UNIVERSIDADE DE SÃO PAULO
CENTRO DE ENERGIA NUCLEAR NA AGRICULTURA

STELLA THOMAZ DE LIMA

Biossíntese do organofosforado anatoxina-a(s) pela cianobactéria *Sphaerospermopsis torques-reginae* ITEP-024

Piracicaba
2016

STELLA THOMAZ DE LIMA

Biossíntese do organofosforado anatoxina-a(s) pela cianobactéria *Sphaerospermopsis torques-reginae* ITEP-024

Versão revisada de acordo com a Resolução CoPGr 6018 de 2011

Dissertação apresentada ao Centro de Energia Nuclear na Agricultura da Universidde de São Paulo para obtenção do título de Mestre em Ciências

Área de Concentração: Biologia na Agricultura e no Ambiente

Orientador: Profa. Dra. Marli de Fátima Fiore

Piracicaba

2016

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

Lima, Stella Thomaz de
Biossíntese do organofosforado anatoxina-a(s) pela cianobactéria Sphaerospermopsis torques-reginae ITEP-024 / Stella Thomaz de Lima; orientadora Marli de Fátima Fiore. - - versão revisada de acordo com a Resolução CoPGr 6018 de 2011. - - Piracicaba, 2016.
92 f.: il.

Dissertação (Mestrado – 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. Cianotoxina 2. Genoma 3. Inibidores de proteases 4. Metabolismo secundário 5. Sequenciamento I. Título

CDU 579.22 : 577.2

Dedico

Aos meus pais, Claudia e Nivaldo, pelo suporte e amor incondicional.

Ofereço

Ao meu marido, Fernando, pelo amor, cumplicidade, compreensão, paciência e companheirismo.

AGRADECIMENTOS

Aos meus pais pelo apoio, investimento e suporte na minha formação. Sem vocês eu não chegaria até aqui.

À Profa. Dra. Marli de Fátima Fiore, pela orientação e ensinamentos. Obrigada pela oportunidade concedida e pela confiança depositada.

Ao meu marido pelo carinho e amor de todos os dias.

Ao meu irmão, Nicolas, que me inspira a ser melhor a cada dia.

Aos queridos, Rosangela e André, que sempre me apoiaram.

Aos meus avós Maria (in memorian), Paulo (in memorian) e Diva (in memorian) pela influência na minha formação e pelo amor sempre presente. Saudades.

Ao Programa de Pós-Graduação do Centro de Energia Nuclear na Agricultura da Universidade de São Paulo, pela estrutura oferecida.

Ao Prof. Dr. Augusto Etchegaray pela enorme contribuição, dedicação e discussões tão produtivas durante a execução do trabalho.

Ao Prof. Dr. Alessandro de Mello Varani e Prof. Dr. Ernani Pinto pela colaboração.

Aos companheiros de laboratório Karina, Marcelo, Diego, Talita, Rafael, Paula, Bruno W., Janaína e Vinícius.

Aos amigos Bruno C. e Ana Paula e Gabriela, agradeço as risadas, companhias durante o café, conversar e ajuda em relação a dissertação, mas em especial agradeço a amizade de todos os dias.

Ao Danillo pela ajuda com a pesquisa, orientação e ensinamentos de genômica, pelas discussões sempre produtivas e pela amizade.

A todos os meus amigos queridos de Americana, meu mestre e amigo Marcio, obrigada pelo apoio de sempre.

Aos meus padrinhos Rosangela e Marcos, que mesmo longe sempre me apoiaram e acreditaram no meu potencial.

Aos meus sogros Angela e Valmir, aos meus cunhados Amanda, Raul, Alessandra, Patrícia, Henrique, Ana Flávia e Diego, pelo apoio.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro.

À Fundação de Amparo à Pesquisa do Estado de São Paulo pelo apoio financeiro.

Aos funcionários do CENA em especial a Ana Paula do Lab. de Biologia Celular e Molecular Neuda, e ao Fábio e a Neuda da Secretaria de Pós-Graduação, agradeço todo o apoio e ajuda nos momentos solicitados.

À Bibliotecária Chefe Marília e Adriana pela ajuda prestada.

A todos que de alguma forma contribuíram para a realização deste trabalho.

Meus sinceros agradecimentos!

“Quem conheceu a alegria da compreensão
Conquistou um amigo infalível para a vida.
O pensar é para o homem, o que é voar para os pássaros.
Não tome como exemplo uma galinha quando podes ser uma cotovia”
(Albert Einstein)

RESUMO

LIMA, S. T. **Biossíntese do organofosforado anatoxina-a(s) pela cianobactéria *Sphaerospermopsis torques-reginae* ITEP-024**. 2016. 92 f. Dissertação (Mestrado) – Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Piracicaba, 2016.

As cianobactérias são fonte abundante de produtos naturais com diversos arranjos de classes estruturais, incluindo peptídeos, policetídeos, alcaloides, lipídeos e terpenos. Apesar das suas funções indefinidas na natureza, muitas destas moléculas são tóxicas para eucariotos ou possuem atividade farmacológica ou biológica com potencial de utilização como fármacos. A descoberta desses produtos naturais tem sido acelerada com o avanço dos esforços de sequenciamento de genomas cianobacterianos, aproveitando a homologia de sequência entre os genes que codificam peptídeos precursores e abordagens de mineração genômica. Neste trabalho, o genoma da linhagem *Sphaerospermopsis torques-reginae* ITEP-024 produtora da anatoxina-a (s) (ANS) foi sequenciado, montado e realizada a mineração do genoma buscando agrupamentos gênicos envolvidos na biossíntese da ANS e outros produtos naturais. O sequenciamento do genoma obtido utilizando uma plataforma Illumina MiSeq gerou 52.789.680 leituras, e estes dados brutos foram montados em 85 contigs gerando um genoma incompleto de tamanho 5.211.795 pb com cerca de 843 × de cobertura e teor de GC de 37,42%. A mineração dos dados do genoma, utilizando os softwares antiSMASH, RAST e Artemis, levou à identificação e proposição do agrupamento gênico da ANS sintetase. Além disso, as sequências de proteínas do agrupamento gênico ANS foram analisadas usando a ferramenta BLAST do NCBI e o programa Pfam, assim como a busca por domínios conservados no banco de dados do NCBI. O agrupamento gênico inédito da ANS possui oito genes (*ansA-H*) que codificam enzimas que participam diretamente da biossíntese, além da ORF1, responsável pela afinidade da rota biossintética por arginina, aminoácido precursor da síntese de ANS. A via biossintética de ANS proposta neste estudo envolve síntese peptídica ribossômica e é diferente de todas as outras vias biossintéticas de cianotoxinas já descritas. A mineração de dados do genoma da *S. torques-reginae* ITEP-024 também permitiu a identificação de dois agrupamentos gênicos envolvidos na produção de espumigina e anabenopeptina, dois potentes inibidores de proteases. O agrupamento gênico de espumigina é composto por dois genes que codificam uma enzima híbrida NRPS/PKS, e outros quatro genes que codificam uma proteína hipotética, uma desidrogenase, uma redutase e uma acetiltransferase. O agrupamento gênico de anabenopeptina é constituído por quatro genes que codificam enzimas NRPS, e dois genes que codificam uma isopropilmalato sintase e um transportador ABC. As sínteses dessas três moléculas bioactivas por *S. torques-reginae* ITEP-024 foram confirmadas por LC-MS/MS neste estudo e em anteriores. Este estudo contribui para a compreensão dos fatores genéticos, ecológicos e evolutivos relacionados à produção de ANS por cianobactérias e também para o desenvolvimento de métodos rápidos e sensíveis para a detecção e monitoramento de cianobactérias e suas toxinas em águas usadas para abastecimento público.

Palavras-chave: Genoma. Sequenciamento. Metabolismo secundário. Cianotoxina. Inibidores de proteases.

