidratação ou heterociclicização, reações realizadas por proteínas codificadas por genes localizados no genoma de maneira adjacente ao gene do precursor. Apesar de pesquisas sobre bacteriocinas terem se focado principalmente sobre bactérias gram-positivas, cianobactérias são uma fonte prolífica de bacteriocinas significativamente distintas das encontradas em outros filos. Nesse filo, baixas taxas de conservação geram diversidade suficiente para permitir sua divisão em sete grupos, de acordo com a composição de seus domínios, de sua organização gênica e de sua filogenia (WANG; FEWER; SIVONEN, 2011).

Lantipeptídeos são uma subclasse de bacteriocinas formadas pela introdução adicional de lantionina e metil-lantionina, aminoácidos que contêm grupos tioéter (WILLEY; DONK, 2007). Alta diversidade de lantipeptídeos em um gênero de cianobactérias foi relatada como sendo resultante das atividades de uma única enzima sobre 29 precursores ribossômicos diferentes, com papel específico ao hábitat ou à comunidade de origem (LI et al., 2010). Grande diversidade de estruturas, atividades, mecanismos biossintéticos e organismos produtores dessas moléculas tem sido encontrada recentemente, e acredita-se que centenas de novos lantipeptídeos ainda permaneçam a ser descobertos em genomas bacterianos (KNERR; DONK, 2012).

De maneira semelhante às bacteriocinas, cianobactinas e laçopectídeos são sintetizados a partir de peptídeos precursores, que são sintetizados ribossomicamente e modificados pós-traducionalmente. Ao passo que cianobactinas passam por grandes transformações pós-traducionais, lembrando em pouco seus peptídeos precursores, laçopectídeos sofrem um número bem menor de modificações (DONIA; SCHMIDT, 2011; MAKSIMOV; LINK, 2014). A família das cianobactinas é composta tanto por moléculas circulares como lineares de atividade anticancerígena, antiviral, antimalárica ou alelopática, cujas vias estão presentes em aproximadamente 24 % dos gêneros cianobacterianos (JASPARS et al., 2014; LEIKOSKI et al., 2013). Laçopectídeos, por outro lado, recebem esse nome por serem caracterizados por uma estrutura em formato de laço ou nó corrediço, e podem apresentar atividade antimicrobiana, antiviral e antimetastática ou atuar como inibidores enzimáticos e antagonistas de receptores, sendo encontrados em 2,5 % dos genomas bacterianos (MAKSIMOV; PELCZER; LINK, 2012; MAKSIMOV; LINK, 2014).

Duas classes distintas de lipídeos foram preditas no genoma da cianobactéria estudada. Laderanos são moléculas constituídas por anéis múltiplos de ciclobutano cuja função ainda é desconhecida e que são geralmente encontradas em bactérias anaeróbicas que oxidam amônia (RATTRAY et al., 2009; CHABAN et al., 2014). Terpenos são moléculas derivadas de isopreno encontradas em todos os organismos vivos que, de maneira geral, proporcionam algumas características mecânicas necessárias a membranas celulares, ou são secretadas, volatilizadas ou catabolizadas em metabólitos com papel fisiológico ou ecológico, das quais 55.000 moléculas já foram descritas (OURISSON, 1990; SCHWAB et al., 2012). O sesquiterpeno geosmina foi a única substância que apresentou 100 % de similaridade com algum dos agrupamentos gênicos preditos. Três genes estão envolvidos em sua síntese, e juntos somam pouco mais de 5 Kpb (Figura 7).

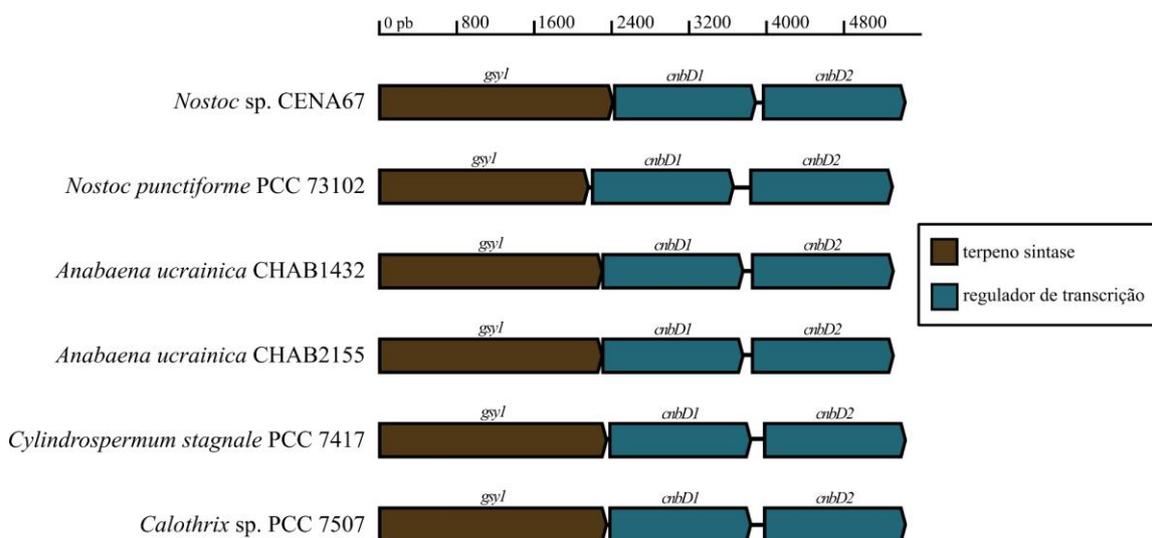


Figura 7 – Estrutura de agrupamentos gênicos presumidamente relacionados à síntese de geosmina em *Nostoc sp. CENA67* e em cianobactérias próximas

Assim como 2-metilisoborneol, a geosmina frequentemente é responsável por alterar as características organolépticas da água potável, já que possui odor e sabor característicos de terra molhada. A biossíntese de geosmina é conhecida em fungos, actinobactérias, proteobactérias e cianobactérias. Actinobactérias são comumente a fonte de geosmina em solos, enquanto cianobactérias são tidas como suas mais importantes produtoras em corpos d'água, especialmente durante episódios de floração. Nesse filo, essa molécula ocorre nas ordens Oscillatoriales e Nostocales, com destaque para espécies do gênero *Nostoc* (SUURNÄKKI et al., 2015).

O outro agrupamento gênico ao qual a síntese de uma molécula já conhecida pôde ser atribuída após curadoria manual das anotações foi o que recebeu o número 22, relacionado à biossíntese de microviridina, que apresentou 50 % de similaridade com o agrupamento biossintético descrito para a variante microviridina K. O agrupamento gênico anotado apresentou doze genes, totalizando aproximadamente 9,5 Kpb, um dos mais longos entre os agrupamentos presentes em outros genomas cianobacterianos que puderam ser descritos (Figura 8). A curadoria manual do outro agrupamento predito na biossíntese de microviridina, que recebeu o número 5, não apontou resultado semelhante.

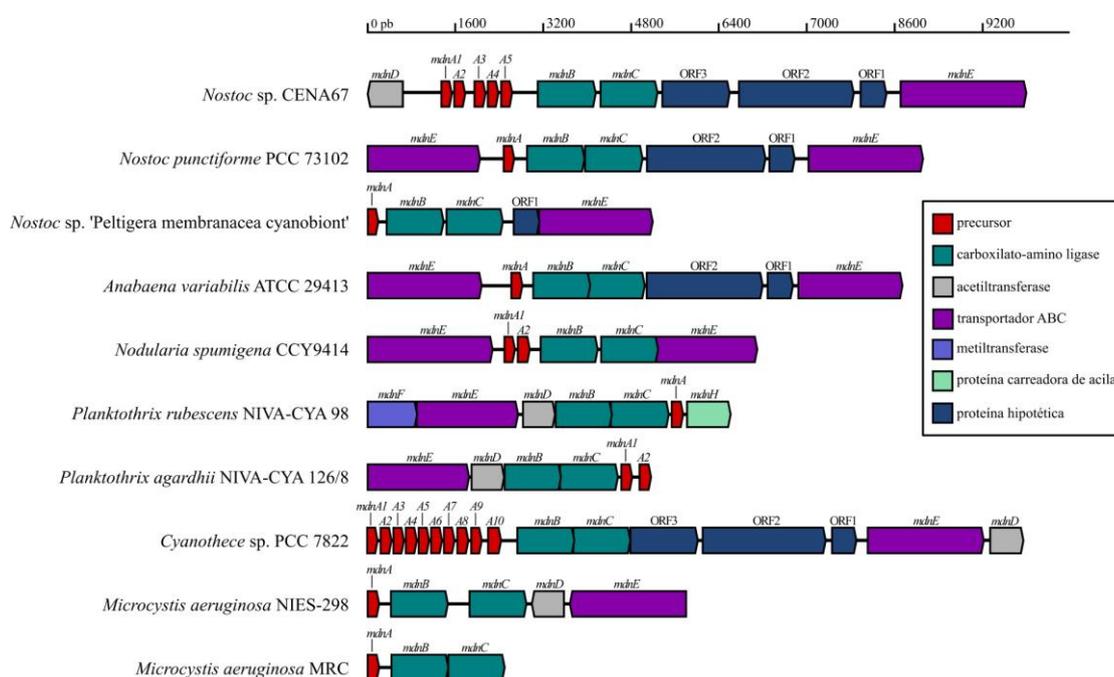


Figura 8 – Estrutura de agrupamentos gênicos presumidamente relacionados à síntese dos inibidores de protease microviridinas no genoma de *Nostoc* sp. CENA67 e nos genomas de outras cianobactérias

Enquanto geosmina apresenta estruturas gênicas com alta conservação, a estrutura de agrupamentos gênicos potencialmente envolvidos na produção de microviridinas é bastante variável. Como observado, a variação mais marcante é observada em relação à quantidade de genes de precursores peptídicos. A diversidade de precursores de microviridina é significativa não apenas quando comparada entre linhagens, mas mesmo dentro de alguns genomas, que podem conter de um a dez genes codificadores de precursores, semelhante ao observado em agrupamentos gênicos envolvidos na biossíntese de cianobactinas (LEIKOSKI et al., 2013). Os cinco genes codificadores de peptídeos precursores de microviridinas encontrados no genoma de CENA67 apresentam identidades que variam de 62 a 91 % entre si e podem gerar de 45 a 49 aminoácidos (Tabela 4).

Tabela 4 – Identidade entre genes que codificam precursores de microviridina encontrados no genoma da linhagem CENA67

Gene	Aminoácidos	<i>mdnA1</i>	<i>mdnA2</i>	<i>mdnA3</i>	<i>mdnA4</i>	<i>mdnA5</i>
<i>mdnA1</i>	45	–				
<i>mdnA2</i>	46	86	–			
<i>mdnA3</i>	47	79	90	–		
<i>mdnA4</i>	45	88	91	86	–	
<i>mdnA5</i>	49	62	80	74	69	–

O alinhamento das sequências de aminoácidos de peptídeos precursores apresenta alta variação, e algumas sequências são especialmente distintas em *Anabaena variabilis* ATCC 29413 e *Nodularia spumigena* CCY9414 (Figura 9). À primeira vista, o alto número de genes que codificam precursores de microviridinas em alguns genomas parece ter se originado a partir de eventos de duplicação gênica. Em diversos casos, as análises filogenéticas de fato parecem fornecer suporte a essa hipótese (Figura 10). Em contrapartida, alguns precursores de espécies diferentes estão mais proximamente relacionados entre si do que com outros precursores do mesmo genoma, o que sugere que possa haver combinações de genes parálogos e xenólogos dentro de alguns agrupamentos. Não é possível descartar, todavia, a hipótese de convergência evolutiva entre esses genes.

```

Nostoc sp. CENA67 mdnA1          MSTNT--VKASDTNA-VPFFARFLEEQSAQ-----ADNT--P-----YV--TKKYPSDWE-DY 70
Nostoc sp. CENA67 mdnA2          MSNNT--VKASDIQA-VPFFARFLEEQSAQ-----PDNT--P-----WA-ETKKYPSDWE-EY 70
Nostoc sp. CENA67 mdnA3          MSNNT--VKASDIQA-VPFFARFLEEQSAQ-----PDNQ--P-----AIPQTLKYPSDWE-DY 70
Nostoc sp. CENA67 mdnA4          MSNNT--VKASDLQA-VPFFARFLEEQSAQ-----ADNT--P-----YT--TKKYPSDWE-EY 70
Nostoc sp. CENA67 mdnA5          MSTS----AALDIQA-VPFFARFLVEEPPK---PSEEEPT--P-----IPIWTLKYPSDWE-ER 70
Nostoc punctiforme PCC 73102 mdnA MPTNT--VKTVDVVA-VPFFARFLEEQATE-----GTEV--P-----W--TYKFPSDLE-DR 70
Nostoc sp. 'Peltigera membranaceae symbiont' mdnA MSSN---L-----NTQA-VPFFARFLVEEKPP-----EKPNDN--P-----PPIYTFKYPSDWE-DR 70
Anabaena variabilis ATCC 29413 mdnA MLQWYRSLPVSWKSKL-QKVQRLLSLGLSS--ILQIWKTNNS-----TILYLKLR----- 70
Nodularia spumigena CCY9414 mdnA1 MGKT--FWFTE-----YPDPLLEEAGDI-----PDLDFTF-----YLIMKLLQKGFPTDS 70
Nodularia spumigena CCY9414 mdnA2 MTT----ATLANIEA-VPFFARFLAAEPP--ETPAPQPEEQPLP-----PPIFTLKYPSDWE-DC 70
Planktothrix rubescens NIVA-CYA 98 mdnA MSKN---IKVSTGSA-VPFFARFLSEQDTET-----GDSTSTDIP-----TI-WTLKYPSDWE-DS 70
Planktothrix agardhii NIVA-CYA 126/8 nvdE/mdnA1 MSKN---VKVSAPKA-VPFFARFLAEQAVE-----ANNSNSA-----PYGNTMKYPSDWE-EY 70
Planktothrix agardhii NIVA-CYA 126/8 nvdF/mdnA2 MSKN---IKVSTGSA-VPFFARFLSEQDTET-----GDSTSTDIP-----TI-WTFKYPSDWE-DS 70
Cyanothece sp. PCC 7822 mdnA1      MSKNT--GKQKEIKA-VPFFARFLEEQAQ-----NETA--P-----YQ-NTLKYPSDWE-DY 70
Cyanothece sp. PCC 7822 mdnA2      MSKNT--GKQKEIKA-VPFFARFLEEQAQ-----NETA--P-----YQ-NTLKYPSDWE-DY 70
Cyanothece sp. PCC 7822 mdnA3      MSKNT--GKAKEIKA-VPFFARFLEEQAQ-----NETA--P-----TV--TRKYPSDWE-DY 70
Cyanothece sp. PCC 7822 mdnA4      MSKNT--GKAKEIKA-VPFFARFLEEQAQ-----NETA--P-----YV--TKKYPSDWE-DY 70
Cyanothece sp. PCC 7822 mdnA5      MSKNT--GKAKEIKA-VPFFARFLEEQAQ-----NETA--P-----YV--TKKYPSDWE-EY 70
Cyanothece sp. PCC 7822 mdnA6      MSKNT--GKAKEIKA-VPFFARFLEEQAQ-----NETA--P-----YV--TKKYPSDWE-EY 70
Cyanothece sp. PCC 7822 mdnA7      MSKNT--GKAKEIKA-VPFFARFLEEQAQ-----NETA--P-----YQ-NTLKYPSDWE-EY 70
Cyanothece sp. PCC 7822 mdnA8      MSKNT--GKQKEIKV-VPFFARFLEEQAQ-----SETEALP-----PA--TLKYPSDWE-EY 70
Cyanothece sp. PCC 7822 mdnA9      MSKNT--GKAREIRAVVPFFARFLEEQAQ-----PAN--P-----LP-FTLKYPSDWE-DE 70
Cyanothece sp. PCC 7822 mdnA10     MSNN---AAALPITNA-KPFFARFLATQEETP---PTEKPAET--PAPFPFPEFPPIFTLKYPSDWE-DS 70
Microcystis aeruginosa NIES-298 mdnA M-----AYPNDQQGKA-LPFFARFLSVSKEESSI---KSPSPE--P-----FTGTTLKYPSDWE-EY 70
Microcystis aeruginosa MRC mdnA   M-----AYPNDQQGKA-LPFFARFLSVSKEES---SIKSPS--PD-----HEISTRKYPSDWE-EW 70
*                                     :*

```

Figura 9 – Alinhamento de sequências de aminoácidos de peptídeos precursores de microviridinas, evidenciando sua variedade

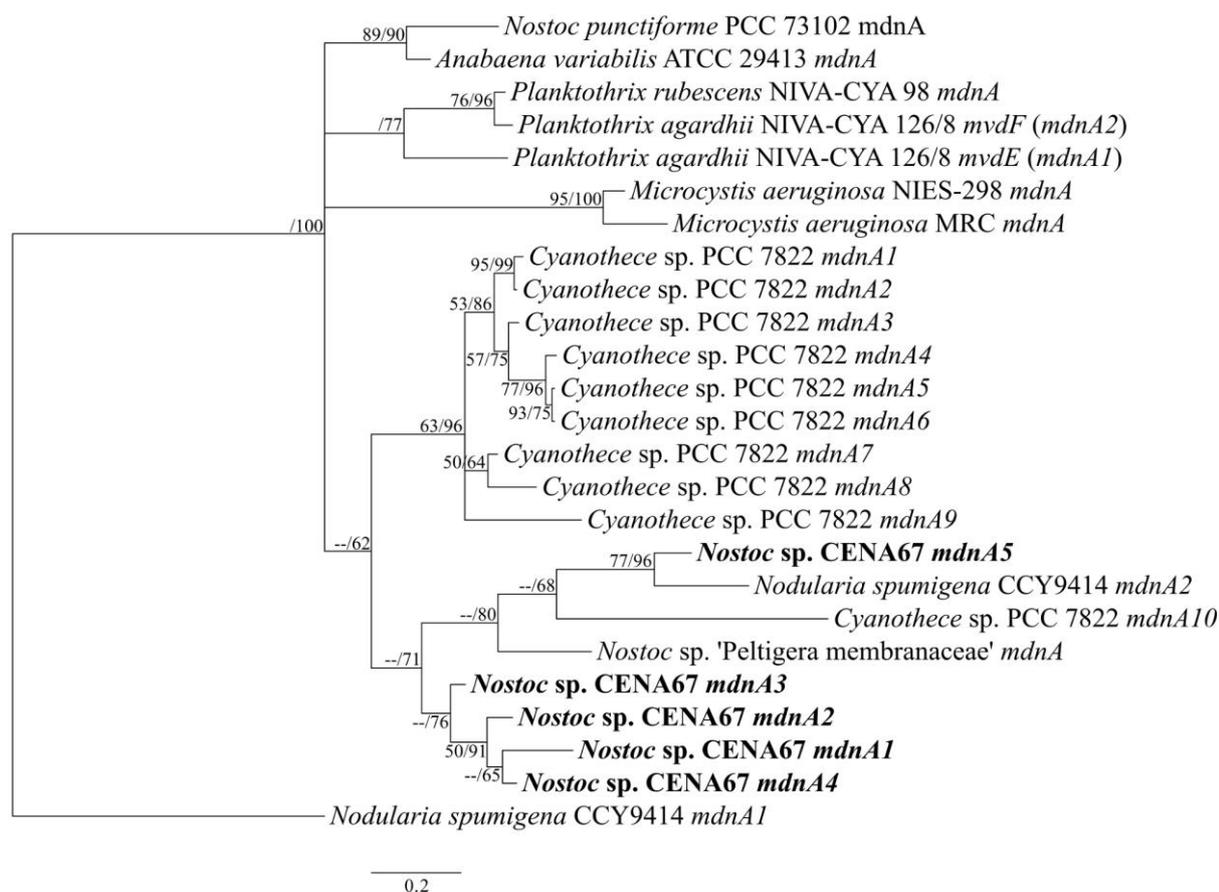


Figura 10 – Árvore filogenética reconstruída por inferência bayesiana com sequências de aminoácidos de peptídeos precursores de microviridinas em genomas de cianobactérias. Sequências da cianobactéria estudada no presente trabalho estão destacadas em negrito. Probabilidades posteriores maiores que 50 % se encontram nos nós dos ramos, precedidas pelos percentuais de reamostragem dos ramos também observados na árvore reconstruída por máxima verossimilhança

A microviridina é um depsipeptídeo tricíclico de ampla ocorrência em alguns gêneros que, na maioria de suas variantes, atua como inibidor de proteases (ZIEMERT et al., 2010). A primeira variante de microviridina, descoberta em 1990, recebeu esse nome por ter sido isolada de *Microcystis viridis* (ISHITSUKA et al., 1990). Novas variantes desse grupo de moléculas continuam a ser encontradas em amostras ambientais e isolados de cianobactérias (GATTE-PICCHI et al., 2014). A diversidade molecular das microviridinas é provavelmente maior que o que se acredita atualmente, tendo em vista que, conforme constatado nas predições de agrupamentos gênicos, há maior diversidade de precursores que o conhecido atualmente, que possivelmente permitem a expressão de novas variantes.

O sequenciamento do genoma da actinobactéria *Streptomyces coelicolor* é um importante marco na descoberta de produtos naturais por métodos genômicos (BENTLEY et al., 2002). A maioria das atuais ferramentas de predição de metabólitos secundários foi

desenvolvida com base em vias biossintéticas de actinobactérias, e por consequência disso as análises realizadas pelas mesmas nem sempre se mostram adequadas para genomas cianobacterianos, os quais frequentemente codificam vias para a síntese de moléculas cuja ocorrência em outros filos é incomum (MICALLEF et al., 2015). Logo, é possível que outras vias biossintéticas codificadas pelo genoma-alvo sejam descobertas à medida que as ferramentas bioinformáticas sejam aprimoradas.

Nostoc sp. CENA67 produz moléculas antitumorais que são capazes de inibir o desenvolvimento de células de câncer pulmonar e de carcinoma de cólon, fato que a coloca como uma interessante fonte de moléculas de importância biotecnológica e farmacológica (SILVA-STENICO et al., 2013a). Todavia, as moléculas responsáveis por essa ação não foram caracterizadas química ou geneticamente, portanto não foi possível identificar seus agrupamentos. Em cianobactérias, análises metabolômicas e genômicas têm se mostrado complementares, e sua combinação desponta como uma estratégia sólida para a descoberta de novos produtos naturais, permitindo tanto verificar se um agrupamento gênico hipotético está sendo expresso como identificar os genes envolvidos na produção de uma determinada molécula (KLEIGREWE et al., 2015). A complementação das predições no genoma de *Nostoc* sp. CENA67 com análises químicas pode ajudar a esclarecer alguns dos agrupamentos gênicos órfãos, cujas moléculas ainda não podem ter suas estruturas preditas com precisão.

6.4 Análises genômicas comparativas

Cromossomos de linhagens identificadas no gênero *Nostoc* que possuem sequências genômicas disponíveis podem variar em tamanho de 5,35 a 8,23 Mpb e codificar entre 5.287 e 6.750 genes. Em comparação com essas sequências, o genoma da linhagem *Nostoc* sp. CENA67 apresenta características bastante próximas às da linhagem simbiote *N. punctiforme* PCC 73102, incluindo em tamanho, percentual GC, número de genes e quantidade de sequências codificadoras de RNA (Tabela 5).

Tabela 5 – Características gerais do genoma de *Nostoc* sp. CENA67 comparadas às sequências cromossômicas de linhagens relacionadas ao gênero *Nostoc*

Linhagem	Hábitat de Origem	Tamanho (Mpb)	Percentual GC	Genes	RNA	Referência
<i>Nostoc</i> sp. CENA67	terra preta antropogênica	8,2	41,4	7.190	84	este trabalho
<i>Nostoc</i> sp. PCC 7120 (NC_003272)	água doce	6,41	41,3	5.287	64	KANEKO et al., 2001
<i>N. azollae</i> 0708 (NC_014248)	simbionte de <i>Azolla filiculoides</i>	5,35	38,4	4.989	55	RAN et al., 2010
<i>N. punctiforme</i> PCC 73102 (NC_010628)	simbionte de <i>Macrozamia riedlei</i>	8,23	41,4	6.750	88	MEEKS et al., 2001
<i>Nostoc</i> sp. PCC 7107 (NC_019676)	água doce	6,33	40,4	5.423	90	SHIH et al., 2013
<i>Nostoc</i> sp. PCC 7524 (NC_019684)	fonte termal	6,64	41,5	5.427	62	SHIH et al., 2013

No tocante a agrupamentos gênicos potencialmente envolvidos na síntese de metabólitos secundários, foram preditos 5, 11, 13, 15 e 21 agrupamentos nos genomas de *N. azollae* 0708, *Nostoc* sp. PCC 7107, *Nostoc* sp. PCC 7120, *Nostoc* sp. PCC 7524 e *N. punctiforme* PCC 73120, respectivamente. A exemplo do observado em CENA67, a maioria desses agrupamentos não apresenta similaridade significativa com agrupamentos conhecidos, com exceção de três agrupamentos detectados em PCC 73120 e relacionados à biossíntese de nostopeptolídeo, geosmina e glicolípídeos de heterócitos.