ABSTRACT

LIMA, S. T. **Biosynthesis of the organophosphate anatoxin-a(s) by the cyanobacterium *Sphaerospermopsis torques-reginae* ITEP-024**. 2016. 92 f. Dissertação (Mestrado) – Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Piracicaba, 2016.

Cyanobacteria are prolific source of natural products with diverse array of structural classes, including peptides, polyketides, alkaloids, lipids, and terpenes. Despite of their undefined roles in nature, many of these molecules are toxic for eukaryotes or possess pharmacological or biological activity with potential application as drug candidates. The discovery of these natural products has been accelerated with the advance of cyanobacterial genome sequencing efforts by taking advantage of sequence homology among genes encoding precursor peptides and genomic mining approaches. In this work, the genome of the anatoxin-a(s)-producing *S. torques-reginae* strain ITEP-024 was sequenced and assembled, and genome mining was performed for the searching of gene clusters involved in the anatoxin-a(s) (ANS) biosynthesis and other natural products. The genome sequencing obtained using an Illumina MiSeq plataform generated 52,789,680 read, and this raw data was assembled into 85 contigs that gave a draft genome size of 5,211,795 bp with ca. 843 × coverage and GC content of 37.42 %. The genome mining data, using the softwares antiSMASH, RAST and Artemis, led to the identification and proposition of the ANS synthetase gene cluster. Furthermore, protein sequences of the ANS cluster were analyzed using BLAST tool of the NCBI and the Pfam software, as well as the search for conserved domains in the NCBI database. The unique ANS gene cluster has eight genes (*ansA-H*) encoding enzymes directly involved in the biosynthesis, in addition to ORF1 responsible for the affinity of the biosynthetic pathway for arginine, amino acid precursor of the ANS synthesis. The proposed ANS biosynthetic pathway in this study involves ribosomal peptide synthesis and is different from all other biosynthetic pathways already described for cyanotoxins. Genome mining data of the *S. torques-reginae* ITEP-024 also allowed the identification of two gene clusters involved in the production of spumigin and anabaenopeptin, two potent protease inhibitors. The spumigin gene cluster is composed of two genes that encode a NRPS/PKS hybrid enzyme, and four other genes encoding a hypothetical protein, a dehydrogenase a reductase and an acetyltransferase. The anabaenopeptin gene cluster consists of four genes encoding NRPS enzymes, and two genes encoding an isopropylmalate synthase and an ABC transporter. The syntheses of these three bioactive molecules by the *S. torques-reginae* ITEP-024 were confirmed by LC-MS/MS in this study and in previous ones. This study contributes to the understanding of the genetic, ecological and evolutionary factors related to ANS production by cyanobacteria and also for the development of rapid and sensitive methods for the detection and monitoring of cyanobacteria and their toxins in water used for public supply.

Keywords: Genome. Sequencing. Secondary metabolism. Cyanotoxin. Protease inhibitors.

LISTA DE FIGURAS

Figura 1 - Gráfico em barras que representa as bioatividades de produtos naturais de cianobactérias; abreviação: RMD, resistência multidrogas (adaptado de DITTIMANN et al., 2015).....22

Figura 2 - Representação esquemática dos domínios enzimáticos de: A) NRPS e B) PKS. Abreviações: C: domínio de condensação; A: domínio de adenilação; PCP: proteína carreadora de peptidil; MT: metiltransferase; E: epimerase; AT: aciltransferase; ACP: proteína carreadora de acila; KS: cetossintase; KR: cetorredutase; DH: desidratase; ER: enoil-redutase; TE: tioesterase (adaptado de KEHR et al., 2011)..... 28

Figura 3 - Estrutura química da anatoxina-a(s) (adaptado de DÖRR et al., 2010) 30

Figura 4 - Etapas da via biossintética hipotética proposta para anatoxina-a(s) por Moore e colaboradores (Adaptado de MOORE et al., 1993)..... 32

Figura 5 - Mecanismo proposto de ação da ANS. A) Representa a ligação covalente que ocorre entre a molécula de ANS e o sítio ativo da enzima acetilcolinesterase. B) Representação esquemática do processo biológico de inibição irreversível da acetilcolinesterase pela cianotoxina ANS (adaptado de VALERIO et al., 2010) 33

Figura 6 - Estrutura química geral da espumigina. R1 e R2 ilustram os substituintes das variantes nas respectivas posições. As variações na estrutura são mostradas esquematicamente de acordo com suas posições. Abreviações: Hpla, ácido hidroxifenilacético; Hpla(Ac), ácido acetil hidroxifenilacético; Hty, homotirosina; Leu, leucina; Tyr, tirosina; Me-Pro, metil prolina; Pro, prolina; Argol, argininol; Argal, argininal; Arg, arginina; Me-Lys, metil lisina; Agm, agmatina.....35

Figura 7 - Estrutura química da anabenopeptina encontrada na linhagem *Sphaerospermopsis torques-reginae* ITEP-024; abreviações: Ile, isoleucina; Leu, leucina; Hty, homotirosina; MeAla, metil alanina; Phe, fenilalanina 37

Figura 8 - Porcentagem de projetos de sequenciamento de genomas cianobacterianos em relação aos outros filos bacterianos disponível no “Genomes Online Database” (GOLD) em 15/02/2016. Fonte: http://genomesonline.org/cgi-bin/GOLD/index.cgi 38

Figura 9 - Qualidade de leituras obtidas em sequenciamento na plafaforma MiSeq (Illumina) do genoma da cianobactéria *S. torques-reginae* ITEP-024. (A) e (B): conjunto dos pares R1 e R2 (respectivamente).....53

Figura 10 – Representação esquemática proposta do agrupamento gênico da anatoxina-a(s)....57

Figura 11 - Proposta de via de biossíntese da anatoxina-a(s) da linhagem *S. torques-reginae* ITEP-024 baseada em enzimas preditas em seu genoma 61

Figura 12 - Representação esquemática do agrupamento gênico da espumigina. Abreviações: A: adenilação; KR: cetorredutase; PCP: proteína carreadora de peptidil; C: condensação; Tyr: tirosina; E: epimerização; Pro: prolina; Arg: arginina; TD: tioéster redutase; PH: proteína hipotética; DH: desidrogenase; AT: aciltransferase; RE: redutase 67

Figura 13 - A: Proposta de biossíntese da espumigina da linhagem *S. torques-reginae* ITEP-024; B: Estrutura da variante produzida pela linhagem *S. torques-reginae* ITEP-02 69

Figura 14 - Representação esquemática do agrupamento gênico de produção de anabenopeptina. Abreviações: Ile: isoleucina; PCP: proteína carreadora de peptidil; C: condensação; Lys: lisina; E: epimerização; Hty: homotirosina; MeAla: metilalanina; MT: metiltranferase; Phe: fenilalanina e TE: tioesterase; IMS: isopropilmalato sintase; ABCt: transportador ABC 71

Figura 15 - Árvore filogenética reconstruída por inferência bayesiana com sequências concatenadas de nucleotídeos de seis agrupamentos gênicos de anabenopeptina. A linhagem utilizada neste estudo está em negrito. Probabilidades posteriores estão apresentadas nos nós dos ramos 72

Figura 16 - A: Proposta de biossíntese da anabenopeptina da linhagem *S. torques-reginae* ITEP-024; B: Estrutura da variante produzida pela linhagem *S. torques-reginae* ITEP-024 74

Figura 17 - Análise por LC-MS/MS do extrato da cultura de *S. torques-reginae* ITEP-024. Espectro de íon produzido para anabenopeptina. O primeiro pico à direita representa o íon molecular com sinal de *m/z* de 808,4553 ($[M+H]^+$).....75

Figura 18 - Estrutura dos agrupamentos gênicos presumidamente relacionados à síntese de espumigina e anabenopeptina em Sphaerospermopsis torques-reginae ITEP-024 e em cianobactérias 77

LISTA DE TABELAS

Tabela 1 - Agrupamentos gênicos de produtos naturais de cianobactérias (adaptado de DITTIMANN et al., 2015) 24
Tabela 2 - Relação de genomas de cianobactérias da ordem Nostocales com sequenciamento completo e incompleto (NCBI, janeiro de 2016) 40
Tabela 3 - Locais de registro da espécie *S. torques-reginae* no Brasil e algumas características físico-químicas desses locais (adaptado de WERNER et al., 201 44
Tabela 4 - Agrupamentos gênicos preditos no genoma de *Sphaerospermopsis torques-reginae* ITEP-024 pelo programa antiSMASH 54
Tabela 5 - Comparação entre as sequências de proteínas da AnsA com CS2 e VioC 59
Tabela 6 - Funções deduzidas das ORFs do agrupamento gênico de anatoxina-a(s) da linhagem *S. torques-reginae* ITEP-024 63
Tabela 7 - Comparação de polipeptídeos codificados pelos agrupamentos gênicos de espumiginas baseada no comprimento da proteína codificada (aa) e similaridade da proteína com *S. torques reginae* ITEP-024 66
Tabela 8 - Funções propostas das proteínas codificadas pelo agrupamento gênico de biossíntese de espumigina na linhagem *S. torques-reginae* ITEP-024 e maiores valores de BLAST de proteína 67
Tabela 9 - Funções propostas das proteínas codificadas pelo agrupamento gênico da biossíntese da anabenopetina na linhagem *S. torques-reginae* ITEP-024 e maiores valores de BLAST de proteína 71
Tabela 10 - Comparação de polipeptídeos codificados pelos agrupamentos gênicos de anabenopeptinas, incluindo o comprimento da proteína codificada (aa) e similaridade da sequência proteica com a sequência equivalente encontrada em *S. torques reginae* ITEP-024 72

SUMÁRIO

1 INTRODUÇÃO.....19
2 REVISÃO DA LITERATURA 21
2.1 Biossíntese de Produtos Naturais Cianobacterianos.....21
2.1.1 A neurotoxina anatoxina-a(s) 28
2.1.2 Espumigina 33
2.1.3 Anabenopeptina 35
2.2 Genômica de cianobactérias 36
2.3 A espécie *Sphaerospermopsis torques-reginae* e sua distribuição no Brasil 42
3 HIPÓTESE.....45
4 OBJETIVO 46
5 MATERIAL E MÉTODOS 47
5.1 Origem da linhagem, condições de cultivo e produção de biomassa da *Sphaerospermopsis torques-reginae* ITEP-024.....47
5.2 Sequenciamento do genoma da *Sphaerospermopsis torques-reginae* ITEP-024.....47
5.2.1 Extração de DNA total 47
5.2.2 Construção da biblioteca genômica e sequenciamento 48
5.2.3 Processamento das sequências e montagem.....49
5.2.4 Anotação e detecção dos agrupamentos gênicos de anatoxina-a(s), espumigina e anabenopeptina 49
5.2.5 Análise filogenética de agrupamentos gênicos de anabenopeptinas 50
5.3 Análises em cromatografia líquida acoplada a espectrometria de massas (LC-MS/MS)...50
6 RESULTADOS E DISCUSSÕES 52
6.1 Sequenciamento, montagem e anotação genômica 52
6.2 Detecção de potenciais agrupamentos gênicos envolvidos na síntese de metabólitos secundários.....53
6.3 Identificação do agrupamento gênico envolvido na síntese da anatoxina-a(s) 54
6.4 Identificação do agrupamento gênico envolvido na síntese da espumigina.....65
6.5 Identificação do agrupamento gênico envolvido na síntese da anabenopeptina 69
6.6 Interação entre os agrupamentos gênicos de anabenopeptina e espumigina.....74
7 CONCLUSÕES.....77
REFERÊNCIAS 79

1 INTRODUÇÃO

A anatoxina-a(s) (ANS), diferentemente das outras neurotoxinas cianobacterianas (saxitoxinas, anatoxina-a e homoanatoxina-a), as quais são alcaloides, é um organofosforado natural, sendo que o “s” do nome deriva da excessiva salivação observada em camundongos expostos a esta cianotoxina (CARMICHAEL; GORHAM, 1978). Ela é o único organofosforado de origem microbiana conhecido. A ANS foi descoberta na linhagem *Anabaena flos-aquae* NRC 525-17, isolada de uma floração ocorrida em 1965 na lagoa Bufallo, província de Saskatchewan, Canadá (CARMICHAEL; GORHAM, 1978). A sua estrutura química foi descrita como um éster metílico de uma N-hidroxiguanidina cíclica ($C_7H_{17}N_4O_4P$) de massa molecular 252 Da (MATSUNAGA et al., 1989). Estudos farmacológicos demonstraram que animais expostos a ANS apresentaram sintomas característicos de uma excessiva estimulação colinérgica, como salivação, lacrimejamento, incontinência urinária, fasciculação muscular e paralisia respiratória (MAHMOOD; CARMICHAEL, 1986). Posteriormente, foi demonstrado que ela é uma potente e irreversível inibidora das acetilcolinesterases atuando da mesma forma que os inseticidas organofosforados paraoxon, fisostigmina, piridostigmina e a arma química sarin (COOK et al., 1988; MAHMOOD; CARMICHAEL, 1987; PETRIKOVICS et al., 2004; PITA et al., 2003; TUOVINEN et al., 1999; VAN APELDOORN, 2007). A dose letal média (DL_{50}) reconhecida pela Organização Mundial da Saúde (WHO) para ANS é de 20 $\mu g \cdot kg^{-1}$ quando administrada via intraperitoneal em camundongos (CHORUS; BARTRAM, 1999).

A ocorrência de ANS foi relatada em alguns países como Canadá, Dinamarca, Estados Unidos e Brasil. A linhagem produtora *Anabaena lemmermannii* PH-160, isolada de um lago na Dinamarca, foi implicada na morte de pássaros (HENRIKSEN et al., 1997; ONODERA et al., 1997). Nos Estados Unidos, toxicose associada à ANS foi relatada em cães em Dakota do Sul (MAHMOOD et al. 1988) e em porcos e patos em Illinois (COOK et al. 1988). No Brasil, a atividade de inibição da acetilcolinesterase foi observada em extratos de *S. torques-reginae* linhagens ITEP-024, ITEP-025 e ITEP-026, isoladas de florações do reservatório Tapacurá em Recife, PE (MOLICA et al., 2005). Posteriormente, a produção de ANS por essas três linhagens de *S. torques-reginae* foi confirmada por espectrometria de massas (DÖRR et al., 2010). Amostras de florações com dominância de *Anabaena spiroides*, ocorrida em um lago ornamental da Universidade Federal do Rio Grande, RS, e *Anabaena crassa*, ocorrida no reservatório Faxinal, Caxias do Sul, RS, foram positivas em bioensaios e testes de inibição da atividade da acetilcolinesterase para ANS (BECKER et al., 2010; MONSERRAT et al. 2001).

Estudos sobre ocorrência de *Sphaerospermopsis* e *Anabaena* em corpos d'água brasileiros mostraram que estes gêneros, potencialmente produtores de ANS, são bastante comuns (HONDA, 2009; WERNER; LAUGHINGHOUSE IV, 2009; WERNER et al., 2012). A portaria MS 2914/11, que trata da normatização da qualidade da água para consumo humano do Ministério da Saúde, recomenda a análise da presença da ANS em águas usadas para consumo humano. Assim, métodos rápidos e sensíveis para a detecção e monitoramento de cianobactérias e suas toxinas são de fundamental interesse para os órgãos públicos, principalmente aqueles responsáveis pela regulação da qualidade de água no país, uma vez que efeitos indesejáveis ou fatais podem ser causados pela presença dessas substâncias na água.