O genoma de *Nostoc* sp. CENA67 tem como vizinho mais próximo o genoma de *N. punctiforme* PCC 73102 (Tabela 6). Com exceção deste, o genoma sequenciado é mais próximo de *Nodularia spumigena* CCY9414 que de outros genomas de cianobactérias identificadas no gênero *Nostoc*, como *Nostoc* sp. PCC 7120, e *Trichormus (Nostoc) azollae* 0708. *Anabaena variabilis* ATCC 29413 também se encontra entre os genomas mais próximos a CENA67. Nota-se, ainda, certa proximidade com genomas do gênero *Cyanothece*, provavelmente em decorrência do baixo número de genomas próximos depositados.

Tabela 6 – Vizinhos mais próximos à linhagem CENA67, de acordo com o servidor RAST

Nº RAST	Pontuação	Linhagem
63737.11	507	<i>N. punctiforme</i> PCC 73102
313624.3	506	<i>Nodularia spumigena</i> CCY9414
63737.4	498	<i>N. punctiforme</i> PCC 73102
103690.1	472	<i>Nostoc</i> sp. PCC 7120
240292.6	471	<i>Anabaena variabilis</i> ATCC 29413
103690.10	462	<i>Nostoc</i> sp. PCC 7120
240292.6	461	<i>Anabaena variabilis</i> ATCC 29413
551115.6	416	<i>Trichormus azollae</i> 0708
533240.4	403	<i>Cylindrospermopsis raciborskii</i> CS-505
533247.5	331	<i>Raphidiopsis brookii</i> D9
203124.1	293	<i>Trichodesmium erythraeum</i> IMS101
65393.13	287	<i>Cyanothece</i> sp. PCC 7424
203124.1	283	<i>Trichodesmium erythraeum</i> IMS101
65393.13	278	<i>Cyanothece</i> sp. PCC 7424
395962.3	264	<i>Cyanothece</i> sp. PCC 8802
395962.3	255	<i>Cyanothece</i> sp. PCC 8802
376219.3	249	<i>Arthrospira</i> sp. PCC 8005
41431.3	242	<i>Cyanothece</i> sp. PCC 8801
118168.3	240	<i>Microcoleus chthonoplastes</i> PCC 7420
497965.6	233	<i>Cyanothece</i> sp. PCC 7822
41431.3	233	<i>Cyanothece</i> sp. PCC 8801
313612.3	201	<i>Lyngbya</i> sp. PCC 8106
449447.3	199	<i>Microcystis aeruginosa</i> NIES-843
449447.3	191	<i>Microcystis aeruginosa</i> NIES-843
43989.3	183	<i>Cyanothece</i> sp. ATCC 51142
497965.6	180	<i>Cyanothece</i> sp. PCC 7822
391612.3	179	<i>Cyanothece</i> sp. CCY 0110
43989.3	174	<i>Cyanothece</i> sp. ATCC 51142
391612.3	160	<i>Cyanothece</i> sp. CCY 0110
395961.4	153	<i>Cyanothece</i> sp. PCC 7425

Como pode ser observado na Figura 11, a maior parte dos genes de CENA67 possui percentual de identidade relativamente baixo em relação aos genes de três de seus vizinhos mais próximos. De maneira semelhante, a análise de sintenia aponta para variação significativa em relação ao genoma de seu vizinho mais próximo, o que significa que provavelmente há um tempo relativamente elevado de divergência entre esses táxons a partir de sua população ancestral (Figura 12).

MRB	100	99.9	99.8	99.5	99	98	95	90	80	70	60	50	40	30	20	10
MRU	100	99.9	99.8	99.5	99	98	95	90	80	70	60	50	40	30	20	10

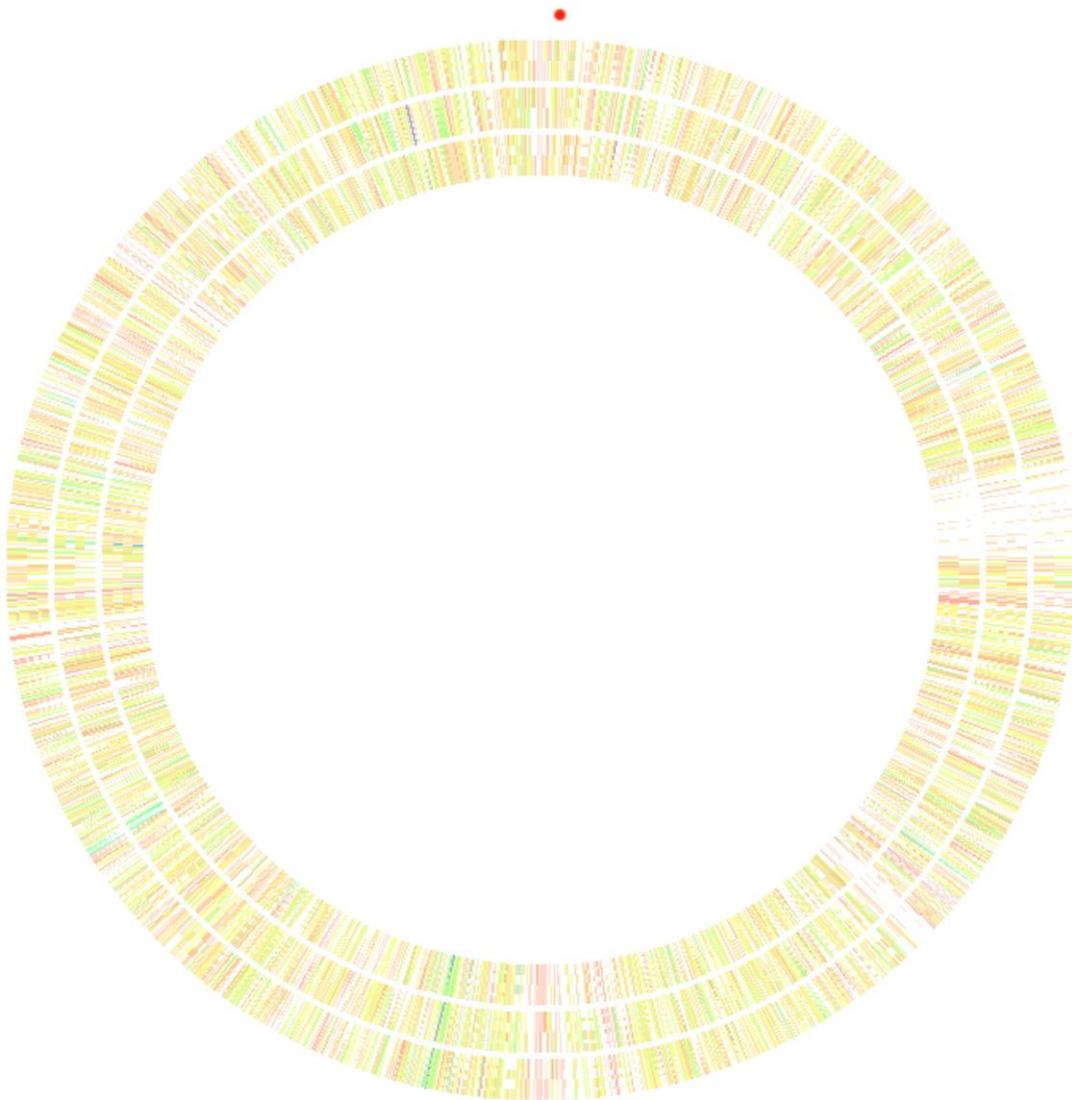


Figura 11 – Percentuais de identidade proteica de CENA67 em comparação com os genomas das linhagens *A. variabilis* ATCC 29413 (círculo exterior), *N. punctiforme* PCC 73102 (centro) e *Nostoc* sp. PCC 7120 (círculo interior). MRB: melhor resultado bidirecional; MRU: melhor resultado unidirecional

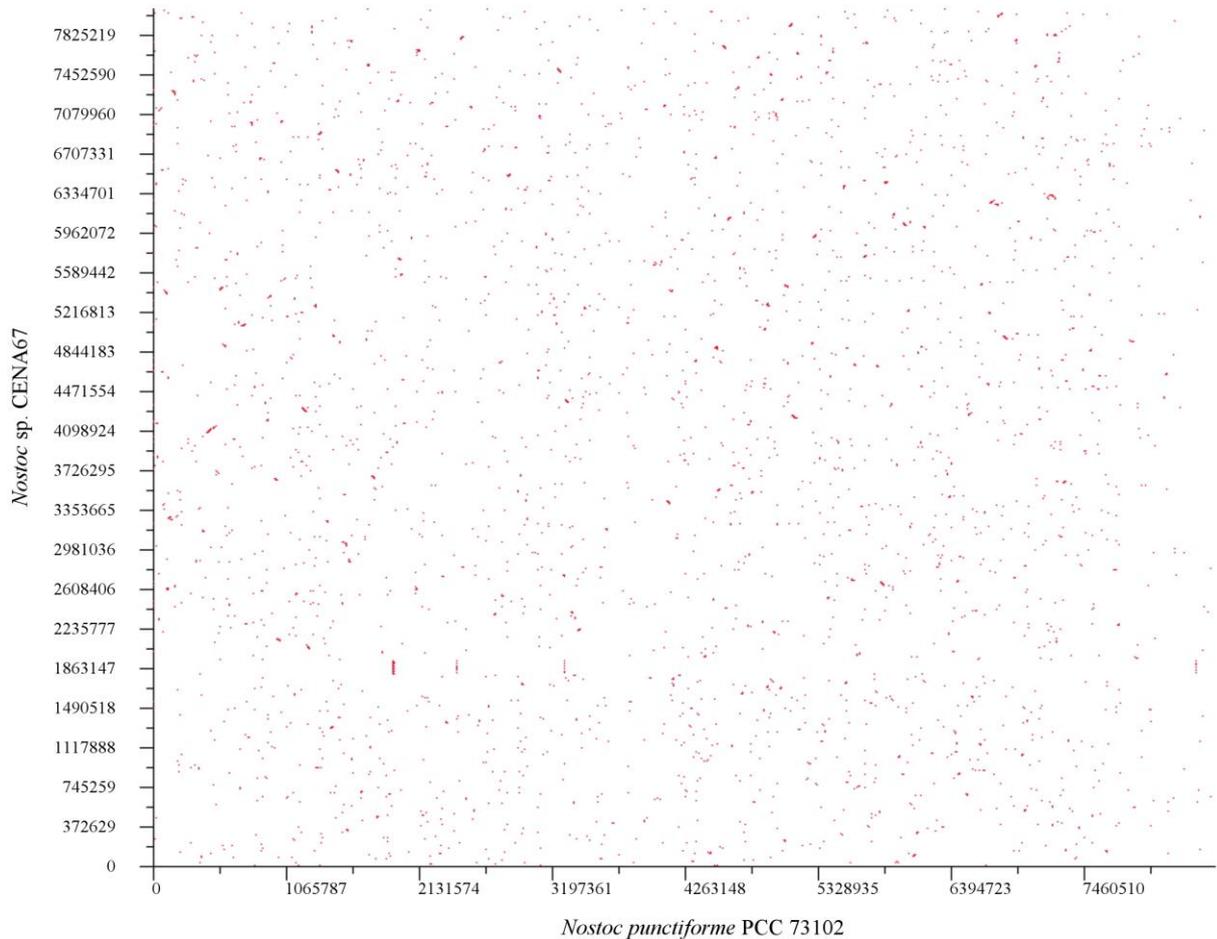


Figura 12 – Sintenia inferida por BLAST entre os genomas de *Nostoc* sp. CENA67 e de seu vizinho genômico mais próximo, *N. punctiforme* PCC 73102

Comparando a região intergênica do DNAr 16S-23S de linhagens próximas ao gênero *Nostoc*, pode ser identificado no genoma de *Nostoc* sp. CENA67 e da maioria das linhagens relacionadas as regiões conservadas D1-D5, genes para RNAs transportadores de isoleucina e alanina e os antiterminadores Box A e Box B (Figura 13). As maiores variações entre linhagens são observadas em sequências da região D1-D1', do antiterminador Box B e da região D5, onde todas apresentam sequências exclusivas. As diferenças mais significativas em regiões conservadas são encontradas em *Nostoc azollae* 0708, que apresenta uma inserção de quatro nucleotídeos em D2, *Desmonostoc* sp. HA7617, que apresenta uma inserção de 30 nucleotídeos na região D5, e *Nostoc muscorum* CENA61, que não possui genes para RNAt em sua sequência.

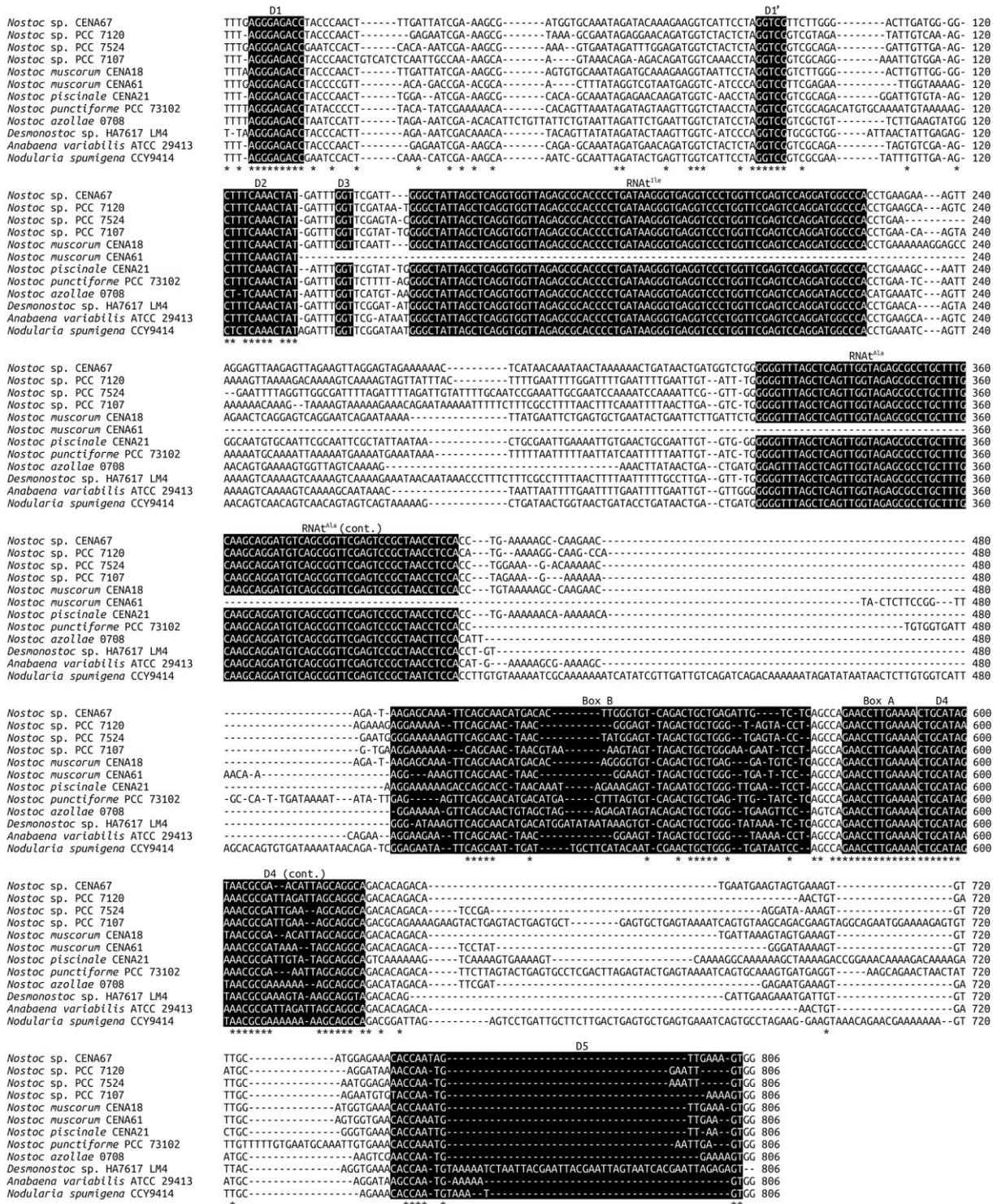


Figura 13 – Alinhamento do espaço intergênico DNAr 16S-23S em *Nostoc* sp. CENA67 e linhagens de táxons próximos, destacando domínios conservados, antiterminadores e genes de RNAt para isoleucina e alanina

As diferenças entre sequências do espaço intergênico DNAr 16S-23S das linhagens comparadas são mais facilmente visualizadas nas estruturas secundárias da hélice D1-D1' (Figura 14) e da hélice Box B (Figura 15). Em ambos os casos, as diferenças se tornam claras no número e tamanho dos círculos formados e nas sequências basais. CENA67 se mostra distinta mesmo de *N. punctiforme* PCC 73102, seu vizinho genômico mais próximo, porém apresenta várias semelhanças com *Desmonostoc* sp. HA7617 LM4. Distinções dessas estruturas têm sido consideradas importantes para a diferenciação de táxons cianobacterianos próximos, como linhagens, espécies ou mesmo gêneros (ITEMAN et al. 2000; JOHANSEN et al., 2011). Novamente nota-se, portanto, a grande variabilidade dentro do gênero *Nostoc* e a necessidade de uma revisão sobre a biologia deste táxon.

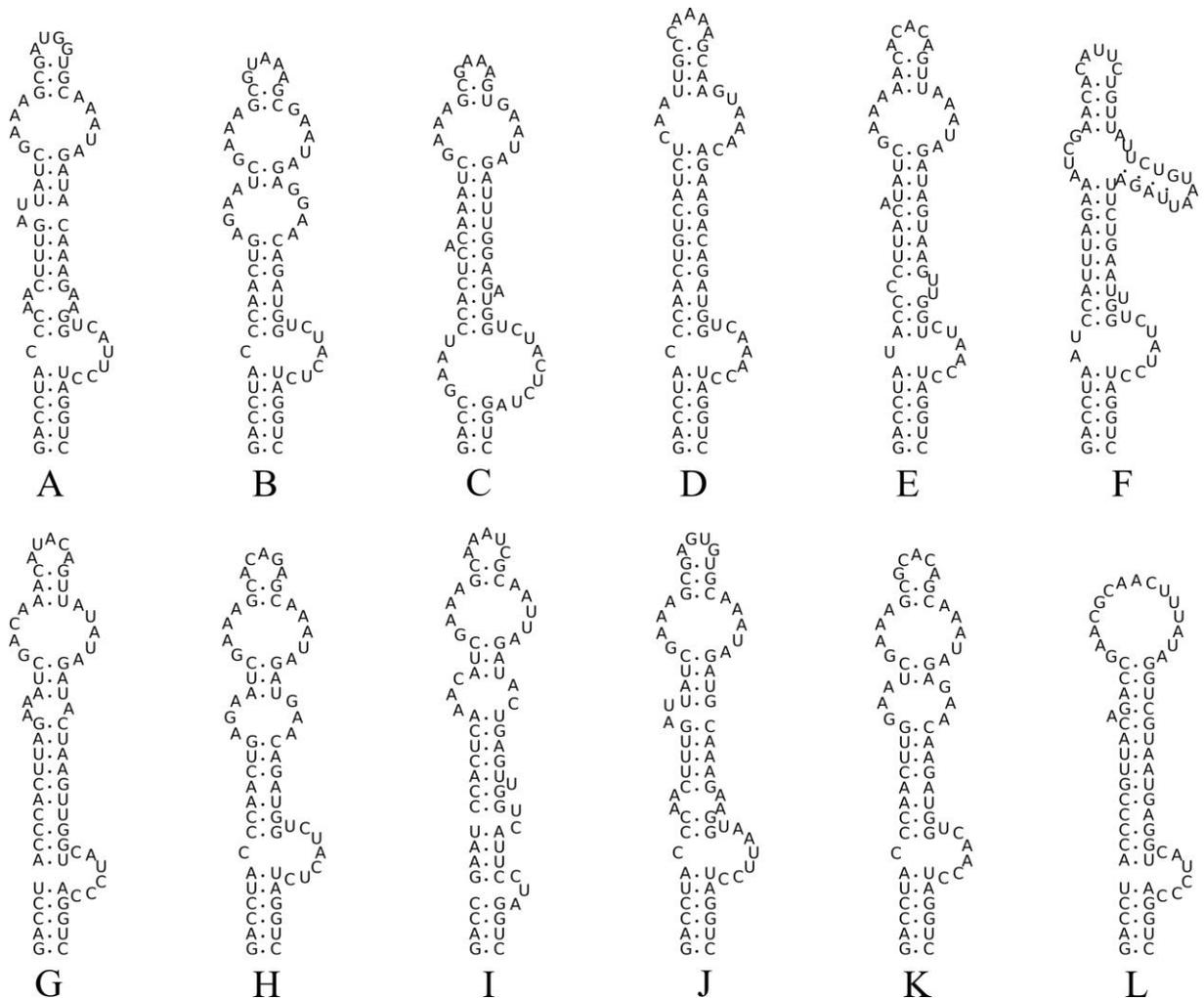


Figura 14 – Estruturas secundárias previstas para a região D1-D1' do espaço intergênico DNAr 16S-23S. A: *Nostoc* sp. CENA67; B: *Nostoc* sp. PCC 7120; C: *Nostoc* sp. PCC 7524; D: *Nostoc* sp. PCC 7107; E: *N. muscorum* CENA18; F: *N. muscorum* CENA61; G: *N. piscinale* CENA21; H: *N. punctiforme* PCC 73102; I: *N. azollae* 0708; J: *Desmonostoc* sp. HA7617 LM4; K: *Anabaena variabilis* ATCC 29413; L: *Nod. spumigena* CCY9414

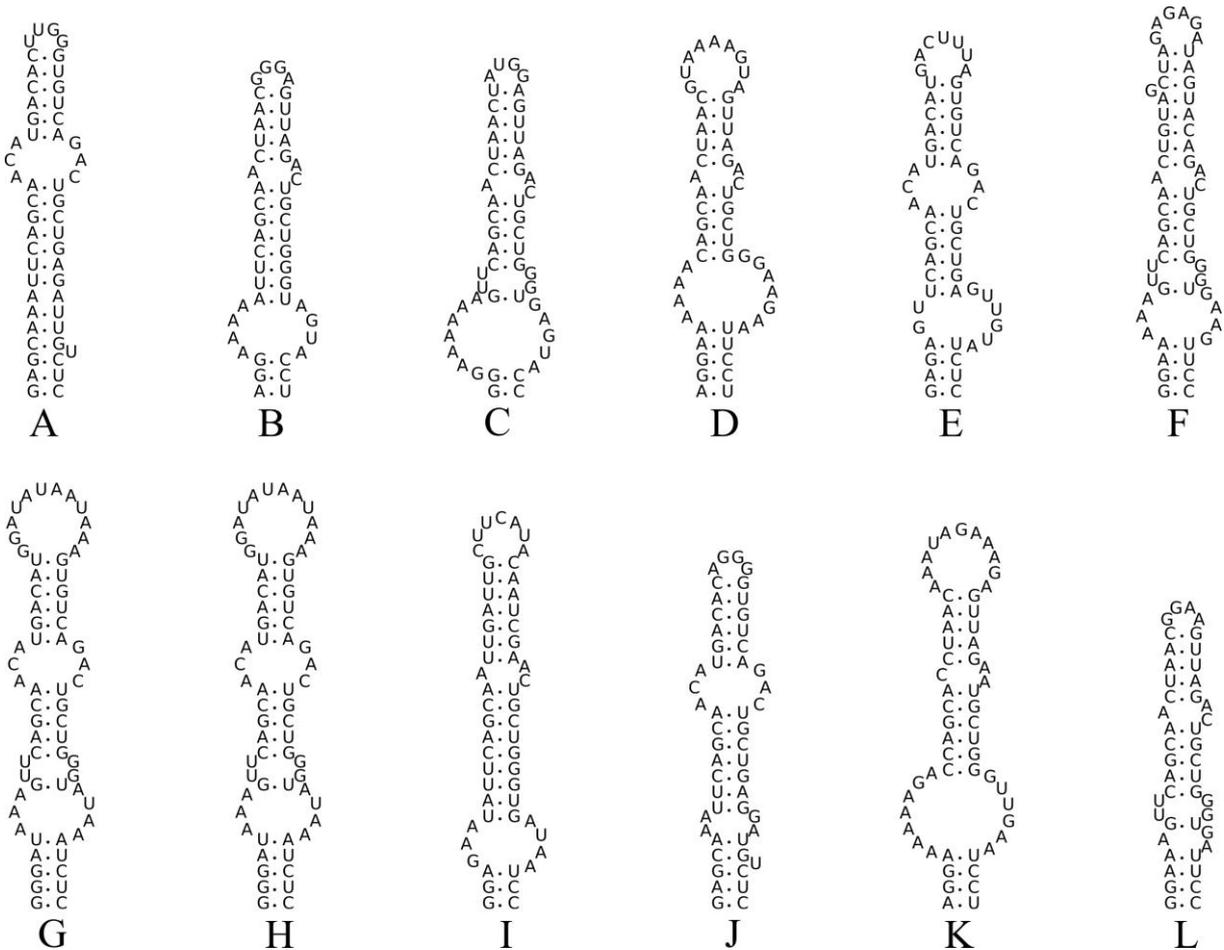


Figura 15 – Predição de estruturas secundárias da região Box B para o espaço intergênico do DNAR 16S-23S. A: *Nostoc* sp. CENA67; B: *Nostoc* sp. PCC 7120; C: *Nostoc* sp. PCC 7524; D: *Nostoc* sp. PCC 7107; E: *N. muscorum* CENA18; F: *N. muscorum* CENA61; G: *N. piscinale* CENA21; H: *N. punctiforme* PCC 73102; I: *N. azollae* 0708; J: *Desmonostoc* sp. HA7617 LM4; K: *Anabaena variabilis* ATCC 29413; L: *Nod. spumigena* CCY9414

Na comparação de proteínas homólogas entre os genomas de *Nostoc* sp. CENA67 e de linhagens já identificadas no gênero *Nostoc* (*Nostoc* sp. PCC 7120, *N. azollae* 0708, *N. punctiforme* PCC 73102, *Nostoc* sp. PCC 7107 e *Nostoc* sp. PCC 7524), verificou-se inicialmente a influência da escolha dentre os algoritmos disponíveis pelo programa GET_HOMOLOGUES sobre os resultados. A análise dos resultados dos algoritmos melhor resultado bidirecional (CONTRERAS-MOREIRA; VINUESA, 2013), COGtriangles (KRISTENSEN et al., 2010) e OrthoMCL (LI; STOECKERT JR; ROOS, 2003) apontou alto consenso dos algoritmos, com a interseção de aproximadamente 90 % dos genes detectados, indicando pouca influência dessa escolha sobre os resultados (Figura 16).