Além da produção de toxinas, as cianobactérias apresentam grande potencial para a produção de outras moléculas bioativas (KEHR; PICCHI; DITTMANN, 2011). Anabenopeptinas e espumiginas são exemplos de produtos naturais sintetizados por cianobactérias. Anabenopeptinas é uma família de hexapeptídeos cíclicos altamente diversificados que apresentam atividade antiproteases e a característica estrutural de uma ligação ureído a qual liga o aminoácido D-lisina do ciclo à cadeia lateral (ROUHIAINEN et al., 2010; CHRISTIANSEN et al., 2011). Espumiginas também são inibidores de proteases tetrapeptídicos lineares que contém na parcela N-terminal um ácido D-hidroxifenilacético, geralmente metil prolina e um derivado aldeído de arginina na porção C-terminal (FEWER et al., 2009; FUJII et al., 1997). Esses produtos naturais possuem grande potencial biotecnológico podendo ser amplamente utilizados em fármacos antirretrovirais, usados, por exemplo, na terapia combinada para o tratamento de infecções pelo vírus HIV, além do tratamento de doenças gástricas e do sistema cardiovascular (MACHADO et al., 2014).

A espécie cianobactéria estudada neste trabalho, a *Sphaerospermopsis torques-reginae*, é conhecida como produtora de ANS (MOLICA et al., 2005; WERNER et al., 2012; DÖRR et al., 2010) e também de espumigina (SANZ et al., 2015). Entretanto, o mecanismo molecular da biossíntese de ANS é desconhecido e o de espumigina não foi ainda descrito para *Sphaerospermopsis*. A obtenção desses conhecimentos pode ser alcançada com o sequenciamento do genoma da *S. torques-reginae*, que representará uma base rica de informação para uma melhor compreensão da biologia, ecologia e evolução dessa espécie, bem como da síntese de ANS, espumigina e anabenopeptina. A descoberta dos genes envolvidos na síntese da toxina ANS permitirá utilizar métodos moleculares para detecção desses genes em amostras de água e determinar os fatores ambientais que regulam a produção de ANS e, portanto, os fatores que afetam a qualidade da água utilizada para consumo e lazer.

2 REVISÃO DA LITERATURA

2.1 Biossíntese de Produtos Naturais Cianobacterianos

Cianobactérias são encontradas em diversos ecossistemas e desempenham um papel fundamental nos ciclos biogeoquímicos da Terra (WHITTON; POTTS, 2000). Provavelmente, devido à sua alta plasticidade funcional esses organismos são produtores de inúmeras substâncias de interesse biotecnológico e farmacológico. Cerca de 1.100 produtos naturais de diferentes estruturas químicas estão distribuídos entre 39 gêneros de cianobactérias e incluídos nas classes de peptídeos, policetídeos, alcaloides, lipídeos e terpenos (DITTIMANN et al., 2015). Dentre essas moléculas cianobacterianas, muitas possuem estruturas complexas, raras e com modificações não usuais. Avanços na exploração de genomas de cianobactérias têm possibilitado a compreensão da lógica de produção da síntese e da maquinaria bioquímica envolvida na síntese desses produtos naturais, no entanto, muitas moléculas ainda apresentam síntese desconhecida ou não foram testadas, ou até mesmo não detectadas (Figura 1).

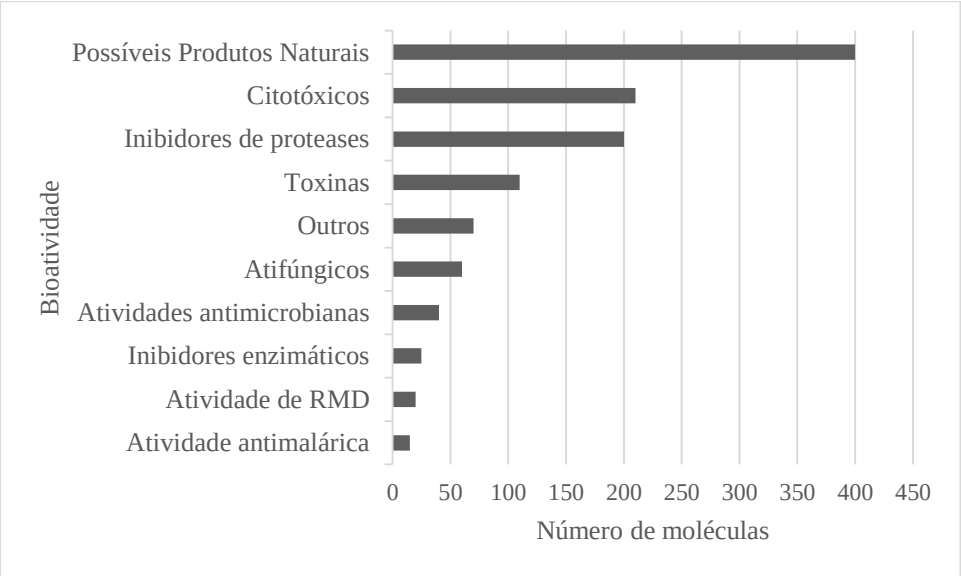


Figura 1 – Gráfico em barras que representa as bioatividades de produtos naturais de cianobactérias; abreviação: RMD, resistência multidrogas (adaptado de DITTIMANN et al., 2015)

As toxinas produzidas por cianobactérias são um exemplo de moléculas bem exploradas dentro do filo. Estas são classificadas de acordo com seus efeitos biológicos: hepatotoxinas (ex.:

microcistinas e nodularinas), neurotoxinas (ex.: saxitoxinas, anatoxina-a, anatoxina-a(s) e homoanatoxina-a), citotoxinas (ex.: cilindropermopsina), toxinas irritantes e gastrointestinais (ex.: aplisiatoxina, debromoaplisiatoxina e limbiatoxina), endotoxinas lipopolissacarídicas, e outras cujos perfis de ação não são ainda completamente conhecidos (microviridinas J, por exemplo) (FIORE et al., 2011). Por muito tempo os estudos de produtos naturais de cianobactérias foram focados nessas toxinas, especialmente na hepatotoxina microcistina (DITTMANN; WIEGAND, 2006). No entanto, um número promissor de substâncias com atividades anticâncer, antiproteases, antifúngica, antiviral, antibacteriana, anti-inflamatória, além de atividade de reversão de drogas multirresistentes, tiveram seus agrupamentos gênicos descobertos em cianobactérias (Tabela 1) (DITIMANN et al., 2015).

Tabela 1 - Agrupamentos gênicos de produtos naturais de cianobactérias (adaptado de DITTIMANN et al., 2015)