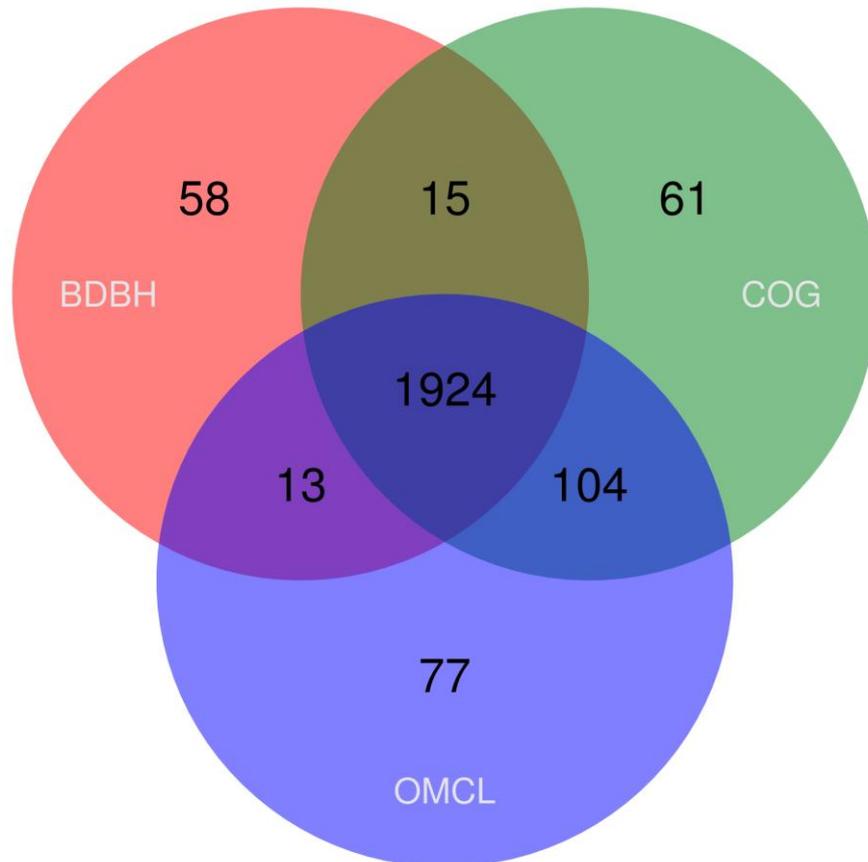


Figura 16 – Comparação de previsões do número de proteínas em comum entre genomas de *Nostoc* spp., indicando o número de proteínas ortólogas encontradas exclusivamente com os algoritmos melhor resultado bidirecional (BDBH), COGtriangles (COG) ou OrthoMCL (OMCL) e simultaneamente entre eles

O algoritmo OrthoMCL acusou o número de ortólogos entre CENA67 e as linhagens PCC 7120, 0708, PCC 73102, PCC 7107 e PCC 7524 como sendo 3.929, 2.565, 4.536, 4.043 e 3.878, respectivamente, e apontou a existência de 1.256 genes parálogos no genoma de *Nostoc* sp. CENA67 (Tabela 7). Esses fatores sugerem tendência ao cosmopolitismo e podem ser parcialmente responsáveis sobre o tamanho genômico relativamente elevado dessa linhagem. Nas análises do pangenoma e do genoma essencial de *Nostoc* spp., baseadas no total de 45.295 genes contidos nos genomas comparados, curvas obtidas com o algoritmo OrthoMCL estimaram o número de genes codificadores de proteínas do genoma essencial em 2.130, enquanto o pangenoma foi estimado por esse algoritmo como sendo composto por 10.590 genes (Figura 17). Contudo, as curvas refletiram a necessidade da disponibilização de maior número de genomas deste grupo.

Tabela 7 – Número de genes parálogos detectados em genomas de linhagens relacionadas ao gênero *Nostoc* e de genes ortólogos que apresentam em comum. O percentual de identidade médio entre linhagens se encontra entre parênteses

Linhagem	Parálogos	1	2	3	4	5	6
1. <i>Nostoc</i> sp. CENA67	1.256	–					
2. <i>Nostoc</i> sp. PCC 7120	569	3.929 (75,6)	–				
3. <i>N. azollae</i> 0708	1.236	2.565 (75,5)	2.561 (75,3)	–			
4. <i>N. punctiforme</i> PCC 73102	1.392	4.536 (74,6)	4.042 (73,7)	2.614 (75,4)	–		
5. <i>Nostoc</i> sp. PCC 7107	649	4.043 (75,8)	4.207 (75,5)	2.599 (75,0)	4.028 (74,0)	–	
6. <i>Nostoc</i> sp. PCC 7524	376	3.878 (78,5)	4.082 (81,4)	2.599 (75,3)	3.832 (76,8)	3.809 (77,9)	–

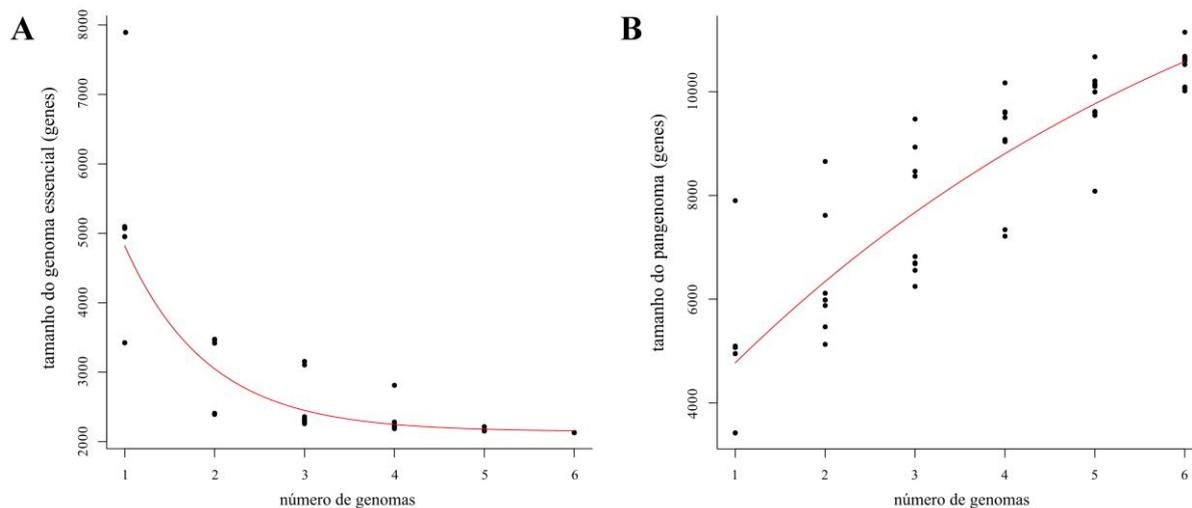


Figura 17 – Curvas de estimativas para os tamanhos do genoma essencial (A) e do pangenoma (B) de cianobactérias relacionadas ao gênero *Nostoc* obtidos com o algoritmo OrthoMCL.

O termo “pangenoma” foi cunhado em 2005 para expressar a diversidade genômica de isolados de *Streptococcus agalactiae* (TETTELIN et al., 2005). A partir desse estudo, observou-se que uma espécie bacteriana poderia ser representada pelo conjunto dos genes que suas linhagens apresentam em comum, presumivelmente essenciais à manutenção de seus processos vitais, e dos genes específicos a uma ou poucas linhagens, dispensáveis à sua

biologia básica e de função acessória, cuja soma constitui o pangenoma da espécie (MEDINI et al., 2005). Desde então, diversas ferramentas têm sido desenvolvidas para o estudo de pangenomas, que permitem estimar quantos genomas adicionais são necessários para caracterizar uma espécie determinada (VERNICOS et al., 2015; XIAO et al., 2015). A análise pangenômica efetuada somente foi possível ao nível de gênero em vista do baixo número de genomas disponíveis para o táxon estudado, porém se mostra válida na exposição do que o genoma obtido traz de novo ao que já está estabelecido, evidenciando a diversidade dos organismos desse táxon e a necessidade de novos sequenciamentos.

N. punctiforme PCC 73102 é uma linhagem de metabolismo facultativamente heterotrófico que tem competência para a simbiose com fungos e plantas terrestres. O cromossomo dessa linhagem é relativamente longo, se aproximando de 8,23 Mpb. Dentre as 7.432 sequências codificantes que apresenta, 45 % codificam proteínas de função conhecida e 29 % parecem ser exclusivos dessa espécie. Além de numerosas sequências de inserção e repetições, são encontrados genes para transposases e enzimas de modificação de DNA, o que aponta alta plasticidade neste genoma. Essa linhagem parece apresentar grande potencial de percepção e resposta a sinais ambientais, possuindo mais de 400 genes envolvidos em processos de transdução de sinal, entre eles proteínas cinases, reguladores de resposta e fatores de transcrição, que possivelmente contribuem também para sua diferenciação celular e capacidade simbiótica (EKMAN et al., 2013; MEEKS et al., 2001).

Como verificado, apesar de *N. punctiforme* PCC 73102 ter sido apontada como próxima da linhagem estudada, elas apresentam diferenças moleculares e ecológicas significativas. As observações indicam que, a despeito de o genoma de algumas linhagens identificadas no gênero *Nostoc* terem sido caracterizados, há ainda desconhecimento considerável a respeito deste grupo de cianobactérias. A diversidade taxonômica tem sido um importante fator na escolha de cianobactérias para sequenciamento genômico (SHIH et al., 2013), porém ainda há grande carência de variedade de táxons e de habitats em bancos de dados de sequências deste filo.

A maioria dos genomas cianobacterianos disponíveis até o momento pertence a micro-organismos aquáticos, porém há diferenças entre a ecologia química de cianobactérias de ambientes terrestres, marinhos e de água doce que tornam divergências entre seu tratamento inevitáveis (LEÃO et al., 2012a). Genomas de cianobactérias terrestres, como *Nostoc* sp. CENA67, permanecem como um campo pouco explorado. Conforme observado nas análises realizadas, o genoma dessa linhagem exhibe algumas características inéditas até o momento, diferentes mesmo dos genomas tidos como mais próximos dentre os atualmente

disponíveis, portanto contribui para o aumento do conhecimento sobre a diversidade genética e bioquímica do filo Cyanobacteria. Os resultados encontrados reforçam a hipótese de que CENA67 constitua um táxon distinto, porém é necessário que mais linhagens próximas sejam isoladas e caracterizadas para que seja possível realizar seu posicionamento taxonômico com a segurança necessária.

A montagem de sequências provenientes de amostras com múltiplos micro-organismos é favorecida em amostras com baixa diversidade e composta por espécies bastante diferentes (DE FILIPPO et al., 2012), como provavelmente encontrado na cultura estudada. Montagens desse tipo geralmente podem ser aprimoradas pela separação prévia das leituras em grupos representativos de um mesmo genoma ou de genomas próximos. Essa separação pode ser tanto baseada em diferenças na composição das sequências, como a conservação no conteúdo de nucleotídeos, quanto na homologia e na similaridade das mesmas com sequências depositadas em bancos de dados. Todavia, leituras curtas geralmente não fornecem informação o suficiente para que sejam separadas apenas por sua composição de nucleotídeos e frequentemente necessitam também de sua comparação com alguma referência. Caso uma sequência proximamente relacionada não esteja disponível, mesmo referências relacionadas apenas ao nível de filo podem ser utilizadas para a separação de leituras (THOMAS; GILBERT; MEYER, 2012).

No presente trabalho, a separação de sequências de acordo com sua identificação taxonômica ocorreu após os procedimentos de montagem. Essa abordagem se fez necessária em decorrência do ineditismo do genoma da cianobactéria estudada, que possivelmente constitui um táxon novo e possui, portanto, um número considerável de segmentos exclusivos, não encontrados nos genomas de referência disponíveis. Conseqüentemente, a separação de leituras em uma etapa anterior à montagem levaria à perda de uma quantidade substancial de sequências. O método utilizado se mostrou adequado para o estudo genômico de cianobactérias provenientes de culturas não axênicas, e contorna as grandes dificuldades encontradas nas etapas de remoção de outros micro-organismos da cultura, as quais, caso sejam possíveis e bem sucedidas, podem acrescentar anos ao tempo necessário para o estudo genômico de uma linhagem.

6.5 Diversidade taxonômica de micro-organismos associados

Como esperado, as sequências mais abundantes na comunidade foram identificadas como Cyanobacteria (43 %), mas sua abundância foi acompanhada proximamente por Alphaproteobacteria (33 %) (Figura 18A). Os mais abundantes gêneros não cianobacterianos identificados foram *Mesorhizobium*, *Sinorhizobium* e *Starkeya*, seguidos por outros gêneros que, em sua maioria, também são tipicamente encontrados em solo (Figura 18B).

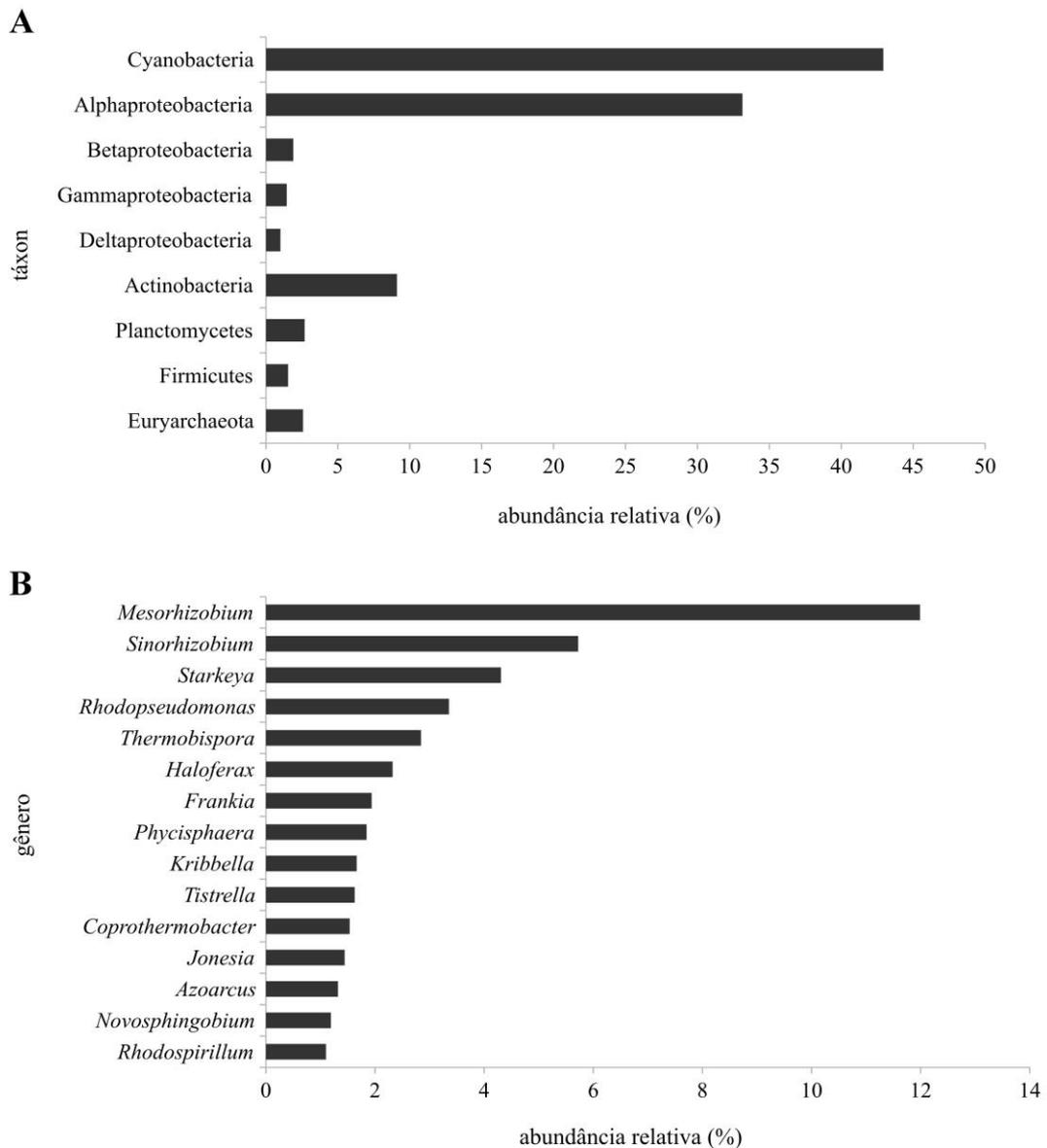


Figura 18 – Abundâncias relativas dos táxons de maior predominância na cultura analisada. A: Abundância de sequências identificadas em categorias taxonômicas de nível mais elevado. B: Gêneros não cianobacterianos mais abundantes

Mesorhizobium, *Sinorhizobium*, *Starkeya*, *Rhodopseudomonas*, *Frankia*, *Azoarcus*, *Novosphingobium* e *Rhodospirillum* são gêneros bacterianos que apresentam muitos membros capazes de fixar nitrogênio. Provavelmente como consequência dessa habilidade, várias espécies de *Mesorhizobium*, *Sinorhizobium*, *Starkeya*, *Frankia* e *Azoarcus* são capazes de estabelecer relações simbióticas mutualistas com plantas (MOULIN et al., 2001; PAWLOWSKI; SIRRENBORG, 2003; KRAUSE et al., 2006; ZAKHIA et al., 2006; MASSON-BOIVIN et al., 2009). Além disso, *Phycisphaera* pode estabelecer associações com algumas espécies de clorófitas (LAGE; BONDOSO, 2014). *Rhodospirillum* e *Rhodopseudomonas* também são capazes de realizar fotossíntese (GIESBERGER, 1947; LARIMER et al., 2004). Porém, como é típico em alfa-proteobactérias, ela é inibida por oxigênio (IMHOFF, 2006). Por outro lado, apesar de RubisCo ativa ter sido descrita em *Haloferax*, este gênero aparentemente é incapaz de crescimento autotrófico (HÜGLER et al., 2003). A fisiologia de *Starkeya* também é interessante, já que pode tanto produzir quanto consumir dióxido de carbono (KAPPLER et al., 2012).

A maioria dos gêneros identificados é relatada como membro usual de comunidades microbianas em habitats de solo. Não obstante, a composição taxonômica da comunidade microbiana associada à linhagem cianobacteriana estudada é incomum devido a um número considerável desses micro-organismos comumente interagir com plantas, principalmente na rizosfera. Uma exceção marcante a esse padrão é a presença do gênero *Coprothermobacter*, cujos membros são normalmente encontrados em ambientes anaeróbicos e de alta temperatura (RAINEY; STACKEBRANDT, 1993; TANDISHABO et al., 2012). Aparentemente, alta abundância de *Mesorhizobium* associado a uma linhagem cianobacteriana capaz de fixação de nitrogênio não é incomum, já que também foi relatada por Brauer et al. (2015).

O índice de alfa-diversidade calculado pelo servidor MG-RAST para este metagenoma indicou 191,954 espécies, o que não é surpreendente, tendo em vista que solos da Amazônia são potenciais reservatórios de diversidade microbiana intocada (BRUCE et al., 2012). Além disso, a curva de rarefação calculada para este metagenoma está próxima de atingir a assíntota (Figura 19), indicando que poucas espécies permanecem por ser descobertas nessa comunidade. Logo, os esforços de amostragem e sequenciamento foram considerados adequados para os propósitos deste trabalho.

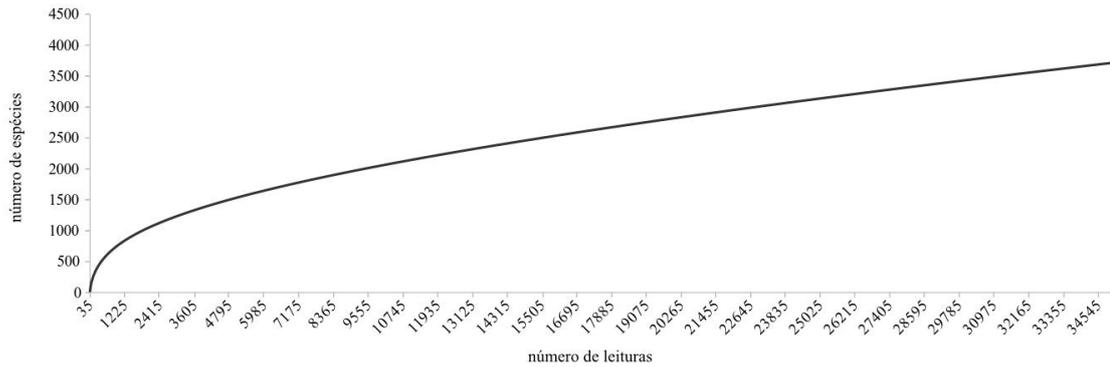


Figura 19 – Curva de rarefação calculada para o metagenoma montado

A seleção de unidades taxonômicas operacionais (UTOs) por referências fechadas a 94 % a partir de sequências de DNAr 16S detectou 364 UTOs não cianobacterianas representativas identificadas nos filos Acidobacteria, Actinobacteria, Bacteroidetes, Chloroflexi, Firmicutes, Planctomycetes e Proteobacteria. Algumas dessas UTOs foram designadas a identidades desconhecidas em níveis taxonômicos variados (Tabela 8). Além disso, EMIRGE montou 11 sequências parciais de DNAr 16S, cuja maioria apresentou identidade significativa com bactérias conhecidas (Tabela 9).

Tabela 8 – Identificação taxonômica de UTOs não cianobacterianas em fragmentos de DNAr 16S detectados em sequências não montadas

Filo	Classe	Ordem	Família	Gênero	UTOs
–	–	–	–	–	3
Acidobacteria	Acidobacteriia	Acidobacteriales	Acidobacteriaceae	–	1
Acidobacteria	Solibacteres	Solibacterales	Desconhecida	–	3
Actinobacteria	Actinobacteria	Actinomycetales	–	–	1
Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	<i>Microbacterium</i>	4
Actinobacteria	Actinobacteria	Actinomycetales	Mycobacteriaceae	<i>Mycobacterium</i>	11
Actinobacteria	Actinobacteria	Actinomycetales	Propionibacteriaceae	<i>Propionibacterium</i>	1
Bacteroidetes	Bacteroidia	Bacteroidales	–	–	1
Chloroflexi	Ktedonobacteria	–	–	–	2
Firmicutes	Bacilli	Bacillales	Bacillaceae	<i>Bacillus</i>	3
Planctomycetes	Phycisphaerae	Phycisphaerales	Desconhecida	–	1
Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	–	2
Planctomycetes	Planctomycetia	Gemmatales	Gemmataceae	<i>Gemmata</i>	13
Planctomycetes	Planctomycetia	Gemmatales	Isosphaeraceae	–	1
Proteobacteria	Alphaproteobacteria	–	–	–	3

(continua)

Tabela 8 (conclusão)

Filo	Classe	Ordem	Família	Gênero	UTOs
Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	–	4
Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Desconhecido	8
Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	<i>Phenylobacterium</i>	2
Proteobacteria	Alphaproteobacteria	Rhizobiales	–	–	24
Proteobacteria	Alphaproteobacteria	Rhizobiales	Aurantimonadaceae	–	1
Proteobacteria	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	–	5
Proteobacteria	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	<i>Afipia</i>	5
Proteobacteria	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	<i>Bosea</i>	1
Proteobacteria	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	<i>Bradyrhizobium</i>	3
Proteobacteria	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	Desconhecido	13
Proteobacteria	Alphaproteobacteria	Rhizobiales	Desconhecida	–	44
Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	–	5
Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	<i>Devosia</i>	6
Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	<i>Hyphomicrobium</i>	17
Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	<i>Rhodoplanes</i>	1
Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylocystaceae	–	1
Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	–	6
Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	Desconhecido	4
Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	<i>Mesorhizobium</i>	13
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	–	39
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	<i>Agrobacterium</i>	1
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	–	28
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	<i>Sinorhizobium</i>	5
Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	–	5
Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	–	1
Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	–	9
Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	–	7
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	–	7
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	<i>Blastomonas</i>	1
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Desconhecido	9
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	<i>Novosphingobium</i>	11
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	<i>Sphingobium</i>	1
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	<i>Sphingomonas</i>	24
Proteobacteria	Gammaproteobacteria	–	–	–	1
Proteobacteria	Gammaproteobacteria	Chromatiales	–	–	1
Proteobacteria	Gammaproteobacteria	Xanthomonadales	–	–	1
Proteobacteria	–	–	–	–	3

Os resultados obtidos por análises de DNAr diferem significativamente dos resultados taxonômicos baseados em análises de todo o metagenoma. Contudo, já que as sequências de DNAr estão representadas aleatoriamente no conjunto de dados, elas não podem ser tomadas como sendo representativas da comunidade total, diferentemente das análises do metagenoma total. Apesar disso, essa abordagem é adotada por algumas metodologias de análise, como a implementada pelo servidor EBI Metagenomics (HUNTER et al., 2013), a qual detecta DNAr 16S em dados metagenômicos e infere a composição taxonômica da comunidade a partir dessas sequências. Logo, resultados obtidos com base nesse método devem ser abordados com cautela.

Tabela 9 – Similaridade de sequências de DNAr 16S montadas pelo programa EMIRGE em relação a organismos conhecidos depositados no banco de dados GenBank obtida com a ferramenta BLAST

Sequência	Comprimento (pb)	Sequência Mais Próxima	Cobertura	Identidade
2	1283	<i>Ensifer</i> sp. ZNC0028 (KM585589)	100 %	100 %
4	1392	<i>Gemmata</i> sp. Ha1-2 (GQ889431)	96 %	93 %
6	1528	<i>Burkholderia cenocepacia</i> HI2424 (CP000459)	100 %	100 %
10	1362	<i>Paracraurococcus</i> sp. ORS 1473 (AJ968702)	100 %	99 %
12	1441	<i>Afipia</i> sp. 1 MRT-114 (EF371496)	100 %	99 %
22	1251	<i>Bosea lupini</i> R-45681 (NR_108514)	100 %	99 %
35	1430	<i>Nostoc ellipsosporum</i> V (AJ630450)	97 %	98 %
42	1360	<i>Bryobacter aggregatus</i> MOB76 (AM887762)	98 %	97 %
52	1284	<i>Caulobacter leidyi</i> (AF331660)	100 %	100 %
53	1296	<i>Nordella oligomobilis</i> N21 (NR_114615)	96 %	99 %
74	1113	<i>Hyphomicrobium aestuarii</i> ATCC 27483 (NR_104954)	100 %	99 %

Bactérias de novos táxons podem ser encontradas em associação com cianobactérias, como exemplificado por *Hoeflea anabaenae*, encontrada em heterócitos de *Anabaena*, e *Gracilimonas tropica*, proveniente de uma cultura de *Synechococcus* (CHOI et al., 2009; STEVENSON et al., 2011). Uma das sequências de DNAr 16S obtidas apresentou 93 % de identidade com uma sequência de *Gemmata*, um indicativo de que possivelmente é

representante de um novo gênero da família Planctomycetaceae. Alguns planctomicetes podem ser encontrados em forte associação com cianobactérias, e florações de cianobactérias frequentemente estimulam o aumento do número de planctomicetes (WARD et al., 2006). A análise da comunidade microbiana associada a *Microcystis* spp. apontou que, ao passo que proteobactérias dominavam as comunidades presentes em sistemas de enriquecimento, planctomicetes predominavam entre os micro-organismos mais fortemente aderidos às células das cianobactérias analisadas (CAI et al., 2013).

Como as análises de DNAr apresentaram poucos gêneros em comum com os gêneros apontados como mais abundantes na análise de metagenoma total, elas provavelmente revelaram táxons de maior raridade nas amostras, que possivelmente fazem parte da chamada “biosfera rara”, ou o conjunto de micro-organismos que permanecem sob baixa abundância em determinados ambientes (SOGIN et al., 2006). Esse método levantou um número considerável de UTOs em táxons desconhecidos, e algumas UTOs também foram incluídas em categorias taxonômicas que conhecidamente possuem um grande número de bactérias que atualmente não são cultivadas, como Ktedonobacteria (YOKOTA, 2012).

Atualmente, a maior parte das bactérias não é passível de desenvolvimento sob condições de cultura (RAPPE; GIOVANNONI, 2003). Dentre essas bactérias, micro-organismos obrigatoriamente simbiotes dependem de nutrientes ou de interações celulares providas por outros micro-organismos para o seu desenvolvimento e apresentam, portanto, baixas possibilidades de ser cultivados (WILSON; PIEL, 2013). Apesar de cocultivo já ter sido proposto como uma opção alternativa para o desenvolvimento de estudos *in vitro* de bactérias previamente não cultivadas (VARTOUKIAN; PALMER; WADE, 2010; STEWART, 2012), baixa ênfase foi dada a cianobactérias como parceiras de cultivo. Não obstante, bactérias fotoautotróficas e fixadoras de nitrogênio do filo Cyanobacteria estão em uma ótima posição para desempenhar essa função, como consequência de sua extraordinária capacidade metabólica.

Alta diversidade bacteriana acompanha florações cianobacterianas, e pode incluir táxons ainda não descritos, incluindo possíveis novas espécies, gêneros, famílias ou mesmo ordens (BERG et al., 2009; TUOMAINEN et al., 2006). Macrocolônias de algumas espécies do gênero *Nostoc* também constituem ambientes interessantes para a descoberta de novos táxons microbianos, já que fornecem condições únicas, não encontradas em outros ambientes (INTHASOTTI; PATHOM-AREE, 2015). Conforme se verifica nas análises de DNAr 16S, cianobactérias relacionadas ao gênero *Nostoc* que não formam macrocolônias também estabelecem nichos viáveis para o desenvolvimento de micro-organismos ainda

desconhecidos, e, em consórcio, podem prover as condições necessárias para que alguns micro-organismos previamente não cultivados se desenvolvam sob cocultivo, mesmo sob baixa abundância.

Dentre as sequências encontradas no metagenoma montado e que, a exemplo do genoma de *Nostoc* sp. CENA67, puderam ser reconhecidas e somadas até constituírem rascunhos genômicos, foram encontradas sequências identificadas como pertencentes às alfa-proteobactérias *Bradyrhizobium diazoefficiens*, *Bradyrhizobium japonicum* e *Hyphomicrobium nitratorans* e à betaproteobactéria *Burkholderia lata* (Tabela 10). Apesar de ser possível encontrar e identificar sequências de outras bactérias conhecidas no metagenoma, elas se apresentam bastante fragmentadas, ou sua soma não alcança tamanhos próximos aos de genomas de referência. Além disso, há a possibilidade de que outros rascunhos genômicos presentes no metagenoma montado pertençam a táxons novos e, dessa forma, não tenham sido identificados em espécies conhecidas, o que dificulta sua recuperação. De fato, aproximadamente 89 % das sequências foram classificadas como desconhecidas pelo programa Kraken. Contudo, essa pode ser uma limitação do método utilizado.

Tabela 10 – Características gerais dos rascunhos genômicos de algumas bactérias associadas à cianobactéria *Nostoc* sp. CENA67

Organismo	Sequências	Tamanho (Mpb)	Percentual GC	Cobertura Média	Genes	RNA
<i>Bradyrhizobium diazoefficiens</i>	266	9,28	64,21	7 ×	9.101	49
<i>Bradyrhizobium japonicum</i>	289	8,23	63,72	6 ×	8.026	69
<i>Burkholderia lata</i>	139	7,53	66,65	8 ×	6.860	56
<i>Hyphomicrobium nitratorans</i>	23	4,19	62,28	14 ×	3.833	52

Não obstante o baixo número de trabalhos que trataram da diversidade de micro-organismos associados a cianobactérias, a maioria dos quais se restringiu a ambientes aquáticos, alguns táxons relacionados aos genomas obtidos já foram relatados em simbiose com organismos do filo Cyanobacteria. *Burkholderia* e *Bradyrhizobium* foram encontradas anteriormente em associação com *Microcystis aeruginosa* (BERG et al., 2009; LI et al., 2011; SHI et al., 2011). Bactérias da família Burkholderiaceae também foram observadas em associação com *Anabaena*, *Aphanizomenon*, *Microcystis aeruginosa*, *Cylindrospermopsis raciborskii* e *Nostoc commune* (EILER; BERTILSSON, 2004; GRAHAM et al., 2014; SHI et al., 2009), e uma espécie do gênero *Hyphomicrobium* também esteve presente em culturas de enriquecimento de *Microcystis* spp. (CAI et al., 2013).

Todavia, com exceção dos genomas de *Blastomonas* sp. e *Rhodobacter* sp., provenientes de uma cultura não axênica de *Cyanobium* sp. (LIMA et al., 2014a; 2014b; 2014c), há poucas informações sobre a genômica de bactérias heterotróficas associadas a cianobactérias em culturas não axênicas.

Abordagens metagenômicas geralmente são baseadas na identificação e na separação de sequências prévia ou posteriormente à montagem do metagenoma, com base em bancos de dados de referência e/ou em diferenças de composição entre sequências (DRÖGUE; MCHARDY, 2012; MANDE; MOHAMMED; GHOSH, 2012). Alternativamente, uma estratégia de montagem e separação integradas, baseada em perfis diferenciais de cobertura genômica, está apresentando resultados promissores (ALBERTSEN et al., 2013; NIELSEN et al., 2014). Alguns programas que implementam esse método, tais como CONCOCT (ALNEBERG et al., 2014), GroopM (IMELFORT et al., 2014) e MetaBAT (KANG et al., 2014), têm sido disponibilizados.

Devido à pouca diferença de cobertura entre alguns genomas amostrados e às relações próximas entre os organismos simbióticos da cultura estudada, as abordagens que independem de bancos de dados não se mostram adequadas aos dados analisados, já que podem ter dificuldade em discriminar sequências e, conseqüentemente, gerar resultados híbridos e de baixa confiabilidade. Dessa forma, apesar de possivelmente haver táxons desconhecidos na amostra sequenciada, a comparação das sequências com referências já conhecidas se mostra como a alternativa mais confiável para a situação encontrada. À medida que os bancos de dados públicos são aprimorados, novas respostas podem ser disponibilizadas.

Há relatos de especificidade de comunidades microbianas a algumas espécies cianobacterianas (BAGATINI et al., 2014; SHI et al., 2009). Algumas cianobactérias aparentam de fato possuir certa influência sobre a composição taxonômica de suas comunidades associadas, exercitando controle através de mecanismos de competição e de facilitação, como a liberação de carbono e nitrogênio fixados (BRAUER et al., 2015). Além disso, metabólitos secundários de efeito alelopático, incluindo inibidores de protease e bacteriocinas, podem ser liberados por diversas cianobactérias e essas moléculas poderiam, portanto, impactar a comunidade e atuar como agentes de seleção de micro-organismos (LEÃO et al., 2012b; SILVA-STENICO et al., 2011; 2013a; 2013b; WANG; FEWER; SIVONEN, 2011). Logo, os organismos identificados provavelmente são o produto de uma série de interações entre a linhagem cianobacteriana e os micro-organismos que, ao longo do procedimento de cultura, remanesceram, enquanto outros eram eliminados por encontrar condições pouco favoráveis, exacerbadas pela competição.

6.6 Caracterização funcional da comunidade associada

As anotações dos rascunhos genômicos montados revelaram um grande número de sequências não incluídas em subsistemas no genoma de *Nostoc* sp. (73 %), um número maior que o encontrado nos outros genomas (Figura 20). Genes para a biossíntese de auxinas foram encontrados nos cinco rascunhos genômicos montados. A presença de vias biossintéticas para este fitormônio não é surpreendente, tendo em vista que é provavelmente comum que essas bactérias interajam com plantas em seu ambiente natural.

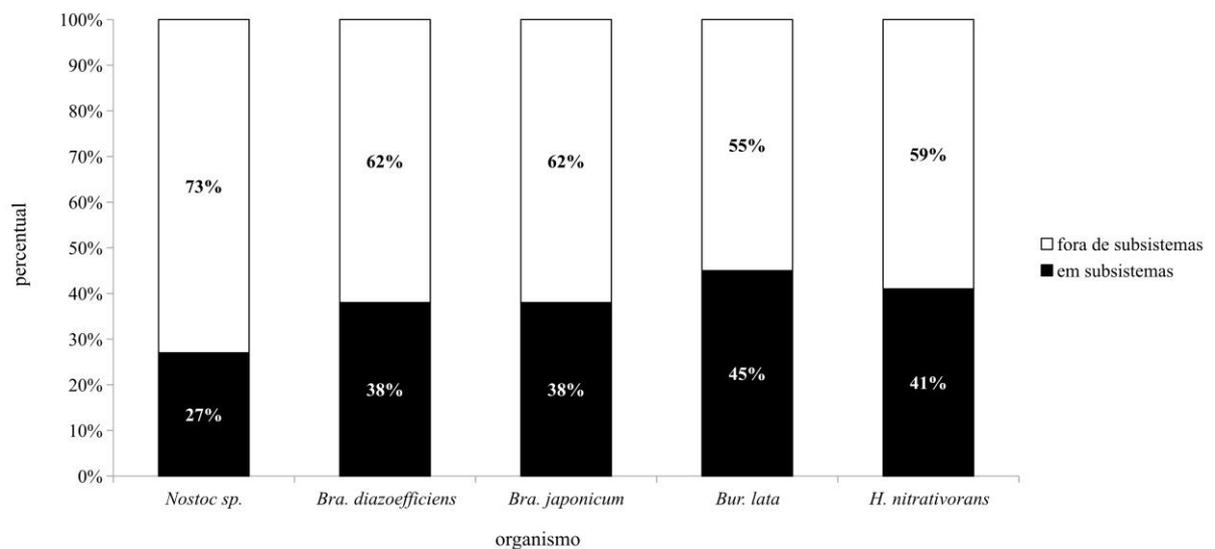


Figura 20 – Cobertura de subsistemas dos genomas obtidos a partir do metagenoma

A anotação apontou 429, 381, 550 e 413 subsistemas nos rascunhos genômicos de *Bra. diazoefficiens*, *Bra. japonicum*, *Bur. lata* e *H. nitratorans*, respectivamente (Figura 21). Como esperado, genes para fotossíntese foram encontrados apenas entre as sequências cianobacterianas dentre os cinco genomas obtidos. Todavia, genes para fixação de nitrogênio também foram encontrados apenas entre sequências cianobacterianas, apesar de os outros genomas pertencerem a gêneros que normalmente são capazes de fixar nitrogênio atmosférico (Tabela 11).

Tabela 11 – Subsistemas relacionados ao metabolismo de nitrogênio detectados nos rascunhos genômicos montados

Subsistema	<i>Nostoc sp.</i>	<i>Bra. diazoefficiens</i>	<i>Bra. japonicum</i>	<i>Bur. lata</i>	<i>H. nitratorans</i>
Fixação de nitrogênio	19	–	–	–	–
Hidrólise de cianato	8	–	–	–	–
Amonificação de nitrato/nitrito	9	25	14	14	14
Assimilação de amônia	12	18	–	17	29
Utilização de alantoína	–	–	9	7	–
Estresse nitrosativo	–	3	1	1	5
Nitrilase	–	–	1	2	–
Amidase / ureia e nitrila hidratase	–	2	1	–	–
Redutase desnitrificante	–	5	–	–	12
Desnitrificação	–	11	–	–	24

Apesar de ser possível detectar genes a partir de sequências curtas, a montagem de metagenomas fornece maior número de informações aos programas de predição funcional (THOMAS; GILBERT; MEYER, 2012), assegurando maior confiança em seus resultados. Logo, as análises funcionais da comunidade microbiana associada à cianobactéria *Nostoc sp. CENA67* foram realizadas sobre as sequências metagenômicas após a montagem. A anotação do metagenoma da comunidade associada revelou uma ampla gama de subsistemas (Figura 22). A comparação dos subsistemas do genoma cianobacteriano às anotações do metagenoma de sua comunidade associada apontou nove subsistemas exclusivos (Tabela 12). O procedimento oposto – a busca por subsistemas exclusivos à comunidade associada – revela 285 subsistemas, porém nenhum aparenta ser fundamental para cianobactérias.

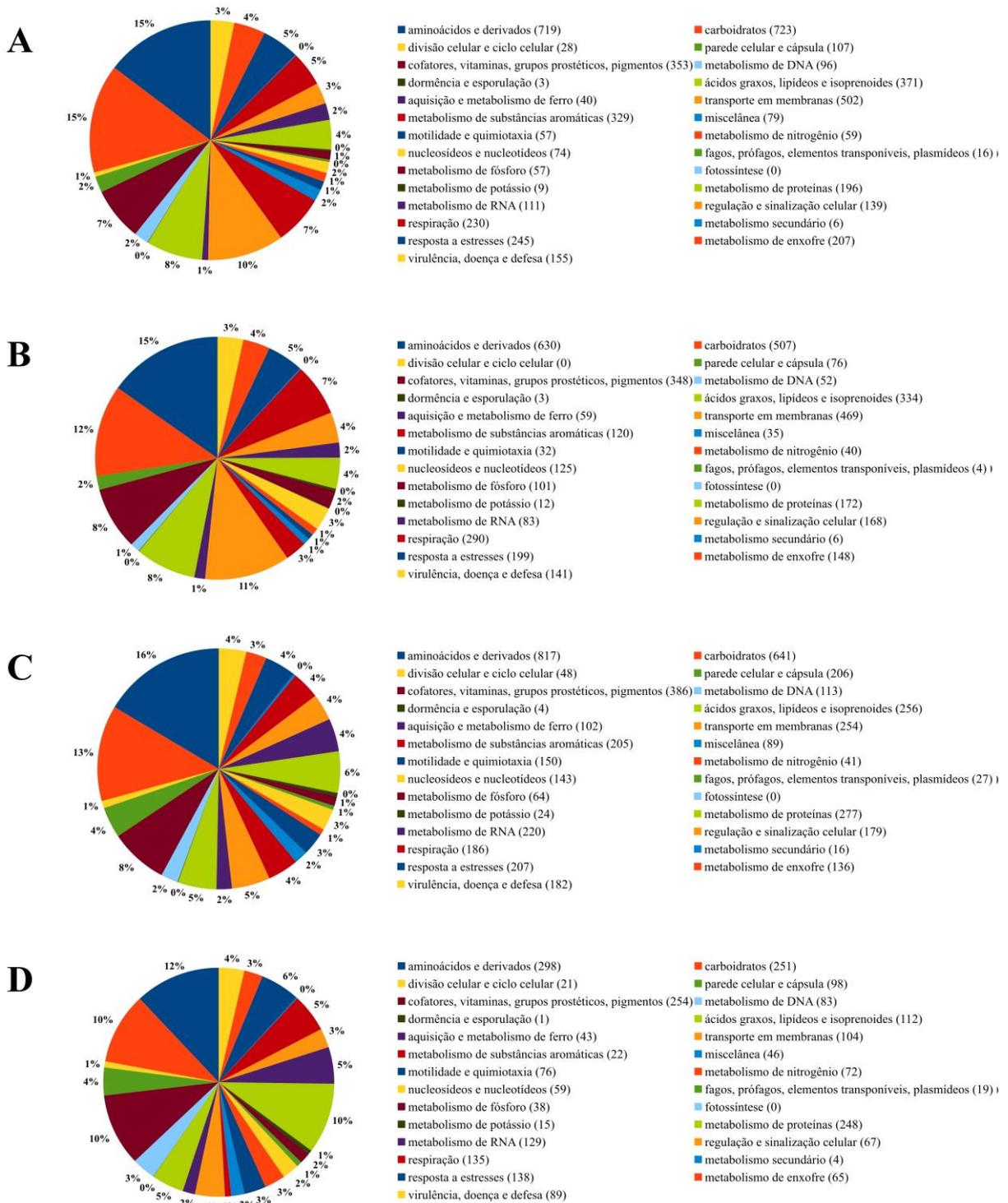


Figura 21 – Distribuição de subsistemas nos rascunhos genômicos de bactérias heterotróficas obtidos neste trabalho. A: *Bra. diaoefficiens*; B: *Bra. japonicum*; C: *Bur. lata*; D: *H. nitrativorans*

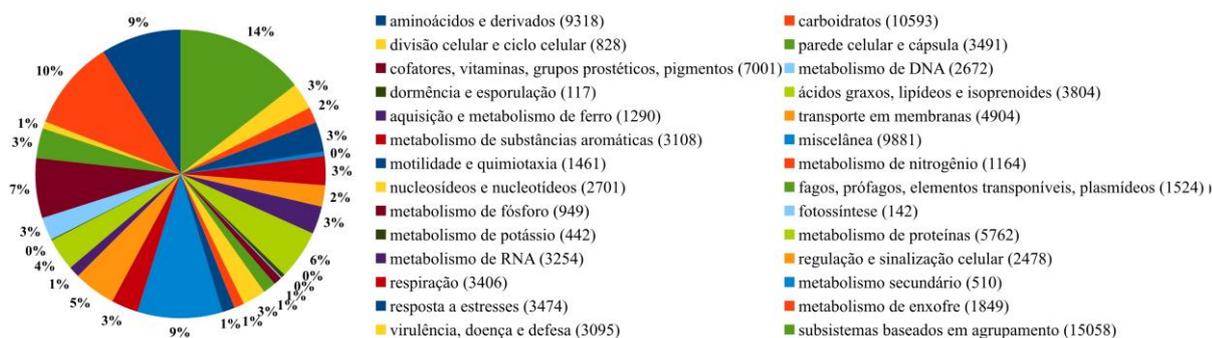


Figura 22 – Distribuição de subsistemas da comunidade associada à cianobactéria CENA67

Aparentemente, os subsistemas exclusivos ao genoma cianobacteriano não favorecem qualquer tipo de interação entre estas células e sua comunidade associada, já que a maioria deles está relacionada direta ou indiretamente à fisiologia cianobacteriana comum. Vias exclusivas também foram observadas no mapeamento metabólico (Figura 23).

Contudo, é possível que moléculas desconhecidas desempenhem algum papel nessa interação, tendo em vista que anotações do servidor MG-RAST apontaram 46,8 % do metagenoma da comunidade associada como proteínas desconhecidas. Além disso, entre 24,2 % (de acordo com comparações ao banco de dados COG) e 84,6 % (com base no banco de dados NOG) das sequências classificadas em categorias funcionais foram incluídas em processos pouco caracterizados.

Interações entre cianobactérias e heterótrofos em consórcios, apesar de específicas a cada produtor primário, geralmente têm como padrão a troca mutualista de oxigênio e carbono, liberados pelas células cianobacterianas, por metabólitos produzidos por células heterotróficas, como nutrientes e vitaminas (COLE et al., 2014). Algumas análises transcritômicas fornecem suporte a essa observação, e também indicam que uma situação semelhante pode ser observada em relação ao nitrogênio fixado por cianobactérias diazotróficas em cocultivo com bactérias heterotróficas (BELIAEV et al., 2014). Além do trabalho presente, a associação entre cianobactérias diazotróficas e outras bactérias também potencialmente diazotróficas, bem como associações entre cianobactérias e bactérias fotossintéticas anoxigênicas ou bactérias mais comumente encontradas em raízes de plantas, foi relatada anteriormente (GRAHAM et al., 2014; SÁNCHEZ et al., 2005). A fixação de nitrogênio pode estimular a simbiose mesmo entre cianobactérias de espécies diazotróficas (MOMPER et al., 2015).