Moléculas	**Agrupamentos gênicos**	**Linhagem Produtora**	**Categoria**	**Tamanho (kb)**	**Ano de descoberta**
Apratoxina	-	*Lyngbya bouillonii*	NRPS-PKS	57,4	2011
Barbamida	AF516145	*Lyngbya majuscula* 19L	NRPS-PKS	26	2002
Criptoficina	EF159954	*Nostoc* sp. ATCC 53789	NRPS-PKS	40,3	2006
Curacina A	AY652953	*Lyngbya majuscula* 19L	NRPS-PKS	64	2004
Cilindrospermopsina	EU140798	*Cylindrospermopsis raciborskii* AWT205	NRPS-PKS	43	2008
Hectoclorina	AY974560	*Lyngbya majuscula* JHB	NRPS-PKS	38	2007
Jamaicamida	AY522504	*Lyngbya majuscula* JHB	NRPS-PKS	58	2004
Microcistina	AF183408	*Microcystis aeruginosa* PCC 7806	NRPS-PKS	55	2000
Nodularina	AY210783	*Nodularia spumigena* NSOR10	NRPS-PKS	48	2004
Nostociclopeptídeo	AY167420	*Nostoc* sp. ATCC53789	NRPS-PKS	33	2004
Nostoficina	JF430079	*Nostoc* sp. 152	NRPS-PKS	45,1	2011
Puwainaficinas	KM078884	*Cylindrospermum alatosporum* CCALA 988	NRPS-PKS	56,7	2014
Aeruginosina	AM778955	*Microcystis aeruginosa* PCC 7806	NRPS	27	2009
Anabaenopeptilida	AJ269505	*Anabaena* sp. 90	NRPS	29	2000
Anabenopeptina	GU174493	*Anabaena* sp. 90	NRPS	32	2010
Hassallidinas	KJ502174	*Anabaena* sp. SYKE 748A	NRPS	59	2014
Lingbiatoxina	AY588942	*Lyngbya majuscula*	NRPS	11,3	2004
Microginina	AM990464	*Planktothrix agardhii* NIVA-CYA 126/8	NRPS	20,5	2009
Nostopeptolida	AF204805	*Nostoc* sp. GSV224	NRPS	40	2003
Espumigina	CP007203	*Nodularia spumigena* CCY 9414	NRPS	21	2009
Anatoxina-a	FJ477836	*Oscillatoria* sp. PCC 6506	PKS	29	2009
Nosperina	JQ975876	*Nostoc* sp.	PKS	22,4	2013
Aeruginosamida	CAIH01000183	*Microcystis aeruginosa* PCC 9432	Cianobactina	12,4	2013
Aestuaramida	AAVU01000047	*Lyngbya aestuarii* PCC 8106	Cianobactina	14,8	2014
Anaciclamida	FJ461733	*Anabaena* sp. 90	Cianobactina	11	2010
Cianothecamida	CP001344	*Cyanothece sp.* PCC 7425	Cianobactina	16,5	2012
Microciclamida	AM931579	*Microcystis aeruginosa* PCC 7806	Cianobactina	12,6	2008
Patellamida	AY986476	*Prochloron didemni*	Cianobactina	11	2005
Piricyclamida	JQ951924	*Microcystis aeruginosa* PCC 7005	Cianobactina	9,3	2012
Prenilagaramida	HQ655154	*Planktothrix agardhii* NIES-596	Cianobactina	13,4	2011

continua

Moléculas	Agrupamentos gênicos	Linhagem Produtora	Categoria	Tamanho (kb)	Ano de descoberta
Tenueciclamida	EU290741	*Nostoc spongiaeforme var. tenue str. Carmeli*	Cianobactina	11,8	2008
Tricamida	CP000393	*Trichodesmium erythraeum* ISM101	Cianobactina	12,5	2006
Truncamida	EU290742	*Prochloron sp.* 06037A	Cianobactina	11,1	2008
Viridisamida	CP003614	*Oscillatoria nigro-viridis* PCC 7112	Cianobactina	13	2013
Microviridina	AM943877	*Microcystis aeruginosa* NIES-298	Microviridina	5,8	2008
Proclorosina	BX548175	*Prochlorococcus marinus* MIT9313	Lantipeptídeo	P	2010
Scitonemina	CP001037	*Nostoc punctiforme* ATCC 29133	Indole-alcaloide	29,6	2007
Saxitoxina	DQ787200	*Cylindrospermopsis raciborskii* T3	Alcaloide	37,6	2008
Geosmina	CP001037	*Nostoc punctiforme* PCC 73102	Terpeno	2,3	2009
2-Metilisoborneol	HQ830028	*Pseudanabaena sp.* dqh15	Terpeno	5,9	2011
Ambiguina	KF664586	*Fischerella ambigua* UTEX 1903	Indole-Terpeno	41,7	2014
Welwitindolinona	KF811479	*Hapalosiphon welwitschii* UTEX B 1830	Indole-Terpeno	36,6	2014
Welwitindolinona	KJ767018	*Westiella intricata* UH HT-29-1	Indole-Terpeno	59,2	2014
Micosporine-2-glcina	AB854644	*Aphanothece halophytica*	NRPS-like	4	2014
Shinorina	CP000117	*Anabaena variabilis* ATCC 29413	NRPS-like	6,8	2010
Cilindrociclofana	JX477167	*Cylindrospermum licheniforme* UTEX 'B 2014'	Ciclofano	26,2	2012

conclusão

A diversidade dos produtos naturais de cianobactéria é conferida pela diversidade das rotas biossintética e das estruturas químicas de seus produtos finais. Essas rotas podem variar amplamente, desde a utilização de sete domínios catalíticos para síntese da pequena molécula de lingbiatoxina, por exemplo, até o uso de 64 domínios catalíticos para a síntese de curacina, (DITTMANN et al., 2015). Existem duas vias de produção de substâncias em cianobactérias, a via ribossômica e a via não ribossômica.

A via ribossômica para a síntese proteica depende da informação contida nos códons do RNA mensageiro que se associa aos ribossomos, os quais consistem de RNA ribossômico e proteínas. O pareamento de bases dos códons no ribossomo resulta na captação de aminoácidos específicos associados/ligados a moléculas de RNA transportador e a formação do peptídeo se dá pela ligação peptídica que é catalisada entre os aminoácidos (ARNISON et al., 2013). Em cianobactérias a via ribossômica é responsável pela produção de produtos naturais como, microviridinas e cianobactinas, e está limitada a 20 aminoácidos proteinogênicos conhecidos, no entanto, modificações pós-traducionais conferem a essas moléculas uma grande diversidade química e estrutural em relação aos peptídeos ribosomais típicos. As cianobactinas, por exemplo, constituem uma ampla família de peptídeos e a característica de como o peptídeo central é sintetizado por essas vias, atua como um guia na busca por novos produtos naturais, já que o tipo de modificação pós-traducional que ocorre (por exemplo, heterociclização, oxidação, metilação, prenilação) promove uma assinatura no produto final (DITTIMANN et al., 2015). Esses peptídeos ribossomais exibem uma grande variedade de conformações estruturais e considerável potencial bioativo (OMAN; VAN DER DONK, 2010). Os atuais esforços nos estudos de mineração genômica apontam que uma parte substancial dos produtos naturais cianobacterianos é produzida por agrupamentos gênicos de biossíntese de peptídeos ribossomais (DITTIMANN et al., 2015).

A via não ribossômica de produção de produtos naturais é uma via biossintética bastante utilizada por micro-organismos na produção de antibióticos e toxinas (WELKER; VON DÖHREN, 2006; JONES et al., 2010). A síntese não ribossômica utiliza uma grande variedade de substratos, muitos dos quais são aminoácidos não proteicos, hidroxiácidos e substâncias policetônicas, especialmente elaborados para serem incorporados na estrutura peptídica. Portanto, na estrutura de uma mesma toxina, por exemplo, observa-se ligações peptídicas, funções cetônicas, insaturações e funções aromáticas. A biossíntese de metabólitos não ribossomais é catalisada por enzimas conhecidas como peptídeo sintetase não ribossômica (NRPS, do inglês “non-ribosomal peptide synthetase”) e policetídeo sintase (PKS, do inglês “polyketide synthase”).

A síntese utilizando as enzimas PKS e NRPS, diferentemente da síntese ribossômica, pode envolver cerca de 300 substratos gerando uma grande diversidade de estruturas (FINKING; MARAHIEL, 2004; KEHR et al., 2011). As NRPS possuem uma estrutura modular (Figura 2A), sendo que cada módulo é responsável pela incorporação ou modificação do aminoácido (MARAHIEL et al., 1997; VON DÖHREN et al., 1997). Estes módulos são constituídos por diferentes domínios, entre eles o domínio de adenilação (A), responsável pela seleção e ativação do aminoácido que será incorporado, ou seja, controla a sequência primária do peptídeo final; o domínio proteína carreadora de peptidil (PCP do inglês *Peptidyl Carrier Protein*), responsável pelo transporte dos aminoácidos de um centro catalítico para o próximo; o domínio de condensação (C), responsável pela formação da ligação peptídica, sendo, por isso, o módulo central da síntese não ribossômica por garantir o alongamento da cadeia polipeptídica; e o domínio de tioesterase (TE), que é importante para a liberação do produto final (FINKING; MARAHIEL, 2004; BARRIOS-LLERENA et al., 2007). Porém, outros domínios também podem fazer parte da estrutura de uma NRPS, entre eles o metiltransferase (MT), epimerase (E) e redutase (Re) (ROUNGE et al., 2009), ou, ainda, os módulos podem conter apenas um domínio (WELKER; VON DOHREN, 2006).