Tabela 12 – Subsistemas exclusivos ao genoma de *Nostoc* sp. CENA67 quando comparado com o metagenoma de sua comunidade associada

Subsistema	Categoria	Ocorrências
Formação de heterócitos em cianobactérias	Divisão celular e ciclo celular	8
Degradação de clorofila	Cofatores, vitaminas, grupos prostéticos, pigmentos, tetrapirróis	5
At2g44920 At1g12250	Miscelânea, genômica comparativa planta-procarioto	4
Agrupamento Cmr CRISP	Metabolismo de DNA, CRISPs	2
Caspases bacterianas	Regulação e sinalização celular, morte celular programada e sistemas toxina-antitoxina	2
Biossíntese de glicosil-glicerídeos em plantas	Ácidos graxos, lipídeos e isoprenoides	1
Hidrolases agindo em ligações carbono-enxofre	Metabolismo de enxofre	1
Desvio cianobacteriano no ciclo do ácido tricarbóxico	Metabolismo central, ciclo do ácido tricarbóxico	1
Adesão de <i>Campylobacter</i>	Virulência, doença e defesa, adesão	1

A abundância das células cianobacterianas na comunidade é mantida em nível mais elevado em relação às células de outras espécies, provavelmente devido ao fato de que as condições fornecidas pelo meio de cultura estão direcionadas às necessidades da fisiologia cianobacteriana. Consequentemente, apesar de genes para fotossíntese e fixação de nitrogênio poderem ser encontrados no metagenoma da comunidade associada, as células cianobacterianas possivelmente são a maior fonte de carbono e nitrogênio fixado neste microambiente, dada a ausência de fontes de carbono e nitrogênio na composição do meio de cultura. É provável que a cianobactéria esteja fornecendo a pelo menos alguns membros de sua comunidade associada não apenas o carbono fixado que receberiam de plantas, mas também nitrogênio fixado. A falta de genes para fixação de nitrogênio nos rascunhos genômicos de bactérias heterotróficas obtidos sugerem que algumas populações bacterianas podem estar se adaptando à grande oferta de metabólitos cianobacterianos à sua disposição pela eliminação de genes que se tornaram redundantes neste contexto, o que, por sua vez, pode torná-las mais dependentes das células cianobacterianas.

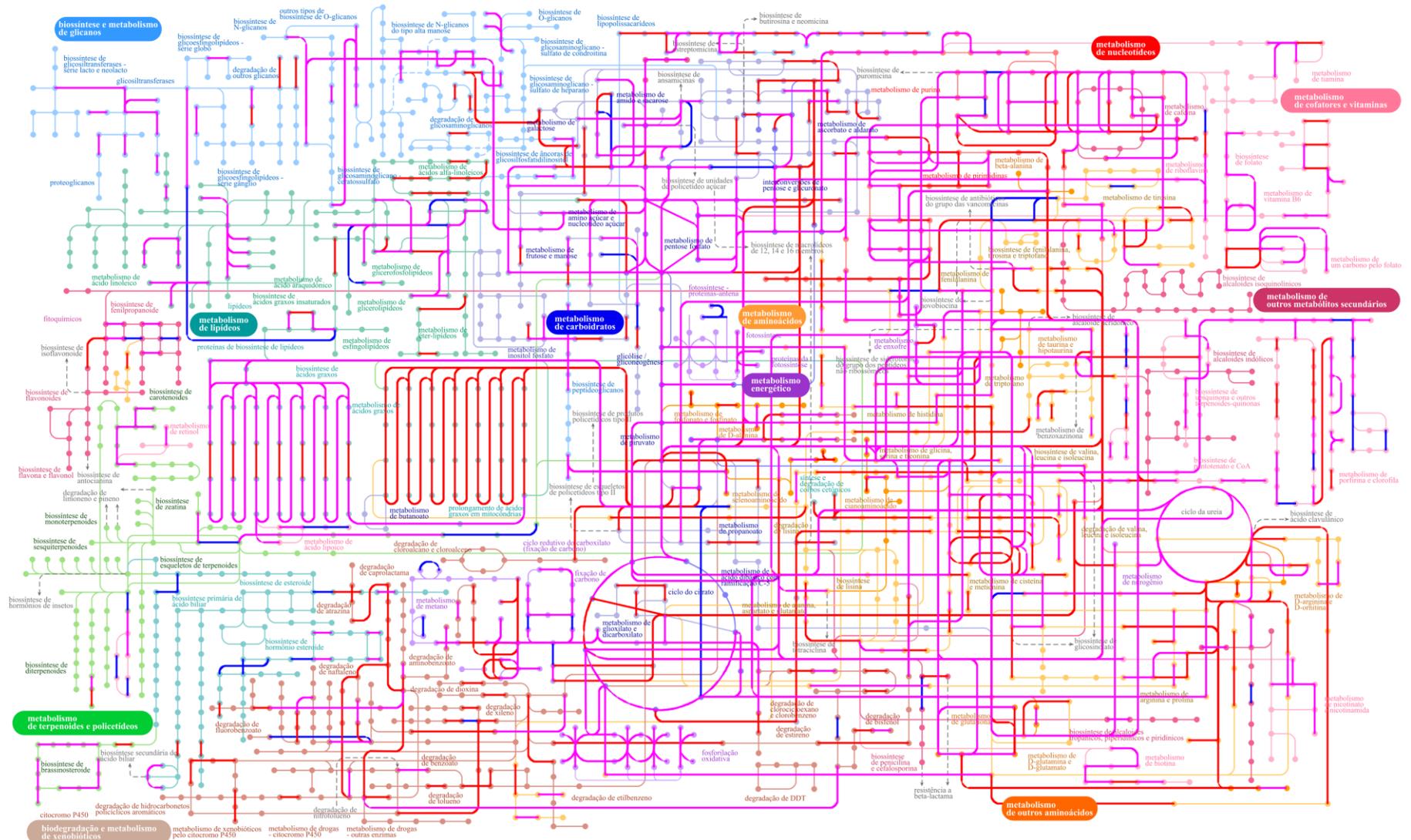


Figura 23 – Comparação de rotas metabólicas presentes no genoma cianobacteriano e/ou no metagenoma de sua comunidade associada obtidas por comparações com o banco de dados KEGG. Rosa: processos comuns à cianobactéria e à comunidade associada. Vermelho: vias específicas à comunidade associada. Azul: vias específicas à cianobactéria

Algumas cianobactérias não diazotróficas parecem não utilizar o nitrogênio fixado por sua comunidade associada, mesmo em ambientes em que a eficiência desta é maior (LÓPEZ-LOZANO et al., 2002). Em contrapartida, marcadores isotópicos indicam que pode ocorrer transferência de fósforo entre cianobactérias e bactérias simbiotes em ambas as direções, de modo que a comunidade associada pode ser vista como uma espécie de banco de fósforo temporário para a cianobactéria (JIANG et al., 2007).

Em relação à captação de ferro, indispensável para a fotossíntese e para a fixação de nitrogênio, entre outros processos, algumas bactérias associadas a cianobactérias são capazes de adquirir esse elemento a partir de fontes diversas, e em taxas maiores que a cianobactéria à qual se associam (ROE et al., 2012), e é possível que haja utilização pela cianobactéria de parte do ferro captado por uma bactéria associada (BELIAEV et al., 2014). Esse fato é importante tendo em vista que genes para sideróforos não são comuns em cianobactérias de alguns ambientes (HOPKINSON; MOREL, 2009). Genes classificados no subsistema de aquisição e metabolismo de ferro e sideróforos foram encontrados tanto no metagenoma da comunidade associada como no genoma da linhagem cianobacteriana estudada, o que sugere independência de ambas em relação às estratégias de obtenção desse elemento.

A variante microviridina J apresenta toxicidade a *Daphnia* (ROHRLACK et al., 2003; 2004). Dentre as dez variantes de microviridina que foram encontradas até o momento, as variantes G e H foram descritas a partir de um espécime de *Nostoc minutum* (MURAKAMI et al., 1997). A maior parte dos estudos de associações em ambientes naturais se concentrou em florações de espécies de *Microcystis*, porém sequências de microviridina também foram encontradas em trombólitos de *Nostoc commune* (GRAHAM et al., 2014). Alta expressão de microviridina foi observada em uma floração de *Microcystis*, o que estimou-se não somente contribuir para sua dominância competitiva, mas também conferir propriedades antipredatórias à comunidade (PENN et al., 2014). Considerando que a associação com determinadas espécies pode beneficiar certas cianobactérias e que o observado para microviridina J poderia ser válido para outras variantes de microviridina, é provável que a produção de microviridina por *Nostoc* sp. CENA67 constitua uma vantagem para os micro-organismos que se associam a esta cianobactéria e que esta empregue microviridina no recrutamento de simbiotes.

Ao passo que um número considerável de sobreposição entre vias foi observada no mapeamento metabólico, alguns metabolismos sugerem a complementariedade entre vias presentes exclusivamente na cianobactéria ou na comunidade associada. Como a maioria das

sequências cianobacterianas não foi identificada em subsistemas conhecidos e quase a metade do metagenoma da comunidade associada foi identificado como proteínas desconhecidas, é possível que fatores ainda não conhecidos expressos pelas células cianobacterianas possam influenciar interações com seus micro-organismos associados. Da mesma maneira, fatores incluídos no número considerável de proteínas desconhecidas e processos pouco caracterizados no metagenoma associado podem desempenhar algum papel nessa interação.

De forma semelhante ao observado em raízes de plantas, certas actinobactérias associadas a cianobactérias do gênero *Nostoc* promovem o crescimento desta como consequência da produção de substâncias reguladoras de crescimento e de antibióticos, e recebem nutrientes em troca desse serviço (INTHASOTTI; PATHOM-AREE, 2015). Na rizosfera, a presença de micro-organismos sem função ecológica aparente é de certa forma explicada pelo fato de que plantas podem digerir micro-organismos em suas raízes como forma de obter nutrientes e promover crescimento (PAUNGFOO-LONHIENNE et al., 2010). Ao passo que diversas bactérias são capazes de degradar células cianobacterianas para consumo (RASHIDAN; BIRD, 2001), algumas cianobactérias, incluindo espécies de *Nostoc*, são capazes de mixotrofia, ou a mistura de metabolismos autotróficos e heterotróficos, e podem apresentar aumento de sua atividade antimicrobiana sob essa condição (NOWRUZI et al., 2012). Logo, hipoteticamente, o estímulo do crescimento e o consumo e a predação de ambas as partes da interação na cultura estudada poderia ocorrer.

Contudo, as interações entre a cianobactéria e sua comunidade associada provavelmente são mais complexas que meras trocas de moléculas orgânicas, já que meios de cultura acrescidos de extratos de cianobactérias não parecem causar impacto positivo sobre o cultivo de bactérias associadas (INTHASOTTI; PATHOM-AREE, 2015). Flutuações na composição da comunidade microbiana associada a cianobactérias ocorrem de acordo com o período em que a colônia se encontra, e incluem a alternância da dominância (SHI et al., 2009; 2011), o que provavelmente resulta em alterações na dinâmica dessa interação.

Em alguns casos, respostas ao estresse oxidativo podem desempenhar um papel nas interações entre cianobactérias e bactérias heterotróficas (MORRIS et al., 2008). De fato, algumas cianobactérias têm seu crescimento estimulado pelo cocultivo com bactérias heterotróficas não somente devido ao aumento da concentração do dióxido de carbono ou à captura de refugos de ciclagem, mas também à redução do estresse oxidativo (BELIAEV et al., 2014; SHER et al., 2011). Além disso, é possível que fatores físicos e morfológicos tenham influência sobre a associação entre a cianobactéria e os

micro-organismos heterótrofos. Cianobactérias aquáticas se beneficiam de agregados microbianos por estes lhes conferirem proteção contra radiação ultravioleta, calor e ozônio (TANG; DZIALLAS; GROSSART, 2011), e interações com micro-organismos associados podem ser responsáveis por induzir a produção de mucilagem e modificar as características morfológicas de colônias de cianobactérias (SHEN et al., 2011). Algumas bactérias podem ter conexões ainda mais fortes com exopolissacarídeos que com as células das cianobactérias que os secretam (SHIRAISHI et al., 2015).

Por ser unicianobacteriano e conter os resultados de anos de cultivo, a cultura estudada proporciona um objeto de estudo com menor número de variáveis que interações observadas em ambientes naturais. Contudo, dada a complexidade e a diversidade das associações entre autótrofos e heterótrofos e em vista do alto número de sequências anotadas em funções desconhecidas, é difícil assegurar neste momento o que *Nostoc* sp. CENA67 poderia ganhar com as interações com sua comunidade associada. É necessário que as sequências e as anotações em bancos de dados públicos se tornem mais abrangentes antes que esta questão seja respondida.

7. CONCLUSÕES

A linhagem CENA67 diverge em certos aspectos de outras linhagens identificadas no gênero *Nostoc*, e pode ser representante de um novo táxon. Contudo, devido ao atual baixo número de linhagens conhecidas relacionadas, é difícil assegurar sua posição taxonômica, sendo necessário isolar novas cianobactérias relacionadas a essa linhagem para que suas relações evolutivas sejam esclarecidas e permitam uma descrição adequada desse táxon. Essa distinção é refletida em seu genoma, que é significativamente distinto dos genomas atualmente disponíveis de *Nostoc* spp. Conseqüentemente, o genoma dessa linhagem apresenta diversos agrupamentos gênicos hipotéticos possivelmente envolvidos na síntese de metabólitos ainda desconhecidos ou de novas variantes de moléculas conhecidas, incluindo grande variedade de microviridinas. Um alto número de micro-organismos associados à cianobactéria em cultura não axênica pertence a táxons comumente relatados em interações com plantas, que têm à sua disposição neste ambiente não apenas o carbono fixado que receberiam destas, mas também nitrogênio biodisponível. Essa condição parece estar sujeitando algumas das populações bacterianas heterotróficas a pressões de seleção que as levam a se adaptar pela eliminação de vias custosas que se tornaram redundantes, como a fixação de nitrogênio, e as tornar mais dependentes da cianobactéria, tornando a axenia da cianobactéria ainda mais difícil de ser alcançada. O baixo conhecimento a respeito da genômica funcional de cianobactérias e de micro-organismos associados, refletido no alto número de funções desconhecidas encontrado, não deixa claro se outros fatores desempenham algum papel nessa interação.

REFERÊNCIAS

- ADAMS, D.G. Symbiotic interactions. In: WHITTON, B.A.; POTTS, M. (Ed.). **The ecology of cyanobacteria**. Dordrecht: Kluwer Academic, 2000. p. 523-561.
- ADAMS, D.G.; BERGMAN, B.; NIERZWICKI-BAUER, S.A.; DUGGAN, P.S.; RAI, A.N.; SCHÜSSLER, A. Cyanobacterial-plant symbioses. In: ROSENBERG, E.; DELONG, E.F.; LORY, S.; STACKBRANDT, E.; THOMPSON, F. (Ed.). **The prokaryotes 1**. 3. ed. Heidelberg: Springer, 2013. p. 359-400.
- AHMED, M.; STAL, L.J.; HASNAIN, S. Association of non-heterocystous cyanobacteria with crop plants. **Plant and Soil**, Dordrecht, v. 336, p. 363-375, 2010.
- ALBERTSEN, M.; HUGENHOLTZ, P.; SKARSHEWSKI, A.; NIELSEN, K.L.; TYSON, G.W.; NIELSEN, P.H. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. **Nature Biotechnology**, New York, v. 31, p. 533-538, 2013.
- ALLEN, M.B.; ARNON, D.I. Studies on nitrogen-fixing blue-green algae. I. Growth and nitrogen fixation by *Anabaena cylindrica* Lemm. **Plant Physiology**, Rockville, v. 30, p. 366-372, 1955.
- ALNEBERG, J.; BJARNASON, B.S.; DE BRUIJN, I.; SCHIRMER, M.; QUICK, J.; IJAZ, U.Z.; LAHTI, L.; LOMAN, N.J.; ANDERSSON, A.F.; QUINCE, C. Binning metagenomic contigs by coverage and composition. **Nature Methods**, New York, v. 11, p. 1144-1146, 2014.
- ALTSCHUL, S.F.; GISH, W.; MILLER, W.; MYERS, E.W.; LIPMAN, D.J. Basic local alignment search tool. **Journal of Molecular Biology**, London, v. 215, p. 403-410, 1990.
- AZIZ, R.K.; BARTELS, D.; BEST, A.A.; DEJONGH, M.; DISZ, T.; EDWARDS R.A.; FORMSMA, K.; GERDES, S.; GLASS, E.M.; KUBAL, M.; MEYER, F.; OLSEN, G.J.; OLSON, R.; OSTERMAN, A.L.; OVERBEEK, R.A.; MCNEIL, L.K.; PAARMMANN, D.; PACZIAN, T.; PARELLO, B.; PUSCH, G.D.; REICH, C.; STEVENS, R.; VASSIEVA, O.; VONSTEIN, V.; WILKE, A.; ZAGNITKO, O. The RAST server: rapid annotations using subsystems technology. **BMC Genomics**, London, v. 9, p. 75, 2008.
- BAGATINI, I.L.; EILER, A.; BERTILSSON, S.; KLAVENESS, D.; TESSAROLLI, L.P.; VIEIRA, A.A.H. Host-specificity and dynamics in bacterial communities associated with bloom-forming freshwater phytoplankton. **PLoS One**, San Francisco, v. 9, p. e85950, 2014.
- BANKEVICH, A.; NURK, S.; ANTIPOV, D.; GUREVICH, A.A.; DVORKIN, M.; KULIKOV, A.S.; LESIN, V.M.; NIKOLENKO, S.I.; PHAM, S.; PRJIBELSKI, A.D.; PYSHKIN, A.V.; SIROTKIN, A.V.; VYAHHI, N.; TESLER, G.; ALEKSEYEV, M.A.; PEVZNER, P.A. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. **Journal of Computational Biology**, New York, v. 19, p. 455-477, 2014.

BATTISTUZZI, F.U.; FEIJÃO, A.; HEDGES, S.B. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. **BMC Evolutionary Biology**, London, v. 4, p. 44-58, 2004.

BELIAEV, A.S.; ROMINE, M.F.; SERRES, M.; BERNSTEIN, H.C.; LINGGI, B.E.; MARKILLIE, L.M.; ISERN, N.G.; CHRISLER, W.B.; KUCEK, L.A.; HILL, E.A.; PINCHUK, G.E.; BRYANT, D.A.; WILEY, H.S.; FREDRICKSON, J.K.; KONOPKA, A. Inference of interactions in cyanobacterial-heterotrophic co-cultures via transcriptome sequencing. **The ISME Journal**, Oxon, v. 8, p. 2243-2255, 2014.

BENTLEY, S.D.; CHATER, K.F.; CERDENO-TARRAGA, A.M.; CHALLIS, G.L.; THOMSON, N.R.; JAMES, K.D.; HARRIS, D.E.; QUAIL, M.A.; KIESER, H.; HARPER, D.; BATEMAN, A.; BROWN, S.; CHANDRA, G.; CHEN, C.W.; COLLINS, M.; CRONIN, A.; FRASER, A.; GOBLE, A.; HIDALGO, J.; HORNSBY, T.; HOWARTH, S.; HUANG, C.H.; KIESER, T.; LARKE, L.; MURPHY, L.; OLIVER, K.; O'NEIL, S.; RABBINOWITSCH, E.; RAJANDREAM, M.A.; RUTHERFORD, K.; RUTTER, S.; SEEGER, K.; SAUNDERS, D.; SHARP, S.; SQUARES, R.; SQUARES, S.; TAYLOR, K.; WARREN, T.; WIETZORREK, A.; WOODWARD, J.; BARRELL, B.G.; PARKHILL, J.; HOPWOOD, D.A. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). **Nature**, London, v. 417, p. 141–147, 2002.

BERG, K.A.; LYRA, C.; SIVONEN, K.; PAULIN, L.; SUOMALAINEN, S.; TUOMI, P.; RAPALA, J. High diversity of cultivable heterotrophic bacteria in association with cyanobacterial water blooms. **The ISME Journal**, Oxon, v. 3, p. 314-325, 2009.

BERGMAN, B.; RAI, A.N.; RASMUSSEN, U. Cyanobacterial associations. In: ELMERICH, C.; NEWTON, W.E. (Ed.). **Associative and endophytic nitrogen-fixing bacteria and cyanobacterial associations**. New York: Springer, 2007. p. 257-301.

BILLER, S.J.; BERUBE, P.M.; BERTA-THOMPSON, J.W.; KELLY, L.; ROGGENSACK, S.E.; AWAD, L.; ROACHE-JOHNSON, K.H.; DING, H.; GIOVANNONI, S.J.; ROCAP, G.; MOORE, L.R.; CHISHOLM, S.W. Genomes of diverse isolates of the marine cyanobacterium *Prochlorococcus*. **Scientific Data**, London, v. 1, p. 140034, 2014.

BLIN, K.; MEDEMA, M.H.; KAZEMPOUR, D.; FISCHBACH, M.A.; BREITLING, R.; TAKANO, E.; WEBER, T. antiSMASH 2.0 – a versatile platform for genome mining of secondary metabolite producers. **Nucleic Acids Research**, London, v. 41, p. W204-W212, 2013.

BLUNT, J.W.; COPP, B.R.; KEYZERS, R.A.; MUNRO, M.H.G.; PRINSEP, M.R. Marine natural products. **Natural Products Reports**, London, v. 32, p. 116-211, 2015.

BOHUNICKÁ, M.; PIETRASIAK, N.; JOHANSEN, J.R.; BERRENDERO GOMÉZ, E.; HAUER, T.; GAYSINA, L.A.; LUKEŠOVÁ, A. *Roholtiella*, gen. nov. (Nostocales, Cyanobacteria) – a tapering and branching cyanobacteria of the family Nostocaceae. **Phytotaxa**, Auckland, v. 197, p. 84-103, 2015.

BÖRNER, T.; DITTMANN, E. Molecular biology of cyanobacterial toxins. In: HUISMAN, J.; MATTHIJS, H.C.P.; VISSER, P.M. (Ed.). **Harmful cyanobacteria**. Dordrecht: Springer, 2005. p. 25-40.

BRAUER, V.S.; STOMP, M.; BOUVIER, T.; FOUILLAND, E.; LÉBOULANGER, C.; CONFURIUS-GUNS, V.; WEISSING, F.J.; STAL, L.J.; HUISMAN, J. Competition and facilitation between the marine nitrogen-fixing cyanobacterium *Cyanothece* and its associated bacterial community. **Frontiers in Microbiology**, Lausanne, v. 5, p. 795, 2015.

BRENNER, K.; YOU, L.; ARNOLD, F.H. Engineering microbial consortia: a new frontier in synthetic biology. **Trends in Biotechnology**, Cambridge, v. 26, p. 483-489, 2008.

BRETTIN, T.; DAVIS, J.J.; DISZ, T.; EDWARDS, R.A.; GERDES, S.; OLSEN, G.J.; OLSON, R.; OVERBEEK, R.; PARRELLO, B.; PUSCH, G.D.; SHUKLA, M.; THOMASON 3RD, J.A.; STEVENS, R.; VONSTEIN, V.; WATTAM, A.R.; XIA, F. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. **Scientific Reports**, London, v. 5, p. 8365, 2015.

BRUCE, T.; CASTRO, A.; KRUGER, R.; THOMPSON, C.C.; THOMPSON, F.L. Microbial diversity of Brazilian biomes. In: NELSON, K.E.; JONES-NELSON, B. (Ed.). **genomic applications for the developing world**. New York: Springer, 2012. p. 217-247.

CAI, H.Y.; YAN, Z.S.; WANG, A.J.; KRUMHOLZ, L.R.; JIANG, H.L. Analysis of the attached microbial community on mucilaginous cyanobacterial aggregates in the eutrophic lake Taihu reveals the importance of Planctomycetes. **Microbial Ecology**, New York, v. 66, p. 73-83, 2013.

CALTEAU, A.; FEWER, D.P.; LATIFI, A.; COURSIN, T.; LAURENT, T.; JOKELA, J.; KERFELD, C.A.; SIVONEN, K.; PIEL, J.; GUGGER, M. Phylum-wide comparative genomics unravel the diversity of secondary metabolism in Cyanobacteria. **BMC Genomics**, London, v. 15, p. 977, 2014.

CAMACHO, C.; COULOURIS, G.; AVAGYAN, V.; MA, N.; PAPADOPOULOS, J.; BEALER, K.; MADDEN, T.L. BLAST+: architecture and applications. **BMC Bioinformatics**, London, v. 10, p. 421, 2009.

CAPORASO, J.G.; KUCZYNSKI, J.; STOMBAUGH, J.; BITTINGER, K.; BUSHMAN, F.D.; COSTELLO, E.K.; FIERER, N.; PEÑA, A.G.; GOODRICH, J.K.; GORDON, J.I.; HUTTLEY, G.A.; KELLEY, S.T.; KNIGHTS, D.; KOENIG, J.E.; LEY, R.E.; LOZUPONE, C.A.; MCDONALD, D.; MUEGGE, B.D.; PIRRUNG, M.; REEDER, J.; SEVINSKY, J.R.; TURNBAUGH, P.J.; WALTERS, W.A.; WIDMANN, J.; YATSUNENKO, T.; ZANEVELD, J.; KNIGHT, R. QIIME allows analysis of high-throughput community sequencing data. **Nature Methods**, New York, v. 7, p. 335-336, 2010a.

CAPORASO, J.G.; BITTINGER, K.; BUSHMAN, F.D.; DESANTIS, T.Z.; ANDERSEN, G.L.; KNIGHT, R. PyNAST: a flexible tool for aligning sequences to a template alignment. **Bioinformatics**, Oxford, v. 26, p. 266-267, 2010b.

CASAMATTA, D.A.; GOMEZ, S.R.; JOHANSEN, J.R. *Rexia erecta* gen. nov. et sp. nov., two newly described cyanobacterial taxa from the Great Smoky Mountains National Park (USA). **Hydrobiologia**, Dordrecht, v. 561, p. 13-26, 2006.

CHABAN, V.V.; NIELSEN, M.B.; KOPEC, W.; KHANDELIA, H. Insights into the role of cyclic ladderane lipids in bacteria from computer simulations. **Chemistry and Physics of Lipids**, Amsterdam, v. 181, p. 76-82, 2014.

CHAISSON, M.J.; BRINZA, D.; PEVZNER, P.A. De novo fragment assembly with short mate-paired reads: does the read length matter? **Genome Research**, Woodbury, v. 19, p. 336-346, 2009.

CHAPMAN, M.J.; MARGULIS, L. Morphogenesis by symbiogenesis. **International Microbiology**, Barcelona, v. 1, p. 319-326, 1998.

CHIN, C.S.; ALEXANDER, D.H.; MARKS, P.; KLAMMER, A.A.; DRAKE, J.; HEINER, C.; CLUM, A.; COPELAND, A.; HUDDLESTON, J.; EICHLER, E.E.; TURNER, S.W.; KORLACH, J. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. **Nature Methods**, New York, v. 10, p. 563-569, 2013.

CHOI, G.G.; BAE, M.S.; AHN, C.Y.; OH, H.M. Induction of axenic culture of *Arthrospira* (*Spirulina*) *platensis* based on antibiotic sensitivity of contaminating bacteria. **Biotechnology Letters**, v. 30, p. 87-92, 2008.

CHOI, D.H.; ZHANG, G.I.; NOH, J.H.; KIM, W.S.; CHO, B.C. *Gracilimonas tropica* gen. nov., sp. nov., isolated from a *Synechococcus* culture. **International Journal of Systematic and Evolutionary Microbiology**, Reading, v. 59, p. 1167-1172, 2009.

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

CIRÉS, S.; WÖRMER, L.; BALLOT, A.; AGHA, R.; WIEDNER, C.; VELÁZQUEZ, D.; CASERO, M.C.; QUESADA, A. Phylogeography of cylindrospermopsin and paralytic shellfish toxin-producing Nostocales cyanobacteria from Mediterranean Europe (Spain). **Applied and Environmental Microbiology**, Baltimore, v. 80, p. 1359-1370, 2014.

COHEN, Y.; GUREVITZ, M. The cyanobacteria – ecology, physiology and molecular genetics. In: DWORKIN, M.; FALKOW, S.; ROSENBERG, E.; SCHLEIFER, K.H.; STACKEBRANDT, E. (Ed.). **The prokaryotes** 4. 3. ed. New York: Springer, 2006. p. 1074-1098.

COLE, J.K.; HUTCHISON, J.R.; RENSLOW, R.S.; KIM, Y.M.; CHRISLER, W.B.; ENGELMANN, H.E.; DOHNALKOVA, A.C.; HU, D.; METZ, T.O.; FREDRICKSON, J.K.; LINDEMANN, S.R. Phototrophic biofilm assembly in microbial-mat-derived unicyanobacterial consortia: model systems for the study of autotroph-heterotroph interactions. **Frontiers in Microbiology**, Lausanne, v. 5, p. 109, 2014.

EILER, A.; BERTILSSON, S. Composition of freshwater cyanobacterial communities associated with cyanobacterial blooms in four Swedish lakes. **Environmental Microbiology**, Oxford, v. 6, p. 1228-1243, 2004.

CONTRERAS-MOREIRA, B.; VINUESA, P. GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. **Applied and Environmental Microbiology**, Baltimore, v. 79, p. 7696-7701, 2013.

CORRE, C.; CHALLIS, G.L. New natural product biosynthetic chemistry discovered by genome mining. **Natural Products Reports**, London, v. 26, p. 977-986, 2009.

COSTA, M.L.; BEHLING, H.; SUGUIO, K.; KÄMPF, N.; KERN, D.C. Paisagens amazônicas sob a ocupação do homem pré-histórico: uma visão geológica. In: TEIXEIRA, W.G.; KERN, D.C.; MADARI, B.E.; LIMA, H.N.; WOODS, W. (Ed.). **As terras pretas de índio da Amazônia**. Manaus: Embrapa Amazônia Oriental, 2009. p. 15-38.

COTTER, P.D.; ROSS, R.P.; HILL, C. Bacteriocins – a viable alternative to antibiotics? **Nature Reviews Microbiology**, London, v. 11, p. 95-105, 2013.

CRAWFORD, J.M.; CLARDY, J. Microbial genome mining answers longstanding biosynthetic questions. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v. 109, p. 7589-7590, 2012.

DARRIBA, D.; TABOADA, G.L.; DOALLO, R.; POSADA, D. jModelTest 2: more models, new heuristics and parallel computing. **Nature Methods**, New York, v. 9, p. 772, 2012.

DE FILIPPO, C.; RAMAZZOTTI, M.; FONTANA, P.; CAVALIERI, D. Bioinformatic approaches for functional annotation and pathway inference in metagenomics data. **Briefings in Bioinformatics**, London, v. 13, p. 696-710, 2012.

DE PHILIPPIS, R.; COLICA, G.; MICHELETTI, E. Exopolysaccharide-producing cyanobacteria in heavy metal removal from water: molecular basis and practical applicability of the biosorption process. **Applied Microbiology and Biotechnology**, Berlin, v. 92, p. 697-708, 2011.

DE RIJK, P.; WUYTS, J.; DE WACHTER, R. RnaViz2: an improved representation of RNA secondary structure. **Bioinformatics**, Oxford, v. 19, p. 299-300, 2003.

DEMBITSKY, V.M.; ŘEZANKA, T. Metabolites produced by nitrogen-fixing *Nostoc* species. **Folia Microbiologica**, Praha, v. 50, p. 363-393, 2005.

DESANTIS, T.Z.; HUGENHOLTZ, P.; LARSEN, N.; ROJAS, M.; BRODIE, E.L.; KELLER, K.; HUBER, T.; DALEVI, D.; HU, P.; ANDERSEN, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. **Applied and Environmental Microbiology**, Baltimore, v. 72, p. 5069-5072, 2006.

DI RIENZI, S.C.; SHARON, I.; WRIGHTON, K.C.; KOREN, O.; HUG, L.A.; THOMAS, B.C.; GOODRICH, J.K.; BELL, J.T.; SPECTOR, T.D.; BANFIELD, J.F.; LEY, R.E. The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. **eLife**, Cambridge, v. 2, p. e01102, 2013.

DITTMANN, E.; NEILAN, B.A.; BÖRNER, T. Molecular biology of the peptide and polyketide biosynthesis in bacteria. **Applied Microbiology and Biotechnology**, Berlin, v. 57, p. 467-473, 2001.

DODDS, W.K.; GUDDER, D.A.; MOLLENHAUER, D. The ecology of *Nostoc*. **Journal of Phycology**, Baltimore, v. 31, p. 2-18, 1995.

DONADIO, S.; MONCIARDINI, P.; SOSIO, M. Polyketide synthases and nonribosomal peptide synthetases: the emerging view from bacterial genomics. **Natural Products Reports**, London, v. 24, p. 1073-1109, 2007.

DONIA, M.S.; SCHMIDT, E.W. Linking chemistry and genetics in the growing cyanobactin natural products family. **Chemistry & Biology**, Cambridge, v. 18, p. 508-519, 2011.

DRÖGUE, J.; MCHARDY, A.C. Taxonomic binning of metagenome samples generated by next-generation sequencing technologies. **Briefings in Bioinformatics**, London, v. 13, p. 646-655, 2012.

EDGAR, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, London, v. 32, p. 1792-1797, 2004.

EDGAR, R.C. Search and clustering orders of magnitude faster than BLAST. **Bioinformatics**, Oxford, v. 26, p. 2460-2461, 2010.

EDWARDS, R.A.; OLSON, R.; DISZ, T.; PUSCH, G.D.; VONSTEIN, V.; STEVENS, R.; OVERBEEK, R. Real time metagenomics: usings k-mers to annotate metagenomes. **Bioinformatics**, Oxford, v. 28, p. 3316-3317, 2012.

EISEN, J.A.; FRASER, C.M. Phylogenomics: intersection of evolution and genomics. **Science**, Washington, DC, v. 300, p. 1706-1707, 2003.

EKMAN, M.; PICOSI, S.; CAMPBELL, E.L.; MEEKS, J.C.; FLORES, E. A *Nostoc punctiforme* sugar transporter necessary to establish a cyanobacterium-plant symbiosis. **Plant Physiology**, Rockville, v. 161, p. 1984-1992, 2013.

ENGINE, N.; CHOI, H.; ESQUENAZI, E.; ROTTACKER, E.C.; ELLISMAN, M.H.; DORRESTEIN, P.C.; GERWICK, W.H. Underestimated biodiversity as a major explanation for the perceived rich secondary metabolite capacity of the genus *Lyngbya*. **Environmental Microbiology**, Oxford, v. 13, p. 1601-1610, 2011.

EWING, B.; GREEN, P. Base-calling of automated sequencer traces using Phred: II. Error probabilities. **Genome Research**, Woodbury, v. 8, p. 186-194, 1998.

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

FAUST, K.; RAES, J. Microbial interactions: from networks to models. **Nature Reviews Microbiology**, London, v. 10, p. 538-550, 2012.

FELSENSTEIN, J. Evolutionary trees from DNA sequences: a maximum likelihood approach. **Journal of Molecular Evolution**, New York, v. 17, p. 368-376, 1981.

FIGLIORE, M.F.; NEILAN, B.A.; COPP, J.N.; RODRIGUES, J.L.M.; TSAI, S.M.; LEE, H.; TREVORS, J.T. Characterization of nitrogen-fixing cyanobacteria in the Brazilian Amazon floodplain. **Water Research**, Oxford, v. 39, p. 5017-5026, 2005.

GATTE-PICCHI, D.; WEIZ, A.; ISHIDA, K.; HERTWECK, C.; DITTMANN, E. Functional analysis of environmental DNA-derived microviridins provides new insights into the diversity of the tricyclic peptide family. **Applied and Environmental Microbiology**, Baltimore, v. 80, p. 1380-1387, 2014.

GEHRINGER, M.M.; ADLER, L.; ROBERTS, A.A.; MOFFITT, M.C.; MIHALI, T.K.; MILLS, T.J.T.; FIEKER, C.; NEILAN, B.A. Nodularin, a cyanobacterial toxin, is synthesized *in planta* by symbiotic *Nostoc* sp. **The ISME Journal**, Oxon, v. 6, p. 1834-1847, 2012.

GENUÁRIO, D.B.; VAZ, M.G.M.V.; HENTSCHKE, G.S.; SANT'ANNA, C.L.; FIGLIORE, M.F. *Halotia* gen. nov., a phylogenetically and physiologically coherent cyanobacterial genus isolated from marine coastal environments. **International Journal of Systematic and Evolutionary Microbiology**, Reading, v. 65, p. 663-675, 2015.

GIESBERGER, G. Some observations on the culture, physiology and morphology of some brown-red *Rhodospirillum*-species. **Antonie van Leeuwenhoek**, Wageningen, v. 13, p. 135-148, 1947.

GORDON, D.; ABAJIAN, C.; GREEN, P. Consed: a graphical tool for sequence finishing. **Genome Research**, Woodbury, v. 8, p. 195-202, 1998.

GRAHAM, L.E.; KNACK, J.J.; PIOTROWSKI, M.J.; WILCOX, L.W.; COOK, M.E.; WELLMAN, C.H.; TAYLOR, W.; LEWIS, L.A.; ARANCIBIA-AVILA, P. Lacustrine *Nostoc* (Nostocales) and associated microbiome generate a new type of modern clotted microbialite. **Journal of Phycology**, Baltimore, v. 50, p. 280-291, 2014.

GRIESE, M.; LANGE, C.; SOPPA, J. Ploidy in cyanobacteria. **FEMS Microbiology Letters**, Amsterdam, v. 323, p. 124-131, 2011.

GROSSMAN, J.M.; O'NEIL, B.E.; TSAI, S.M.; LIANG, B.; NEVES, E.; LEHMANN, J.; THIES, J.E. Amazonian anthrosols support similar microbial communities that differ distinctly from those extant in adjacent, unmodified soils of the same mineralogy. **Microbial Ecology**, New York, v. 60, p. 192-205, 2010.

GUGGER, M.F.; HOFFMANN, L. Polyphyly of true branching cyanobacteria (Stigonematales). **International Journal of Systematic and Evolutionary Microbiology**, Reading, v. 54, p. 349-357, 2004.

GUINDON, S.; GASCUEL, O. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. **Systematic Biology**, Austin, v. 52, p. 696-704, 2003.

GUPTA, R.S. Protein signatures (molecular synapomorphies) that are distinctive characteristics of the major cyanobacterial clades. **International Journal of Systematic and Evolutionary Microbiology**, Reading, v. 59, p. 2510-2526, 2009.

GUPTA, R.S.; MATHEWS, D.W. Signature proteins for the major clades of Cyanobacteria. **BMC Evolutionary Biology**, London, v. 10, p. 24, 2010.

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

GUSEV, M.V.; BAULINA, O.I.; GORELOVA, O.A.; LOBAKOVA, E.S.; KORZHENEVSKAYA, T.G. Artificial cyanobacterium-plant symbioses. In: RAI, A.N.; BERGMAN, B.; RASMUSSEN, U. (Ed.). **Cyanobacteria in symbiosis**. Dordrecht: Kluwer Academic, 2003. p. 253-312.

HAREL, A.; KARKAR, S.; CHENG, S.; FALKOWSKI, P.G.; BHATTACHARYA, D. Deciphering primordial cyanobacterial genome functions from protein network analysis. **Current Biology**, London, v. 25, p. 628-634, 2015.

HASEGAWA, M.; KISHINO, K.; YANO, T. dating the human-ape splitting by a molecular clock of mitochondrial DNA. **Journal of Molecular Evolution**, New York, v. 22, p. 160-174, 1985.

HENSON, B.J.; WATSON, L.E.; BARNUM, S.R. Molecular differentiation of the heterocystous cyanobacteria, *Nostoc* and *Anabaena*, based on complete *nifD* sequences. **Current Microbiology**, New York, v. 45, p. 161-164, 2002.

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

HOLLAND, A.; KINNEAR, S. Interpreting the possible ecological role(s) of cyanotoxins: compounds for competitive advantage and/or physiological aide? **Marine Drugs**, Basel, v. 11, p. 2239-2258, 2013.

HOPKINSON, B.M.; MOREL, F.M.M. The role of siderophores in iron acquisition by photosynthetic marine microorganisms. **Biometals**, Tucson, v. 22, p. 659-669, 2009.

HROUZEK, P.; LUKEŠOVÁ, A.; MAREŠ, J.; VENTURA, S. Description of the cyanobacterial genus *Desmonostoc* gen. nov. including *D. muscorum* comb. nov. as a distinct, phylogenetically coherent taxon related to the genus *Nostoc*. **Fottea**, Olomouc, v. 13, p. 201-213, 2013.

HUANG, J.; VAN ALLER, G.S.; TAYLOR, A.N.; KERRIGAN, J.J.; LIU, W.S.; TRULLI, J.M.; LAI, Z.; HOLMES, D.; AUBART, K.M.; BROWN, J.R.; ZALACAIN, M. Phylogenomic and biochemical characterization of three *Legionella pneumophila* polypeptide deformylases. **Journal of Bacteriology**, Baltimore, v. 188, p. 5249-57, 2006.

HUDDLESTON, J.; RANADE, S.; MALIG, M.; ANTONACCI, F.; CHAISSON, M.; HON, L.; SUDMANT, P.H.; GRAVES, T.A.; ALKAN, C.; DENNIS, M.Y.; WILSON, R.K.; TURNER, S.W.; KORLACH, J.; EICHLER, E.E. Reconstructing complex regions of genomes using long-read sequencing technology. **Genome Research**, Woodbury, v. 24, p. 688-696, 2014.

HÜGLER, M.; HUBER, H.; STETTER, K.O.; FUCHS, G. Autotrophic CO₂ fixation pathways in archaea (Crenarchaeota). **Archives of Microbiology**, Heidelberg, v. 179, p. 160-173, 2003.

HUMBERT, J.F.; BARBE, V.; LATIFI, A.; GUGGER, M.; CALTEAU, A.; COURSIN, T.; LAJUS, A.; CASTELLI, V.; OZTAS, S.; SAMSON, G.; LONGIN, C.; MEDIGUE, C.; TANDEAU DE MARSAC, N. A tribute to the disorder in the genome of the bloom-forming freshwater cyanobacterium *Microcystis aeruginosa*. **PLOS One**, San Francisco, v. 8, p. e70747, 2013.

HUNTER, S.; CORBETT, M.; DENISE, H.; FRASER, M.; GONZALEZ-BELTRAN, A.; HUNTER, C.; JONES, P.; LEINONEN, R.; MCANULLA, C.; MAGUIRE, E.; MASLEN, J.; MITCHELL, A.; NUKA, G.; OISEL, A.; PESSEAT, S.; RADHAKRISHNAN, R.; ROCCA-SERRA, P.; SCHEREMETJEW, M.; STERK, P.; VAUGHAN, D.; COCHRANE, G.; FIELD, D.; SANSONE, S.A. EBI metagenomics – a new resource for the analysis and archiving of metagenomic data. **Nucleic Acids Research**, London, v. 42, p. D600-D606, 2013.

HUSSAIN, A.; SHAH, S.T.; RAHMAN, H.; IRSHAD, M.; IQBAL, A. Effect of IAA on *in vitro* growth and colonization of *Nostoc* in plant roots. **Frontiers in Plant Science**, Lausanne, v. 6, p. 46, 2015.

IDI, A.; NOR, M.H.M.; WAHAB, M.F.A.; IBRAHIM, Z. Photosynthetic bacteria: an eco-friendly and cheap tool for bioremediation. **Reviews in Environmental Science and Bio/Technology**, Dordrecht, v. 14, p. 271-285, 2015.

IMELFORT, M.; PARKS, D.; WOODCROFT, B.J.; DENNIS, P.; HUGENHOLTZ, P.; TYSON, G.W. GroopM: an automated tool for the recovery of population genomes from related metagenomes. **PeerJ**, London, v. 2, p. e603, 2014.

IMHOFF, J.F. The phototrophic Alpha-Proteobacteria. In: DWORKIN, M.; FALKOW, S.; ROSENBERG, E.; SCHLEIFER, K.H.; STACKEBRANDT, E. (Ed.). **The prokaryotes** 5. 3. ed. New York: Springer, 2006. p. 41-64.

INTHASOTTI, T.; PATHOM-AREE, W. Diversity of actinobacteria associated with *Nostoc commune* Vaucher ex Bornet & Flahault macrocolonies. **Annals of Microbiology**, Milano, 2015. Disponível em: <<http://link.springer.com/article/10.1007%2Fs13213-015-1063-8>>. Acesso em: 27 ago. 2015.

ISHITSUKA, M.O.; KUSUMI, T.; KAKISAWA, H.; KAYA, K.; WATANABE, M.M. Microviridin: a novel tricyclic depsipeptide from the toxic cyanobacterium *Microcystis viridis*. **Journal of the American Chemical Society**, Washington, DC, v. 112, p. 8180-8182, 1990.

ITEMAN, I.; RIPPKA, R.; TANDEAU DE MARSAC, N.; HERDMAN, M. Comparison of conserved structural and regulatory domains within divergent 16S rRNA-23S rRNA spacer sequences of Cyanobacteria. **Microbiology**, Reading, v. 146, p. 1275-1286, 2000.

JASPARS, M. The origin of cyanobactin chemistry and biology. **Chemical Communications**, Cambridge, v. 50, p. 10174-10176, 2014.

JIANG, L.J.; YANG, L.Y.; XIAO, L.; SHI, X.L.; GAO, G.; QIN, B. Quantitative studies on phosphorus transference occurring between *Microcystis aeruginosa* and its attached bacterium (*Pseudomonas* sp.). **Hydrobiologia**, Dordrecht, v. 581, p. 161-165, 2007.

JOHANSEN, J.R.; KOVACIK, L.; CASAMATTA, D.A.; FUČÍKOVÁ, K.; KAŠTOVSKÝ, J. Utility of 16S-23S ITS sequence and secondary structure for recognition of intrageneric and intergeneric limits within cyanobacterial taxa: *Leptolyngbya corticola* sp. nov. (Pseudanabaenaceae, Cyanobacteria). **Nova Hedwigia**, Weinheim, v. 93, p. 283-302, 2011.

KAASALAINEN, U.; FEWER, D.P.; JOKELA, J.; WAHLSTEN, M.; SIVONEN, K.; RIKKINEN, J. Lichen species identity and diversity of cyanobacterial toxins in symbiosis. **New Phytologist**, London, v. 198, p. 647-651, 2013a.

KAASALAINEN, U.; FEWER, D.P.; JOKELA, J.; WAHLSTEN, M.; SIVONEN, K.; RIKKINEN, J. Cyanobacteria produce a high variety of hepatotoxic peptides in lichen symbiosis. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v. 109, p. 5886-5891, 2013b.

KAEBERNICK, M.; NEILAN, B.A. Ecological and molecular investigations of cyanotoxin production. **FEMS Microbiology Ecology**, Amsterdam, v. 35, p. 1-9, 2001.

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

KANEKO, T.; NAKAMURA, Y.; WOLK, C.P.; KURITZ, T.; SASAMOTO, S.; WATANABE, A.; IRIGUCHI, M.; ISHIKAWA, A.; KAWASHIMA, K.; KIMURA, T.; KISHIDA, Y.; KOHARA, M.; MATSUMOTO, M.; MATSUNO, A.; MURAKI, A.; NAKAZAKI, N.; SHIMPO, S.; SUGIMOTO, M.; TAKAZAWA, M.; YAMADA, M.; YASUDA, M.; TABATA, S. Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. **DNA Research**, Oxford, v. 8, p. 205-213, 2001.

KANG, D.D.; FROULA, J.; EGAN, R.; WANG, Z. **A robust statistical framework for reconstructing genomes from metagenomic data**. Cold Spring Harbor: bioRxiv, 2014. Disponível em: <http://biorxiv.org/content/biorxiv/early/2014/11/15/011460.full.pdf>. Acesso em: 27 ago. 2015.

KAPLAN, D. Absorption and adsorption of heavy metals by microalgae. In: RICHMOND, A.; HU, Q. (Ed.). **Handbook of microalgal culture: applied phycology and biotechnology**. 2. ed. New York: Wiley-Blackwell, 2013. p. 602-611.

KAPPLER, U.; DAVENPORT, K.; BEATSON, S.; LUCAS, S.; LAPIDUS, A.; COPELAND, A.; BERRY, K.W.; GLAVINA DEL RIO, T.; HAMMON, N.; DALIN, E.; TICE, H.; PITLUCK, S.; RICHARDSON, P.; BRUCE, D.; GOODWIN, L.A.; HAN, C.; TAPIA, R.; DETTER, J.C.; CHANG, Y.J.; JEFFRIES, C.D.; LAND, M.; HAUSER, L.; KYRPIDES, N.C.; GOKER, M.; IVANOVA, N.; KLENK, H.P.; WOYKE, T. Complete genome sequence of the facultatively chemolithoautotrophic and methylotrophic alpha Proteobacterium *Starkeya novella* type strain (ATCC 8093(T)). **Standards in Genomic Sciences**, East Lansing, v. 7, p. 44-58, 2012.

KAŠTOVSKÝ, J.; BERRENDERO GOMÉZ, E.; HLADIL, J.; JOHANSEN, J.R. *Cyanocohniella calida* gen. et sp. nov. (Cyanobacteria: Aphanizomenonaceae) a new cyanobacterium from the thermal springs from Karlovy Vary, Czech Republic. **Phytotaxa**, Auckland, v. 181, p. 279–292, 2014.

KAUFF, F.; BÜDEL, B. Phylogeny of cyanobacteria: an overview. **Progress in Botany**, Berlin, v. 72, p. 209-224, 2011.

KEHR, J.C.; PICCHI, D.G.; DITTMANN, E. Natural product biosyntheses in cyanobacteria: a treasure trove of unique enzymes. **Beilstein Journal of Organic Chemistry**, Frankfurt, v. 7, p. 1622-1635, 2011.