As enzimas PKS, assim como as NRPS, são enzimas multifuncionais organizadas em módulos (Figura 2B), porém utilizam diferentes tipos de iniciadores e alongadores de cadeia, os ácidos carboxílicos, para a síntese da molécula final. As PKSs podem ser divididas em três principais tipos: I) Enzimas multifuncionais compostas por diversos módulos que atuam uma única vez, ou seja, realizam síntese linear; II) Grandes complexos enzimáticos com um único sítio ativo cada, que possuem funções específicas e uma síntese interativa, podendo atuar mais de uma vez; III) Enzimas com vários módulos de síntese interativa que possuem a função de catalisar reações específicas como iniciação, extensão e ciclização (YU et al., 2012). Normalmente, são formadas pelos domínios cetossintase (KS, do inglês *ketosynthase*), aciltransferase (AT, do inglês *acyltransferase*) e tioesterase (T, do inglês *thioesterase*), por fim, proteína carreadora de acila (ACP, do inglês *acyl carrier protein*), que é o domínio responsável pelo transporte entre os sítios catalíticos. Outros domínios que podem ser comumente encontrados em uma PKS são cetorredutase (KR, do inglês *ketoreductase*), desidratase (DH, do inglês *dehydratase*), enoil-redutase (ER, do inglês *enoyl reductase*) (KEHR et al., 2011).

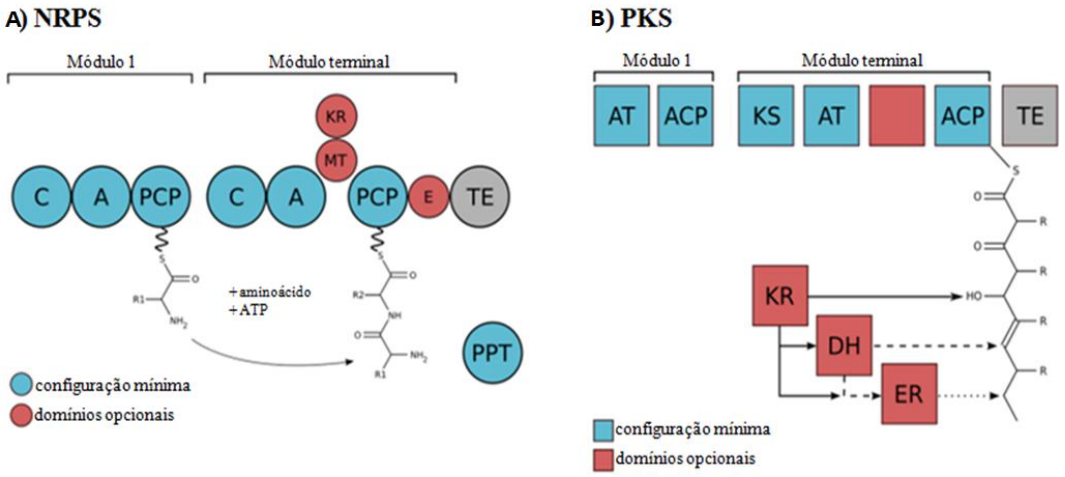


Figura 2 - Representação esquemática dos domínios enzimáticos de: A) NRPS e B) PKS. Abreviações: C: domínio de condensação; A: domínio de adenilação; PCP: proteína carreadora de peptidil; MT: metiltransferase; E: epimerase; AT: aciltransferase; ACP: proteína carreadora de acila; KS: cetossintase; KR: cetorredutase; DH: desidratase; ER: enoil-redutase; TE: tioesterase (adaptado de KEHR et al., 2011).

As cianotoxinas cujas biossínteses já foram descritas, ou seja, microcistina (TILLETT et al., 2000), nodularina (MOFFITT; NEILAN, 2004), saxitoxina (KELLMANN et al., 2008), cilindrospermopsina (MIHALI et al., 2008) e anatoxina-a (MÉJEAN et al., 2009), além dos inibidores de protease anabenopeptina (SIVONEN; BÖRNER, 2008) e espumigina (FUJII et al., 1997), são substâncias de estrutura peptídica sintetizadas por enzimas do tipo NRPS, PKS ou híbridos NRPS/PKS, cujos genes estão organizados na forma de agrupamento no genoma das cianobactérias que as sintetizam (KELLMANN et al., 2008; MIHALI et al., 2008; MÉJEAN et al., 2009; MOFFITT; NEILAN, 2004; TILLETT et al., 2000).

Grande parte da variação na estrutura química de produtos naturais está relacionada diretamente com as vias biossintéticas. Nas NRPS e PKS a falta de especificidade dessas enzimas ou, até mesmo, sua não atividade confere a diversidade química desejada; já para peptídeos ribossomais, as modificações estão inseridas dentro da rota, por múltiplos precursores proteicos, ou por precursores com múltiplas regiões catalíticas, além da ativação ou inativação de determinadas enzimas. Essas modificações possuem uma base genética que pode estar relacionada com recombinação, duplicação ou deleção gênica, perda ou ganho de enzimas ou mudanças por afinidades de substratos (DITTIMANN et al., 2015). Dessa forma, o potencial das cianobactérias na prospecção de produtos naturais vai muito além da exploração da bioatividade desses produtos. O conhecimento sobre a genômica e a bioquímica das enzimas é

particularmente valioso do ponto de vista da biologia sintética e para a construção de bibliotecas de novas substâncias.

O potencial biotecnológico dos produtos naturais de cianobactérias se mostra principalmente na manipulação genética e na expressão heteróloga de substâncias bioativas. Antes, a expressão heteróloga de mega sintetazes, NRPSs e PKSs era vista como um desafio, já que essas enzimas são grandes em tamanho e requerem que o organismo hospedeiro codifique enzimas as quais não estão presentes em seu metabolismo original. O sucesso na expressão heteróloga de algumas moléculas cianobacterianas, como o pequeno agrupamento gênico da lingbiatoxina que foi expresso em *E. coli* e da barbamida expressa em *Streptomyces venezuelae* (DITTIMANN et al., 2015), mostra que estes organismos, que antes eram considerados de difícil manipulação genética, são uma fonte rica de novas moléculas bioativas.

2.1.1 A neurotoxina anatoxina-a(s)

A anatoxina-a(s) foi identificada pela primeira vez na linhagem *Anabaena flos-aquae* NRC-525-17, isolada de uma floração ocorrida em 1965 na lagoa Bufallo, província de Saskatchewan, Canadá (CARMICHAEL; GORHAM, 1978). Esses autores notaram uma excessiva salivação nos ratos de laboratórios expostos aos extratos de *A. flos-aquae* NRC-525-17 e para diferenciar dos sintomas observados com a anatoxina-a eles a designaram anatoxina-a(s), sendo o sulfixo “s” referente à salivação.

As florações de cianobactérias dominadas por *Anabaena lemmermannii* ocorridas em lagos da Dinamarca em julho de 1993 foram implicadas na morte de pássaros (ONODERA et al., 1997). Esses autores analisaram quimicamente extratos da cultura de *A. lemmermannii* e verificaram que ela produzia ANS. A linhagem *A. lemmermannii* PH-160, isolada do lago Knud (Dinamarca), cujo extrato foi analisado quimicamente, também mostrou ser produtora de ANS (HENRIKSEN et al., 1997). Nos Estados Unidos, em 26 de agosto de 1985 as mortes de nove cachorros foram atribuídas às cianobactérias do lago Richmond, Dakota do Sul (MAHMOOD et al., 1988). Esses autores verificaram que a espécie dominante nesse lago era *A. flos-aquae* e bioensaios confirmaram a produção de ANS. Florações também com dominância de *A. flos-aquae* produtora de ANS ocorridas em Illinois, EUA, foram responsabilizadas pelas mortes de porcos e patos (COOK et al., 1989).