KLEIGREWE, K.; ALMALITI, J.; TIAN, I.Y.; KINNEL, R.B.; KOROBAYNIKOV, A.; MONROE, E.A.; DUGGAN, B.M.; DI MARZO, V.; SHERMAN, D.H.; DORRESTEIN, P.C.; GERWICK, L.; GERWICK, W.H. Combining mass spectrometric metabolic profiling with genomic analysis: a powerful approach for discovering natural products from cyanobacteria. **Journal of Natural Products**, Cincinnati, v. 78, p. 1671-1682, 2015.

KELLMANN, R.; MIHALI, T.K.; JEON, Y.J.; PICKFORD, R.; POMATI, F.; NEILAN, B.A. Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria. **Applied and Environmental Microbiology**, Baltimore, v. 74, p. 4044-4053, 2008.

KERN, D.; KÄMPF, N.; WOODS, W.I.; DENEVAN, W.M.; COSTA, M.L.; FRAZÃO, F.J.L.; SOMBROEK, W. Evolução do conhecimento em terra preta de índio. In: TEIXEIRA, W.G.; KERN, D.C.; MADARI, B.E.; LIMA, H.N.; WOODS, W. (Ed.). **As terras pretas de índio da Amazônia**. Manaus: Embrapa Amazônia Oriental, 2009. p. 72-81.

KIDWELL, M.G.; LISCH, D.R. Perspective: transposable elements, parasitic DNA and genome evolution. **Evolution**, Lancaster, v. 55, p. 1-24, 2001.

KNERR, P.J.; DONK, W.A. Discovery, biosynthesis and engineering of lantipeptides. **Annual Review of Biochemistry**, Palo Alto, v. 81, p. 479-505, 2012.

KOMÁREK, J. Modern taxonomic revision of planktic nostocacean cyanobacteria: a short review of genera. **Hydrobiologia**, Dordrecht, v. 639, p. 231-243, 2010a.

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

KOMÁREK, J. Cyanoprokaryota - 3. Teil/ Part 3: Heterocytous genera. In: BÜDEL, B.; GÄRTNER, G.; KRIENITZ, L.; SCHAGERL, M. (Ed.). **Süßwasserflora von Mitteleuropa 19/3**. Heidelberg: Springer/Spektrum, 2013.

KOMÁREK, J.; ANAGNOSTIDIS, K. Modern approach to the classification system of Cyanophytes 4 – Nostocales. **Algological Studies**, Stuttgart, v. 56, p. 247-345, 1989.

KOMÁREK, J.; HAUER, T. **CyanoDB.cz** - On-line database of cyanobacterial genera. - Word-wide electronic publication. České Budějovice, Czech Republic: University of South Bohemia; Institute of Botany, 2013. Disponível em: <http://www.cyanodb.cz>. Acesso em: 27 ago. 2015.

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

KOPF, M.; KLÄHN, S.; VOSS, B.; STÜBER, K.; HUETTEL, B.; REINHARDT, R.; HESS, W.R. Finished genome sequence of the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6714. **Genome Announcements**, Washington, DC, v. 2, p. e00757-14.

KOPYLOVA, E.; NOÉ, L.; TOUZET, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. **Bioinformatics**, Oxford, v. 28, p. 3211-3217, 2012.

KOREIVIENĖ, J.; ANNE, O.; KASPEROVIČIENĖ, J.; BURŠKYTĖ, V. Cyanotoxin management and human health risk mitigation in recreational waters. **Environmental Monitoring and Assessment**, Dordrecht, v. 186, p. 4443-4459, 2014.

KORELUSOVÁ, J. **Phylogeny of heterocytous cyanobacteria (Nostocales and Stigonematales)**. 2008. 33 f. Dissertation (M.Sc. in Botany) - Faculty of Science, University of South Bohemia, České Budějovice, 2008.

KOREN, S.; PHILLIPPY, A.M. One chromosome, one contig: complete microbial genomes from long-read sequencing and assembly. **Current Opinion in Microbiology**, Kidlington, v. 23, p. 110-120, 2015.

KRAUSE, A.; RAMAKUMAR, A.; BARTELS, D.; BATTISTONI, F.; BEKEL, T.; BOCH, J.; BÖHM, M.; FRIEDRICH, F.; HUREK, T.; KRAUSE, L.; LINKE, B.; MCHARDY, A.C.; SARKAR, A.; SCHNEIKER, S.; SYED, A.A.; THAUER, R.; VORHÖLTER, F.J.; PÜHLER, A.; REINHOLD-HUREK, B.; KAISER, O.; GOESMANN, A. Complete genome of the mutualistic, N₂-fixing grass endophyte *Azoarcus* sp. strain BH72. **Nature Biotechnology**, New York, v. 24, p. 1385-1391, 2006.

KRISTENSEN, D.M.; KANNAN, L.; COLEMAN, M.K.; WOLF, Y.; SOROKIN, A.; KOONIN, E.V.; MUSHEGIAN, A. A low-polynomial algorithm for assembling clusters of orthologous groups from intergenomic symmetric best matches. **Bioinformatics**, Oxford, v. 26, p. 1481-1487, 2010.

KUMAR, M.; SINGH, D.P.; PRABHA, R.; SHARMA, A.K. Role of cyanobacteria in nutrient cycle and use efficiency in the soil. In: RAKSHIT, A.; SINGH, H.B.; SEN, A. (Ed.). **Nutrient use efficiency: from basics to advances**. New Delhi: Springer, 2015. p. 163-171.

KURMAYER, R. The toxic cyanobacterium *Nostoc* sp. strain 152 produces highest amounts of microcystin and nostophycin under stress conditions. **Journal of Phycology**, Baltimore, v. 47. p. 200-207, 2011.

LAGE, O.M.; BONDOSO, J. Planctomycetes and macroalgae, a striking association. **Frontiers in Microbiology**, Lausanne, v. 5, p. 267, 2014.

LAND, M.; HAUSER, L.; JUN, S.R.; NOOKAEW, I.; LEUZE, M.R.; AHN, T.H.; KARPINETS, T.; LUND, O.; KORA, G.; WASSENAAR, T.; POUDEL, S.; USSERY, D.W. Insights from 20 years of bacterial genome sequencing. **Functional and Integrative Genomics**, Heidelberg, v. 15, p. 141-161, 2015.

LANE, D.J. 16S/23S rRNA sequencing. In: STACKEBRANDT, E.; GOODFELLOW, M. (Ed.). **Nucleic acid techniques in bacterial systematics**. Chichester: John Wiley & Sons, 1991. p. 115-175.

LARIMER, F.W.; CHAIN, P.; HAUSER, L.; LAMERDIN, J.; MALFATTI, S.; DO, L.; LAN, M.L.; PELLETIER, D.A.; BEATTY, J.T.; LANG, A.S.; TABITA, F.R.; GIBSON, J.L.; HANSON, T.E.; BOBST, C.; TORRES Y TORRES, J.L.; PERES, C.; HARRISON, F.H.; GIBSON, J.; HARWOOD, C.S. Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodospseudomonas palustris*. **Nature Biotechnology**, New York, v. 22, p. 55-61, 2004.

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

LEÃO, P.N.; RAMOS, V.; VALE, M.; MACHADO, J.P.; VASCONCELOS, V.M. Microbial community changes elicited by exposure to cyanobacterial allelochemicals. **Microbial Ecology**, New York, v. 63, p. 85-95, 2012b.

LEIKOSKI, N.; LIU, L.; JOKELA, J.; WAHLSTEN, M.; GUGGER, M.; CALTEAU, A.; PERMI, P.; KERFELD, C.A.; SIVONEN, K.; FEWER, D.P. Genome mining expands the chemical diversity of the cyanobactin family to include highly modified linear peptides. **Chemistry & Biology**, Cambridge, v. 20, p. 1033-1043, 2013.

LI, B.; SHER, D.; KELLY, L.; SHI, Y.; HUANG, K.; KNERR, P.J.; JOEWONO, I.; RUSCH, D.; CHISHOLM, S.W.; DONK, W.A. Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v. 107, p. 10430-10435, 2010.

LI, L.; STOECKERT JR, C.J.; ROOS, D.S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. **Genome Research**, Woodbury, v. 13, p. 2178-2189, 2003.

LI, N.; ZHANG, L.; LI, F.; WANG, Y.; ZHU, Y.; KANG, H.; WANG, S.; QIN, S. Metagenome of microorganisms associated with the toxic cyanobacteria *Microcystis aeruginosa* analyzed using the 454 sequencing platform. **Chinese Journal of Oceanology and Limnology**, Beijing, v. 29, p. 505-513, 2011.

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

LIMA, A.R.J.; SIQUEIRA, A.S.; SANTOS, B.G.S.; SILVA, F.D.F.; INADA, D.T.; LIMA, C.P.; CARDOSO, J.F.; VIANEZ-JÚNIOR, J.L.S.G.; NUNES, M.R.T.; GONÇALVES, E.C. Draft genome sequence of *Blastomonas* sp. strain CACIA 14H2, a heterotrophic bacterium associated with cyanobacteria. **Genome Announcements**, Washington, DC, v. 2, p. e01200-13, 2014a.

LIMA, A.R.J.; SIQUEIRA, A.S.; SANTOS, B.G.S.; SILVA, F.D.F.; INADA, D.T.; LIMA, C.P.; CARDOSO, J.F.; VIANEZ-JÚNIOR, J.L.S.G.; NUNES, M.R.T.; GONÇALVES, E.C. Draft genome sequence of *Rhodobacter* sp. strain CACIA 14H1, a heterotrophic bacterium obtained from a nonaxenic culture of a *Cyanobium* species. **Genome Announcements**, Washington, DC, v. 2, p. e01116-13, 2014b.

LIMA, A.R.J.; SIQUEIRA, A.S.; SANTOS, B.G.S.; SILVA, F.D.F.; LIMA, C.P.; CARDOSO, J.F.; VIANEZ-JÚNIOR, J.L.S.G.; DALL'AGNOL, L.T.; MCCULLOCH, J.A.; NUNES, M.R.T.; GONÇALVES, E.C. Draft genome sequence of the Brazilian *Cyanobium* sp. strain CACIAM 14. **Genome Announcements**, Washington, DC, v. 2, p. e00669-14, 2014c.

LIN, S.; HASS, S.; ZEMOJTEK, T.; XIAO, P.; VINGRON, M.; LI, R. Genome-wide comparison of cyanobacterial transposable elements, potential genetic diversity indicators. **Gene**, Amsterdam, v. 473, p. 139-149, 2011.

LÓPEZ-LOSANO, A.; DIEZ, J.; EL ALAOUI, S.; MORENO-VIVIÁN, C.; GARCÍA-FERNANDEZ, J.M. Nitrate is reduced by heterotrophic bacteria but not transferred to *Prochlorococcus* in non-axenic cultures. **FEMS Microbiology Ecology**, Amsterdam, v. 41, p. 151-160, 2002.

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

MAKSIMOV, M.O.; LINK, A.J. Prospecting genomes for lasso peptides. **Journal of Industrial Microbiology and Biotechnology**, Heidelberg, v. 41, p. 333-344, 2014.

MAKSIMOV, M.O.; PELCZER, I.; LINK, A.J. Precursor-centric genome-mining approach for lasso peptide discovery. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v. 109, p. 15223-15228, 2012.

MANDE, S.S.; MOHAMMED, M.H.; GHOSH, T.S. Classification of metagenomic sequences: methods and challenges. **Briefings in Bioinformatics**, London, v. 13, p. 669-681, 2012.

MARDIS, E.R. Next-generation sequencing platforms. **Annual Review of Analytical Chemistry**, Palo Alto, v. 6, p. 287-303, 2013.

MASSON-BOIVI, C.; GIRAUD, E.; PERRET, X.; BATUT, J. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? **Trends in Microbiology**, Cambridge, v. 17, p. 458-466, 2009.

MATEO, P.; PERONA, E.; BERRENDERO, E.; LEGANÉS, F.; MARTÍN, M.; GOLUBIĆ, S. Life cycle as a stable trait in the evaluation of diversity of *Nostoc* from biofilms in rivers. **FEMS Microbiology Ecology**, Amsterdam, v. 76, p. 185-298, 2011.

MAU, B.; NEWTON, M.; LARGET, B. Bayesian phylogenetic inference via Markov chain Monte Carlo methods. **Biometrics**, Malden, v. 55, p. 1-12, 1999.

MCCOY, R.C.; TAYLOR, R.W.; BLAUWKAMP, T.A.; KELLEY, J.L.; KERTESZ, M.; PUSHKAREV, D.; PETROV, D.A.; FISTON-LAVIER, A.S. Illumina TruSeq synthetic long-reads empower *de novo* assembly and resolve complex, highly-repetitive transposable elements. **PLoS One**, San Francisco, v. 9, p. e106689, 2014.

MCHARDY, A.C.; RIGOUTSOS, I. What's in the mix: phylogenetic classification of metagenome sequence samples. **Current Opinion in Microbiology**, Kidlington, v. 10, p. 499-503, 2010.

MEDEMA, M.H.; BLIN, K.; CIMERMANCIC, P.; JAGER, V.; ZAKRZEWSKI, P.; FISCHBACH, M.A.; WEBER, T.; TAKANO, E.; BREITLING, R. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. **Nucleic Acids Research**, London, v. 39, p. W339-W346, 2011.

MEDINI, D.; DONATI, C.; TETTELIN, H.; MASIGNANI, V.; RAPPUOLI, R. The microbial pan-genome. **Current Opinion in Genetics & Development**, London, v. 15, p. 589-594, 2005.

MEEKS, J.C.; ELHAI, J.; THIEL, T.; POTTS, M.; LARIMER, F.; LAMERDIN, J.; PREDKI, P.; ATLAS, R. An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium. **Photosynthesis Research**, Dordrecht, v. 70, p. 85-106, 2001.

MESSING, J. New M13 vectors for cloning. **Methods in Enzymology**, New York, v. 101, p. 20-78, 1983.

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

MICALLEF, M.L.; D'AGOSTINO, P.M.; AL-SINAWI, B.; NEILAN, B.A.; MOFFITT, M.C. Exploring cyanobacterial genomes for natural product biosynthesis pathways. **Marine Genomics**, Amsterdam, v. 21, p.1-12, 2015.

MIHALI, T.K.; CARMICHAEL, W.W.; NEILAN, B.A. A putative gene cluster from a *Lyngbya wollei* bloom that encodes paralytic shellfish toxin biosynthesis. **PLoS One**, San Francisco, v. 6, 1-9, 2011.

MILLER, C.S.; BAKER, B.J.; THOMAS, B.C.; SINGER, S.W.; BANFIELD, J.F. EMIRGE: reconstruction of full-length ribosomal genes from microbial community short read sequencing data. **Genome Biology**, London, v. 12, p. R44, 2011.

MISHRA, S.; BHARGAVA, P.; ADHIKARY, S.P.; PRADEEP, A.; RAI, L.C. Wighted morphology: a new approach towards phylogenetic assessment of Nostocales (Cyanobacteria). **Protoplasma**, Lipzig, v. 252, p. 145-163, 2015.

MOMPER, L.M.; REESE, B.K.; CARVALHO, G.; LEE, P.; WEBB, E.A. A novel cohabitation between two diazotrophic cyanobacteria in the oligotrophic ocean. **The ISME Journal**, Oxon, v. 9, p. 882-893, 2015.

MORRIS, J.J.; KIRKEGAARD, R.; SZUL, M.J.; JOHNSON, Z.I.; ZINSER, E.R. Facilitation of robust growth of *Prochlorococcus* colonies and dilute liquid cultures by “helper” heterotrophic bacteria. **Applied and Environmental Microbiology**, Baltimore, v. 74, p. 4530-4534, 2008.

MOULIN, L.; MUNIVE, A.; DREYFUS, B.; BOIVIN-MASSON, C. Nodulation of legumes by members of the β -subclass of Proteobacteria. **Nature**, London, v. 411, p. 948-950, 2001.

MOUSTAFA, A.; LORAM, J.E.; HACKETT, J.D.; ANDERSON, D.M.; PLUMLEY, F.G.; BHATTACHARYA, D. Origin of saxitoxin genes in Cyanobacteria. **PLoS ONE**, San Francisco, v. 4, p. 1-10, 2009.

MULKIDJANIAN, A.Y.; KOONIN, E.V.; MAKAROVA, K.S.; MEKHEDOV, S.L.; SOROKIN, A.; WOLF, Y.I.; DUFRESNE, A.; PARTENSKY, F.; BURD, H.; KAZNADZEY, D.; HASELKORN, R.; GALPERIN, M.Y. The cyanobacterial genome core and the origin of photosynthesis. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v. 103, p. 13126-13131, 2006.

MURAKAMI, M.; SUN, Q.; ISHIDA, K.; MATSUDA, H.; OKINO, T.; YAMAGUCHI, K. Microviridins, elastase inhibitors from the cyanobacterium *Nostoc minutum*. **Phytochemistry**, Oxford, v. 45, p. 1197-1202, 1997.

NABOUT, J.C.; ROCHA, B.S.; CARNEIRO, F.M.; SANT'ANNA, C.L. How many species of cyanobacteria are there? Using a discovery curve to predict the species number. **Biodiversity and Conservation**, London, v. 22, p. 2907-2918, 2013.

NAGARKAR, S.; WILLIAMS, G.A.; SUBRAMANIAN, G.; SAHA, S.K. Cyanobacteria-dominated biofilms: a high quality food resource for intertidal grazers. **Hydrobiologia**, Dordrecht, v. 512, p. 89-95, 2004.

NEILAN, B.A.; JACOBS, D.; DEL DOT, T.; BLACKALL, L.L.; HAWKINS, P.R.; COX, P.T.; GOODMAN, A.E. rRNA sequences and evolutionary relationships among toxic and nontoxic cyanobacteria of the genus *Microcystis*. **International Journal of Systematic Bacteriology**, Washington, DC, v. 47, p. 693-697, 1997.

NIELSEN, H.B.; ALMEIDA, M.; JUNCKER, A.S.; RASMUSSEN, S.; LI, J.; SUNAGAWA, S.; PLICHTA, D.R.; GAUTIER, L.; PEDERSEN, A.G.; LE CHATELLIER, E.; PELLETIER, E.; BONDE, I.; NIELSEN, T.; MANICHANH, C.; ARUMUGAN, M.; BATTO, J.M.; SANTOS, M.B.Q.; BLOM, N.; BORRUEL, N.; BURGDORF, K.S.; BOUMEZBEUR, F.; CASELLAS, F.; DORÉ, J.; DWORZYNSKI, P.; GUARNER, F.; HANSEN, T.; HILDEBRAND, F.; KAAS, R.S.; KENNEDY, S.; KRISTIENSEN, K.; KULTIMA, J.R.; LÉONARD, P.; LEVENEZ, F.; LUND, O.; MOUMEN, B.; LE PASLIER, D.; PONS, N.; PEDERSEN, O.; PRIFTI, E.; QIN, J.; RAES, J.; SØRENSEN, S.; TAP, J.; TIMS, S.; USSERY, D.W.; YAMADA, T.; METAHIT CONSORTIUM, RENAULT, P.; SICHERITZ-PONTEN, T.; BORK, P.; WANG, J.; BRUNAK, S.; EHRlich, S.D. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. **Nature Biotechnology**, New York, v. 32, p. 822-828, 2014.

NOWRUZI, B.; KHAVARI-NEJAD, R.A.; SIVONEN, K.; KAZEMI, B.; NAJAFI, F.; NEJADSATTARI, T. A gene expression study on strains of *Nostoc* (Cyanobacteria) revealing antimicrobial activity under mixotrophic conditions. **African Journal of Biotechnology**, Nairobi, v. 11, p. 11296-11308, 2012.

NÜBEL, U.; GARCIA-PICHEL, F.; MUYZER, G. PCR primers to amplify 16S rRNA genes from cyanobacteria. **Applied and Environmental Microbiology**, Baltimore, v. 63, p. 3327-3332, 1997.

NUNNERY, J.K.; MEVERS, E.; GERWICK, W.H. Biologically Active Secondary Metabolites from Marine Cyanobacteria. **Current Opinion in Biotechnology**, Philadelphia, v. 21, p. 787-793, 2010.

OTSUKA, S.; ABE, Y.; FUKUI, R.; NISHIYAMA, M.; SENOO, K. Presence of previously undescribed bacterial taxa in non-axenic *Chlorella* cultures. **Journal of General Applied Microbiology**, Tokyo, v. 54, p. 187-193, 2008.

OURISSON, G. The general role of terpenes and their global significance. **Pure and Applied Chemistry**, Oxford, v. 62, p. 1401-1404, 1990.

OVERBEEK, R.; OLSON, R.; PUSCH, G.D.; OLSEN, G.J.; DAVIS, D.J.; DISZ, T.; EDWARDS, R.A.; GERDES, S.; PARRELLO, B.; SHUKLA, M.; VONSTEIN, V.; WATTAM, A.R.; XIA, F.; STEVENS, R. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). **Nucleic Acids Research**, London, v. 42, p. D206-D214, 2014.

PAERL, H.W.; PINCKNEY, J.L.; STEPPE, T.F. Cyanobacterial-bacterial mat consortia: examining the functional unit of microbial survival and growth in extreme environments. **Environmental Microbiology**, Oxford, v. 2, p. 11-26, 2000.

PALINSKA, K.A.; SUROSZ, W. Taxonomy of cyanobacteria: a contribution to consensus approach. **Hydrobiologia**, Dordrecht, v. 740, p. 1-11, 2014.

PANKE-BUISSE, K.; POOLE, A.C.; GOODRICH, J.K.; LEY, R.E.; KAO-KNIFFIN, J. Selection on soil microbiomes reveals reproducible impacts on plant function. **The ISME Journal**, Oxon, v. 9, p. 980-989, 2015.

PAPAEFTHIMIOU, K.; HROUZEK, P.; MUGNAL, M.A.; LUKESOVA, A.; TURICCHIA, S.; RASMUSSEN, U.; VENTURA, S. Differential patterns of evolution and distribution of the symbiotic behaviour in nostocacean cyanobacteria. **International Journal of Systematic and Evolutionary Microbiology**, Reading, v. 58, p. 553-564, 2008.

PAUNGFOO-LONHIENNE, C.; RENTSCH, D.; ROBATZEK, S.; WEBB, R.I.; SAGULENKO, E.; NÄSHOLM, R.; SCHMIDT, S.; LONHIENNE, T.G.A. Turning the table: plants consume microbes as a source of nutrients. **PLoS One**, San Francisco, v. 5, p. e11915, 2010.

PAWLOWSKI, K.; SIRRENBERG, A. Symbiosis between *Frankia* and actinorhizal plants: root nodules of non-legumes. **Indian Journal of Experimental Biology**, New Delhi, v. 41, p. 1165-1183, 2003.

PEARSON, L.; MIHALI, T.; MOFFITT, M.; KELLMANN, R.; NEILAN, B. On the chemistry, toxicology and genetics of the cyanobacterial toxins microcystin, nodularin, saxitoxin and cylindrospermopsin. **Marine Drugs**, Basel, v. 8, p. 1650-1680, 2010.

PENN, K.; WANG, J.; FERNANDO, S.C.; THOMPSON, J.R. Secondary metabolite gene expression and interplay of bacterial functions in a tropical freshwater cyanobacterial bloom. **The ISME Journal**, Oxon, v. 8, p. 1866-1878, 2014.

POTTS, M. Etymology of the genus name *Nostoc* (Cyanobacteria). **International Journal of Systematic Bacteriology**, Washington, DC, v. 47, p. 584, 1997.

POTTS, M. *Nostoc*. In: WHITTON, B.A.; POTTS, M. (Ed.). **The ecology of cyanobacteria**. Dordrecht: Kluwer Academic, 2000. p. 465-504.

QUICK, J.; QUINLAN, A.; LOMAN, N. A reference bacterial genome dataset generated on the MinION™ portable single-molecule nanopore sequencer. **GigaScience**, London, v. 3, p. 22, 2014.

RAINEY, F.A.; STACKEBRANDT, E. Transfer of the type species of the genus *Thermobacteroides* to the genus *Thermoanaerobacter* as *Thermoanaerobacter acetoethylicus* (Ben-Basst and Zeikus 1981) comb. nov., description of *Coprothermobacter* gen. nov., and reclassification of *Thermobacteroides proteolyticus* as *Coprothermobacter proteolyticus* (Ollivier et al. 1985) comb. nov. **International Journal of Systematic Bacteriology**, Washington, DC, v. 43, p. 857-859, 1993.

RAJANIEMI, P.; HROUZEK, P.; KAŠTOVSKÁ, K.; WILLAME, R.; RANTALA, A.; HOFFMANN, L.; KOMÁREK, J.; SIVONEN, K. Phylogenetic and morphological evaluation of the genera *Anabaena*, *Aphanizomenon*, *Trichormus* and *Nostoc* (Nostocales, Cyanobacteria). **International Journal of Systematic and Evolutionary Microbiology**, Reading, v. 55, p. 11-26, 2005.