No Brasil, uma série de ocorrências foi relatada. A cianobactéria *Anabaena spiroides*, predominante em uma floração ocorrida em um lago ornamental da Universidade Federal do Rio Grande, RS, Brasil, foi considerada responsável pela produção de ANS detectada por

bioensaio e teste enzimático de inibição da atividade da acetilcolinesterase em amostras dessa floração (MONSERRAT et al., 2001). No monitoramento da água do reservatório de Tapacurá, Recife, PE, Brasil, foi coletada uma amostra de floração de cianobactérias em 27 de março de 2002 de onde foram isoladas três linhagens identificadas como *Anabaena spiroides* ITEP-024, ITEP-025 e ITEP-026 (MOLICA et al., 2005). Essas três linhagens apresentaram toxicidade nos bioensaios com camundongos apresentando sintomas típicos daqueles induzidos por ANS. A produção da ANS foi confirmada por teste enzimático de inibição da atividade da acetilcolinesterase, em que os extratos das três linhagens apresentaram inibição de 100%. Análises por espectrometria de massas confirmaram a produção de ANS nessas três linhagens (DÖRR et al., 2010). Posteriormente, estudo de filogenia do gene de RNAr 16S renomearam essas linhagens como *Sphaerospermopsis torques-reginae* (WERNER et al., 2012). Em amostras de água coletadas em janeiro e fevereiro de 2004 no reservatório Faxinal, que abastece a cidade de Caxias do Sul, RS, Brasil, foi constatado, por bioensaio e teste enzimático de inibição da atividade da acetilcolinesterase, a presença de ANS (BECKER et al., 2010). Esses autores relacionaram a detecção dessa toxina com a ocorrência no período de amostragem de florações de *Anabaena crassa*.

A estrutura da molécula de ANS foi elucidada em 1989, revelando um inusitado éster organofosforado natural (N-hidroxiguanidina fosfato de metila - $C_7H_{17}N_4O_4P$) com massa molecular de 252 Da, identificado por espectrometria de massas pelo íon molecular *m/z* 253,1066 ($[M+H]^+$) (MATSUNAGA et al., 1989) (Figura 3).

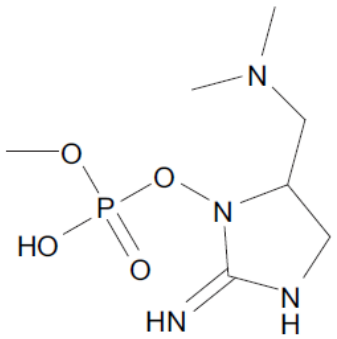


Figura 3 - Estrutura química da anatoxina-a(s) (adaptado de DÖRR et al., 2010).

Em 1992, algumas etapas da biossíntese de anatoxina-a (s) foram propostas (MOORE et al., 1992). A síntese começaria com uma molécula de arginina, que sofreria uma oxidação estéreosseletiva na posição do carbono gama (**1**), seguida por uma substituição nucleofílica intramolecular (SN^2) que retiraria o grupamento hidroxila para formar um anel heterocíclico

contendo o grupo guanidina (**2**) (Figura 4). A partir de mais uma nova oxidação desta arginina modificada (arginina hidroxilada no carbono beta), haveria a formação do intermediário (**3**), que sofreria uma clivagem, possivelmente por um mecanismo retro aldol, liberando um resíduo de glicina (**4**) e formando um aldeído (**5**). Experimentos envolvendo a incorporação de dois átomos de deutério no carbono (C3) da arginina modificada parecem dar suporte aos mecanismos de oxidação seguida de clivagem retro aldol para dar origem ao aldeído (**5**). No entanto, considerando o mecanismo proposto, deveria ter restado um átomo de deutério ligado ao grupamento aldeído, o que não foi confirmado por Moore e colaboradores em 1992. O mecanismo pelo qual a porção dimetilamina é adicionada ao intermediário (**5**) para formar o intermediário (**6**) permanece desconhecido, o que poderia evidenciar um passo enzimático que não pôde ser explicado pela síntese orgânica. O mecanismo proposto por Moore e colaboradores, envolvendo uma aminação redutiva, seguida de duas etapas de metilação do aminogrupo, não explica a perda do átomo de deutério restante que foi observada durante os experimentos. Com a estrutura central montada, a amina terciária (**6**) sofre outra oxidação seletiva (**7**), seguida por fosforilação da hidroxiguanidina resultante, provavelmente por uma via enzimática, dando forma à estrutura final da ANS (**8**) (MOORE et al., 1992; HEMSCHEIDT et al., 1995; SHIMIZU, 1996; MOORE, 1999). O intermediário (**2**) encontrado na biossíntese proposta foi descrito como enduracididina em 1984 por Hatano e colaboradores (HATANO et al., 1984). A enduracididina e sua versão hidroxilada (**3**) são componentes de diversos produtos naturais como manopeptimicina, viomicina e o antibiótico enduracidina (HATANO et al., 1984; HALTLI et al., 2005; YIN; ZABRISKIES, 2006; BURROUGHS et al., 2013). Dessa forma, verificou-se que os carbonos da ANS eram derivados de um aminoácido (MOORE et al., 1992; 1993), no qual os carbonos da estrutura central do triaminopropano e do grupamento guanidino eram derivados de L-arginina; além disso, observou-se na síntese a presença dos intermediários (2S,4S)-4-hidroxiarginina e enduracididina (HEMSCHEIDT et al., 1995). A síntese desse organofosforado, portanto, distingue-se das demais cianotoxinas já descritas, cujas sínteses se desenvolvem pela ação de megassintetases do tipo NRPSs, PKSs ou de híbridos de NRPS/PKS. No caso da ANS, nota-se que a molécula é resultado apenas de modificações sofridas pelo aminoácido precursor, a L-arginina, levando à estrutura final da toxina. Como não há o alongamento da molécula pela adição de outros aminoácidos ou pela adição de outras moléculas, conclui-se que a anatoxina-a(s) não é sintetizada pela via não ribossomal de metabólitos secundários, como ocorre com as outras cianotoxinas já descritas.

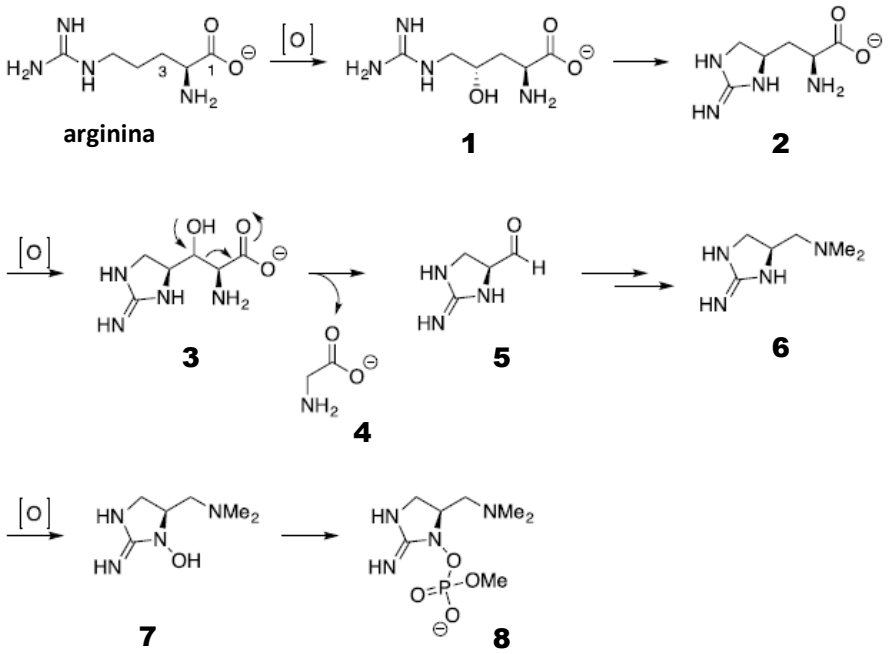
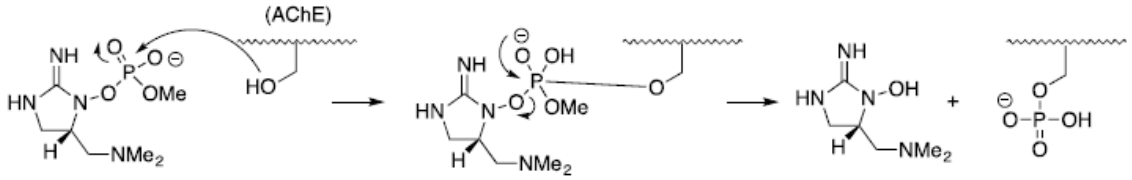