RAMOS, R.T.J.; CARNEIRO, A.R.; SOARES, S.C.; BARBOSA, S.; VARUZZA, L.; ORABONA, G.; TAUCH, A.; AZEVEDO, V.; SCHNEIDER, M. P.; SILVA, A. High efficiency application of a mate-paired library from next-generation sequencing to postlight sequencing: *Corynebacterium pseudotuberculosis* as a case study for microbial *de novo* genome assembly. **Journal of Microbiological Methods**, Amsterdam, v. 95, p. 441-447, 2013.

RAN, L.; LARSSON, J.; VIGIL-STENMAN, T.; NYLANDER, J.A.; ININBERGS, K.; ZHENG, W.W.; LAPIDUS, A.; LOWRY, S.; HASELKORN, R.; BERGMAN, B. Genome erosion in a nitrogen-fixing vertically transmitted endosymbiotic multicellular cyanobacterium. **PLoS One**, San Francisco, v. 8, p. e11486, 2010.

RAPPÉ, M.S.; GIOVANNONI, S.J. The uncultured microbial majority. **Annual Review of Microbiology**, Palo Alto, v. 57, p. 369-394, 2003.

RASHIDAN, K.K.; BIRD, D.F. Role of predatory bacteria in the termination of a cyanobacterial bloom. **Microbial Ecology**, New York, v. 41, p. 97-105, 2001.

RATTRAY, J.E.; STROUS, M.; CAMP, H.J.O.; SCHOUTEN, S.; JETTEN, M.S.M.; DAMSTÉ, J.S.S. A comparative genomics study of genetic products potentially encoding ladderane lipid biosynthesis. **Biology Direct**, London, v. 4, p. 8, 2009.

ŘEHÁKOVÁ, K.; JOHANSEN, J.R.; CASAMATTA, D.A.; XUESONG, L.; VINCENT, J. Morphological and molecular characterization of selected desert soil cyanobacteria: three species new to science including *Mojavia pulchra* gen. et sp. nov. **Phycologia**, Odense, v. 46, p. 481-502, 2007.

RICHARDS, S. It's more than stamp collecting: how genome sequencing can unify biological research. **Trends in Genetics**, Cambridge, v. 31, p. 411-421, 2015.

RICKER, N.; QIAN, H.; FULTHORPE, R.R. The limitations of draft assemblies for understanding prokaryotic adaptation and evolution. **Genomics**, San Diego, v. 100, p. 167-175, 2012.

RIKKINEN, J. Relations between cyanobacterial symbionts in lichens and plants. In: PAWLOWSKI, K. (Ed.). **Microbiology monographs**: prokaryotic symbionts in plants. Berlin/Heidelberg: Springer-Verlag, 2009. v. 8. p. 265-270.

ROE, K.L.; BARBEAU, K.; MANN, E.L.; HAYGOOD, M.G. Acquisition of iron by *Trichodesmium* and associated bacteria in culture. **Environmental Microbiology**, Oxford, v. 14, p. 1681-1695, 2012.

ROHRLACK, T.; CHRISTOFFERSEN, K.; HANSEN, P.; ZHANG, W.; CZARNECKI, O.; HENNING, M.; FASTNER, J.; ERHARD, M.; NEILAN, B.A.; KAEBERNICK, M. Isolation, characterization, and quantitative analysis of microviridin J, a new *Microcystis* metabolite toxic to *Daphnia*. **Journal of Chemical Ecology**, New York, v. 29, p. 1757-1770, 2003.

ROHRLACK, T.; CHRISTOFFERSEN, K.; KAEBERNICK, M.; NEILAN, B.A. Cyanobacterial protease inhibitor microviridin J causes a lethal molting disruption in *Daphnia pulicaria*. **Applied and Environmental Microbiology**, Baltimore, v. 70, p. 5047-5050, 2004.

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

RUTHERFORD, K.; PARKHILL, J.; CROOK, J.; HORSNELL, T.; RICE, P.; RAJANDREAM, M.A.; BARRELL, B. Artemis: sequence visualization and annotation. **Bioinformatics**, Oxford, v. 16, p. 944-945, 2000.

SABATER, S.; MUÑOZ, I. *Nostoc verrucosum* (Cyanobacteria) colonized by a chironomid larva in a Mediterranean stream. **Journal of Phycology**, Baltimore, v. 36, p. 59-61, 2000.

SAITOU, N.; NEI, M. The neighbour-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, Chicago, v. 4, p. 406-425, 1987.

SAMBROOK, J.; RUSSEL, D.W. **Molecular cloning**: a laboratory manual. 3. ed. New York: Cold Spring Harbor Laboratory, 2001.

SÁNCHEZ, O.; DIESTRA, E.; ESTEVE, I.; MAS, J. Molecular characterization of an oil-degrading cyanobacterial consortium. **Microbial Ecology**, New York, v. 50, p. 580-588, 2005.

SAND-JENSEN, K. Ecophysiology of gelatinous *Nostoc* species: unprecedented slow growth and survival in resource-poor and harsh environments. **Annals of Botany**, London, v. 114, p. 17-33, 2014.

SANGER, F.; NICKLEN, S.; COULSON, A.R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v. 74, p. 5463-5467, 1977.

SCHLOSS, P.D.; HANDELSMAN, J. Status of the microbial census. **Microbiology and Molecular Biology Reviews**, Washington, DC, v. 68, p. 686-691, 2004.

SCHMIEDER, R.; EDWARDS, R. Quality control and preprocessing of metagenomic datasets. **Bioinformatics**, Oxford, 27, p. 863-864, 2011a.

SCHMIEDER, R.; EDWARDS, R. Fast identification and removal of sequence contamination from genomic and metagenomic datasets. **PLoS One**, San Francisco, v. 6, p. e17288, 2011b.

SCHNEIDER, G.F.; DEKKER, C. DNA sequencing with nanopores. **Nature Biotechnology**, New York, v. 30, p. 326-328, 2012.

SCHÜBLER, A.M. Molecular phylogeny, taxonomy and evolution of *Geosiphon pyriformis* and arbuscular mycorrhizal fungi. **Plant and Soil**, Dordrecht, v. 244, p. 75-83, 2002.

SCHWAB, W.; FUCHS, C.; HUANG, F.C. Transformation of terpenes into fine chemicals. **European Journal of Lipid Science and Technology**, Weinheim, v. 115, p. 3-8, 2012.

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

SENA, L.; ROJAS, D.; MONTIEL, E.; GONZÁLES, H.; MORET, J.; NARANJO, L. A strategy to obtain axenic cultures of *Arthrospira* spp. cyanobacteria. **World Journal of Microbiology and Biotechnology**, Oxford, v. 27, p. 1045-1053, 2008.

SHARMA, N.K. From natural to human-impacted ecosystems: rationale to investigate the impact of urbanization on cyanobacterial diversity in soils. **Biodiversity and Conservation**, London, v. 24, p. 1007-1015, 2015.

SHEN, B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. **Current Opinion in Chemical Biology**, London, v. 7, p. 285-295, 2003.

SHEN, H.; NIU, Y.; XIE, P.; TAO, M.; YANG, X. Morphological and physiological changes in *Microcystis aeruginosa* as a result of interactions with heterotrophic bacteria. **Freshwater Biology**, Oxford, v. 56, p. 1065-1080, 2011.

SHER, D.; THOMPSON, J.W.; KASHTAN, N.; CROAL, L.; CHISHOLM, S.W. Response of *Prochlorococcus* ecotypes to co-culture with diverse marine bacteria. **The ISME Journal**, Oxon, v. 5, p. 1125-1132, 2011.

SHI, T.; FALKOWSKI, P.G. Genome evolution in cyanobacteria: the stable core and the variable shell. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v. 105, p. 2510-2515, 2008.

SHI, L.; CAI, Y.; YANG, H.; XING, P.; LI, P.; KONG, L.; KONG, F. Phylogenetic diversity and specificity of bacteria associated with *Microcystis aeruginosa* and other cyanobacteria. **Journal of Environmental Sciences**, Amsterdam, v. 21, p. 1581-1590, 2009.

SHI, L.; CAI, Y.; KONG, F.; YU, Y. Changes in abundance and community structure of bacteria associated with buoyant *Microcystis* colonies during the decline of cyanobacterial bloom (autumn-winter transition). **International Journal of Limnology**, Les Ulis, v. 47, p. 355-362, 2011.

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

SHIRAISHI, H. Association of heterotrophic bacteria with aggregated *Arthrospira platensis* exopolysaccharides: implications in the induction of axenic cultures. **Bioscience, Biotechnology, and Biochemistry**, Tokyo, v. 79, p. 331-341, 2015.

SILVA, G.G.Z.; CUEVAS, D.A.; DUTILH, B.E.; EDWARDS, R.A. FOCUS: an alignment-free model to identify organisms in metagenomes using non-negative least squares. **PeerJ**, London, v. 2, p. e425, 2014.

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

SILVA-STENICO, M.E.; KANENO, R.; ZAMBUZI, F.A.; VAZ, M.G.M.V.; ALVARENGA, D.O.; FIORE, M.F. Natural products from cyanobacteria with antimicrobial and antitumor activity. **Current Pharmaceutical Biotechnology**, Hilversum, v. 14, p. 820-828, 2013a.

SILVA-STENICO, M.E.; LORENZI, A.S.; TESCHKE, O.; SILVA, C.S.P.; ETCHEGARAY, A.; FIORE, M.F. Antimicrobial cyanopeptide action on bacterial cells observed with atomic force microscopy. **Current Nanoscience**, Sharjah, v. 9, p. 141-148, 2013b.

SIMM, S.; KELLER, M.; SELYMESI, M.; SCHLEIFF, E. The composition of the global and feature specific cyanobacterial core-genomes. **Frontiers in Microbiology**, Lausanne, v. 6, p. 219, 2015.

SINGH, J.S.; PANDEY, V.C.; SINGH, D.P. Efficient soil microorganisms: a new dimension for sustainable agriculture and environmental development. **Agriculture, Ecosystems & Environment**, Amsterdam, v. 140, p. 339-353, 2011.

SINGH, P.; SINGH, S.S.; ABOAL, M.; MISHRA, A.K. Decoding cyanobacterial phylogeny and molecular evolution using an evonumeric approach. **Protoplasma**, Lipzig, v. 252, p. 519-535, 2015a.

SINGH, N.K.; DHAR, D.W.; TABASSUM, R. Role of cyanobacteria in crop protection. **Proceedings of the National Academy of Sciences of India Section B: Biological Sciences**, Allahabad, 2015. Disponível em: <<http://link.springer.com/article/10.1007%2Fs40011-014-0445-1>>. Acesso em: 27 ago. 2015.

SINGH, R.K.; TIWARI, S.P.; RAI, A.K.; MOHAPATRA, T.M. Cyanobacteria: an emerging source for drug discovery. **Journal of Antibiotics**, Tokyo, v. 64, p. 401-412, 2011.

SMITH, D.R. Death of the genome paper. **Frontiers in Genetics**, Lausanne, v. 4, p. 72, 2013.

SOGIN, M.L.; MORRISON, H.G.; HUBER, J.A.; WELCH, M.D.; HUSE, S.M.; NEAL, P.R.; ARRIETA, J.M.; HERNDL, G.J. Microbial diversity in the deep sea and the underexplored “rare biosphere”. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v. 103, p. 12115-12120, 2006.

SONG, H.; DING, M.Z.; JIA, X.Q.; MA, Q.; YUAN, Y.J. Synthetic microbial consortia: from systematic analysis to construction and applications. **Chemical Society Reviews**, London, v. 43, p. 6954-6981, 2014.

SOO, R.M.; SKENNERTON, C.T.; SEKIGUCHI, Y.; IMELFORT, M.; PAECH, S.J.; DENNIS, P.G.; STEEN, J.A.; PARKS, D.H.; TYSON, G.W.; HUGENHOLTZ, P. An expanded genomic representation of the phylum Cyanobacteria. **Genome Biology and Evolution**, Oxford, v. 6, p. 1031-1045, 2014.

STAMATAKIS, A. RaxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. **Bioinformatics**, Oxford, v. 22, p. 2688-2690, 2006.

STEVENSON, B.S.; SUFLITA, M.T.; STAMPS, B.W.; MOORE, E.R.B.; JOHNSON, C.N.; LAWSON, P.A. *Hoeflea anabaenae* sp. Nov., an epiphytic symbiont that attaches to the heterocytes of a strain of *Anabaena*. **International Journal of Systematic and Evolutionary Microbiology**, Reading, v. 61, p. 2439-2444, 2011.

STEWART, E.J. Growing unculturable bacteria. **Journal of Bacteriology**, Baltimore, v. 194, p. 4151-4160, 2012.

SUURNÄKKI, S.; GOMEZ-SAEZ, G.V.; RANTALA-YLINEN, A.; JOKELA, J.; FEWER, D.P.; SIVONEN, K. Identification of geosmin and 2-methylisoborneol in cyanobacteria and molecular detection methods for the producers of these compounds. **Water Research**, Oxford, v. 68, p. 56-66, 2015.

SVENNING, M.M.; ERIKSSON, T.; RASMUSSEN, U. Phylogeny of symbiotic cyanobacteria within the genus *Nostoc* based 16s rRNA sequence analyses. **Archives of Microbiology**, Heidelberg, v. 183, p. 19-26, 2005.

TAMURA, K.; PETERSON, D.; PETERSON, N.; STECHER, G.; NEI, M.; KHUMAR, S. MEGA5: Molecular evolutionary analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. **Molecular Biology and Evolution**, Chicago, v. 28, p. 2731-2739, 2011.

TANDISHABO, K.; NAKAMURA, K.; UMETSU, K.; TAKAMIZAWA, K. Distribution and role of *Coprothermobacter* spp. in anaerobic digesters. **Journal of Bioscience and Bioengineering**, Osaka, v. 114, p. 518-520, 2012.

TANG, K.W.; DZIALLAS, C.; GROSSART, H.P. Zooplankton and aggregates as refuge for aquatic bacteria: protection from UV, head and ozone stresses used for water treatment. **Environmental Microbiology**, Oxford, v. 13, p. 378-390, 2011.

TEELING, H.; GLÖCKNER, O. Current opportunities and challenges in microbial metagenome analysis – a bioinformatic perspective. **Briefings in Bioinformatics**, London, v. 13, p. 728-742, 2012.

TETTELIN, H.; MAGINANI, V.; CIESLEWICZ, M.J.; DONATI, C.; MEDINI, D.; WARD, N.L.; ANGIUOLI, S.V.; CRABTREE, J.; JONES, A.L.; DURKIN, A.S.; DEBOY, R.T.; DAVIDSEN, T.M.; MORA, M.; SCARSELLI, M.; MARGARIT Y ROS, I.; PETERSON, J.D.; HAUSER, C.R.; SUNDARAM, J.P.; NELSON, W.C.; MADUPU, R.; BRINKAC, L.M.; DODSON, R.J.; ROSOVITZ, M.J.; SULLIVAN, S.A.; DAUGHERTY, S.C.; HAFT, D.H.; SELENGUT, J.; GWINN, M.L.; ZHOU, L.; ZAFAR, N.; KHOURI, H.; RADUNE, D.; DIMITROV, G.; WATKINS, K.; O'CONNOR, K.J.B.; SMITH, S.; UTTERBACK, T.R.; WHITE, O.; RUBENS, C.E.; GRANDI, G.; MADOFF, L.C.; KASPER, D.L.; TELFORD, J.L.; WESSELS, M.R.; RAPPUOLI, R.; FRASER, C.M. Genome analysis of multiple pathogenic isolates of *Streptococcus agalctiae*: implications for the microbial “pan-genome”. **Proceedings of the National Academy of Sciences of the USA**, Washington, DC, v. 102, p. 13950-13955, 2005.

THOMAS, T.; GILBERT, J.; MEYER, F. Metagenomics – a guide from sampling to data analysis. **Microbial Informatics and Experimentation**, London, v. 2, p. 3, 2012.

THOMPSON, C.C.; AMARAL, G.R.; CAMPEÃO, M.; EDWARDS, R.A.; POLZ, M.E.; DUTILH, B.E.; USSERY, D.W.; SAWABE, T.; SWINGS, J.; THOMPSON, F.L. Microbial taxonomy in the post-genomic era: rebuilding from scratch? **Archives of Microbiology**, Heidelberg, v. 197, p. 359-370, 2015.

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

TUOMAINEN, J.; HIETANEN, S.; KUPARINEN, J.; MARTIKAINEN, P.J.; SERVOMAA, K. Community structure of the bacteria associated with *Nodularia* sp. (Cyanobacteria) aggregates in the Baltic Sea. **Microbial Ecology**, New York, v. 52, p. 513-522, 2006.

VANDAMME, P.; POT, B.; GILLIS, M.; DE VOS, P.; KERSTERS, K.; SWINGS, J. Polyphasic taxonomy, a consensus approach to bacterial systematics. **Microbiological Reviews**, Washington, DC, v. 60, n. 2, p. 407-438, 1996.

VARTOUKIAN, S.R.; PALMER, R.M.; WADE, W.G. Strategies for the culture of “unculturable” bacteria. **FEMS Microbiology Letters**, Amsterdam, v. 309, p. 1-7, 2010.

VÁZQUEZ-MARTÍNEZ, G.; RODRIGUEZ, M.H.; HERNÁNDEZ-HERNÁNDEZ, F.; IBARRA, J.E. Strategy to obtain axenic cultures from field-collected samples of the cyanobacterium *Phormidium animalis*. **Journal of Microbiological Methods**, Amsterdam, v. 57, p. 115-121, 2004.

VERNICOS, G.; MEDINI, D.; RILEY, D.R.; TETTELIN, H. Ten years of pan-genome analyses. **Current Opinion in Microbiology**, Kidlington, v. 23, p. 148-154, 2015.

WANG, H.; FEWER, D.; SIVONEN, K. Genome mining demonstrates the widespread occurrence of gene clusters encoding bacteriocins in cyanobacteria. **PLOS One**, San Francisco, v. 6, p. e22384, 2011.

WANG, H.; SIVONEN, K.; ROUHIAINEN, L.; FEWER, D.P.; LYRA, C.; RANTALAYLINEN, A.; VESTOLA, J.; JOKELA, J.; RANTASÄRKKÄ, K.; LI, Z.; LIU, B. Genome-derived insights into the biology of the hepatotoxic bloom-forming cyanobacterium *Anabaena* sp. strain 90. **BMC Genomics**, London, v. 13, p. 613, 2012.

WARD, N.; STALEY, J.T.; FUERST, J.A.; GIOVANNONI, S.; SCHLESNER, H.; STACKEBRANDT, E. The order Planctomycetales, including the genera *Planctomyces*, *Pirellula*, *Gemmata* and *Isosphaera* and the candidatus genera *Brocadia*, *Kuenenia* and *Scalindua*. In: DWORKIN, M.; FALKOW, S.; ROSENBERG, E.; SCHLEIFER, K.H.; STACKEBRANDT, E. (Ed.). **The prokaryotes** 7. 3. ed. New York: Springer, 2006. p. 757-793.

WATERBURY, J.B. The cyanobacteria – isolation, purification and identification. In: DWORKIN, M.; FALKOW, S.; ROSENBERG, E.; SCHLEIFER, K.H.; STACKEBRANDT, E. (Ed.). **The prokaryotes** 4. 3. ed. New York: Springer, 2006. p. 1053-1073.

WEBER, T.; BLIN, K.; DUDELA, S.; KRUG, D.; KIM, H.U.; BRUCCOLERI, R.; LEE, S.Y.; FISCHBACH, M.A.; MULLER, R.; WOHLLEBEN, W.; BREITLING, R.; TAKANO, E.; MEDEMA, M.H. antiSMASH 3.0: a comprehensive resource for the genome mining of biosynthetic gene clusters. **Nucleic Acids Research**, London, v. 43, p. W237-243, 2015.

WIEDNER, K.; SCHNEEWEISS, J.; DIPPOLD, M.A.; GLASER, B. Anthropogenic Dark Earth in Northern Germany – the Nordic analogue to terra preta de índio in Amazonia. **Catena**, Amsterdam, v. 132, p. 114-125, 2015.

WILLEY, J.M.; DONK, W.A. Lantibiotics: peptides of diverse structure and function. **Annual Review of Microbiology**, Palo Alto, v. 61, p. 477-501, 2007.

WILLIAMS, D.; TRIMBLE, W.L.; SHILTS, M.; MEYER, F.; OCHMAN, H. Rapid quantification of sequence repeats to resolve the size, structure and contents of bacterial genomes. **BMC Genomics**, London, v. 14, p. 537, 2013.

WILSON, M.C.; PIEL, J. Metagenomic approaches for exploiting uncultivated bacteria as a resource for novel biosynthetic enzymology. **Chemistry & Biology**, Cambridge, v. 20, p. 636-647, 2013.

WINTER, J.M.; BEHNKEN, S.; HERTWECK, C. Genomics-inspired discovery of natural products. **Current Opinion in Chemical Biology**, London, v. 15, p. 22-31, 2011.

WOOD, D.E.; SALZBERG, S.L. Kraken: ultrafast metagenomic sequence classification using exact alignments. **Genome Biology**, London, v. 15, p. R46, 2014.

WOODS, W. Apresentação. In: TEIXEIRA, W.G.; KERN, D.C.; MADARI, B.E.; LIMA, H.N.; WOODS, W. (Ed.). **As terras pretas de índio da Amazônia**. Manaus: Embrapa Amazônia Oriental, 2009. p. 10.

XIAO, J.; ZHANG, Z.; WU, J.; YU, J. A brief review of software tools for pangenomics. **Genomics Proteomics Bioinformatics**, Amsterdam, v. 13, p. 73-76, 2015.

YAMAGUCHI, H.; SUZUKI, S.; TANABE, Y.; OSANA, Y.; SHIMURA, Y.; ISHIDA, K.I.; KAWACHI, M. Complete genome sequence of *Microcystis aeruginosa* NIES-2549, a bloom-forming cyanobacterium from Lake Kasumigaura, Japan. **Genome Announcements**, Washington, DC, v. 3, p. e00551-15, 2015.

YAMAMURO, M. Importance of epiphytic cyanobacteria as food sources for heterotrophs in a tropical seagrass bed. **Coral Reefs**, Berlin, v. 18, p. 263-271, 1999.

YOKOTA, A. Cultivation of uncultured bacteria of the class Ktedonobacteria in the phylum Chloroflexi. **Makara Journal of Science**, Depok, v. 16, p. 1-8, 2012.

ZAKHIA, F.; JEDER, H.; WILLEMS, A.; GILLIS, M.; DREYFUS, B.; LAJUDIE, P. Diverse bacteria associated with root nodules of spontaneous legumes in Tunisia and first report for *nifH*-like gene within the genera *Microbacterium* and *Starkeya*. **Microbial Ecology**, New York, v. 51, p. 375-393, 2006.

ZANCHETT, G.; OLIVEIRA-FILHO, E.C. Cyanobacteria and cyanotoxins: from impacts on aquatic ecosystems and human health to anticarcinogenic effects. **Toxins**, Basel, v. 5, p. 1896-1917, 2013.

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

ZHANG, X.; DAVENPORT, K.W.; GU, W.; DALIGAULT, H.E.; MUNK, A.C.; TASHIMA, H.; REITENGA, K.; GREEN, L.D.; HAN, C.S. Improving genome assemblies by sequencing PCR products with PacBio. **BioTechniques**, Natick, v. 53, p. 61-62, 2012.

ZHAXYBAYEVA, O.; GOGARTEN, J.P.; CHARLEBOIS, R.L.; DOOLITTLE, F.; PAPKE, R.T. Phylogenetic analyses of cyanobacterial genomes: quantification of horizontal gene transfer events. **Genome Research**, Woodbury, v. 16, p. 1099-1108, 2006.

ZHOU, M.Y.; CLARK, S.E.; GOMEZ-SANCHEZ, C.E. Universal cloning method by TA strategy. **BioTechniques**, Natick, v. 19, p. 34-35, 1995.

ZHUBANOVA, A.A.; ERNAZAROVA, A.K.; KAIYRMANOVA, G.K.; ZAYADAN, B.K.; SAVITSKAYA, I.S.; ABDIEVA, G.Z.; KISTAUBAEVA, A.S.; AKIMBEKOV, N.S. Construction of cyanobacterial-bacterial consortium on the basis of axenic cyanobacterial cultures and heterotrophic bacteria cultures for bioremediation of oil-contaminated soils and water ponds. **Russian Journal of Plant Physiology**, Birmingham, v. 60, p. 555-562, 2013.

ZIEMERT, N.; ISHIDA, K.; WEIZ, A.; HERTWECK, C.; DITTMANN, E. Exploiting the natural diversity of microviridin gene clusters for discovery of novel tricyclic depsipeptides. **Applied and Environmental Microbiology**, Baltimore, v. 76, p. 3568-3574, 2010.

ZUKER, M. Mfold web server for nucleic acid folding and hybridization prediction. **Nucleic Acids Research**, London, v. 31, p. 3406-3415, 2003.

ZUKER, M.; STIEGLER, P. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. **Nucleic Acids Research**, London, v. 9, p. 133-148, 1981.