Figura 4 - Etapas da via biossintética proposta para anatoxina-a(s) por Moore e colaboradores (Adaptado de MOORE et al., 1993).

A ANS é uma potente e irreversível inibidora da enzima acetilcolinesterase (DL_{50} de 20 - 40 $\mu g.kg^{-1}$ i.p. em ratos) que atua apenas no sistema nervoso periférico e tem um mecanismo de ação semelhante aos inseticidas organofosforados paraoxon, fisostigmina, piridostigmina e à arma química sarin (COOK et al., 1988; MAHMOOD; CARMICHAEL, 1987; PETRIKOVICS et al., 2004; PITA et al., 2003; TUOVINEN et al., 1999; VAN APELDOORN, 2007). Foi proposto que a ANS fosforila irreversivelmente um resíduo serina no sítio ativo da acetilcolinesterase. Como mostra a Figura 5A, a hidroxila de um resíduo de serina da acetilcolinesterase é adicionada na porção fosfato da (ANS), o que leva à liberação da hidroxiguanidina. É possível que a ativação da funcionalidade da guanidina ocorra antes da adição da hidroxila da serina para ajudar a promover a fosforilação. É essa fosforilação irreversível que propicia o acúmulo de acetilcolina na fenda sináptica (MATSUNAGA et al., 1989). Esse acúmulo de acetilcolina leva à paralisia neuromuscular e, eventualmente, à morte por insuficiência respiratória. Os sintomas de envenenamento por ANS em mamíferos incluem fraqueza muscular, dificuldade respiratória, rinorreia, e, por fim, convulsões seguidas de morte (ARÁOZ et al., 2010).

A)



B)

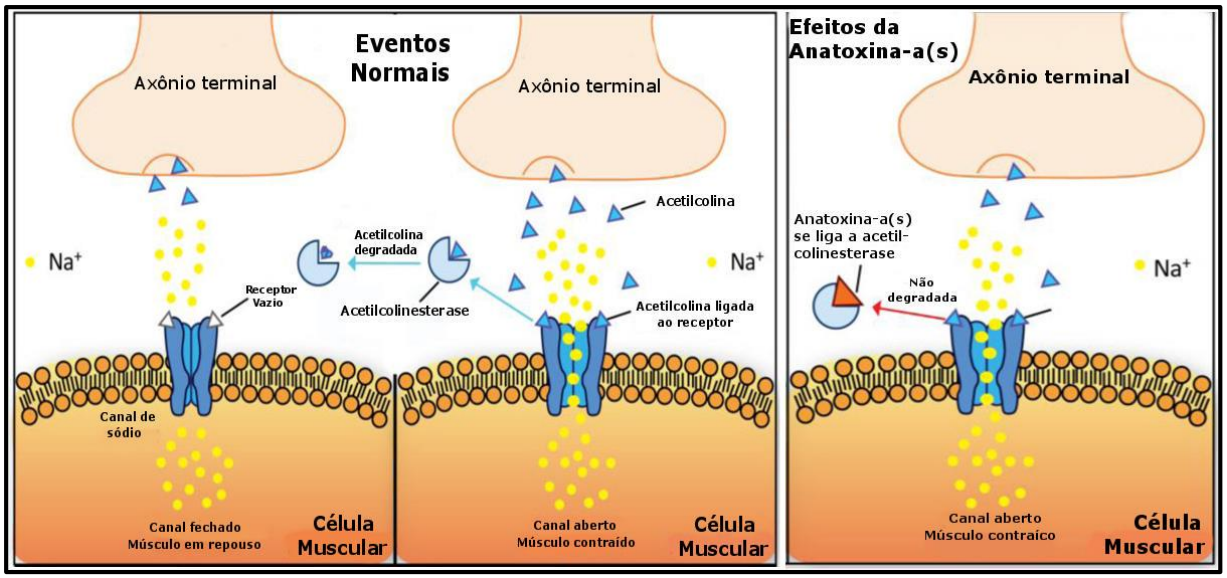


Figura 5 - Mecanismo proposto de ação da ANS. A) Representa a ligação covalente que ocorre entre a molécula de ANS e o sítio ativo da enzima acetilcolinesterase. B) Representação esquemática do processo biológico de inibição irreversível da acetilcolinesterase pela cianotoxina ANS (adaptado de VALERIO et al., 2010)

Apesar da elevada toxicidade da anatoxina-a(s), há uma baixa incidência de casos e relatos de intoxicação em humanos e outros animais, o que talvez possa ser explicado pela sua baixa estabilidade em meios alcalinos (MATSUNAGA et al., 1989) e pela falta de padrões analíticos comerciais disponíveis, fatores que dificultam o desenvolvimento de métodos específicos para determinação e identificação dessa cianotoxina. A detecção de ANS em amostras ambientais é feita em bioensaios com animais e testes de inibição da acetilcolinesterase, métodos que apresentam sérios problemas, como a baixa sensibilidade e especificidade. Além disso, várias questões éticas estão envolvidas em testes com animais, levando à busca de métodos alternativos de análise (DÖRR et al., 2010).

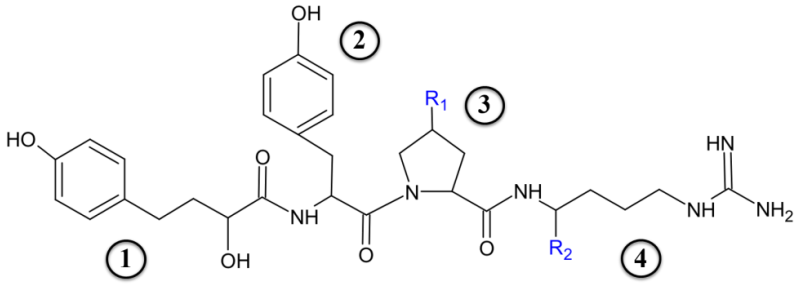
Devido ao seu mecanismo de ação, a ANS tem potencial aplicação na área farmacológica. A acetilcolinesterase está envolvida no encerramento da transmissão do impulso nervoso pela rápida hidrólise do neurotransmissor acetilcolina. Essa reação está envolvida em numerosas vias colinérgicas do sistema nervoso central e periférico. A inativação da enzima, induzida por inibidores como a ANS, conduz ao acúmulo de acetilcolina, levando à hiperestimulação de receptores, prejudicando a neurotransmissão (COLOVIC et al., 2013). Assim, os inibidores de acetilcolinesterase irreversíveis que interagem diretamente com essa

enzima podem ser aplicados como agentes terapêuticos no caso de algumas doenças oftalmológicas, como glaucoma crônico e doenças associadas ao aumento da pressão ocular (COLOVIC et al., 2010), além de poderem ser aplicados em neuropatias como Alzheimer e Parkinson (MCGLEENON et al., 1999). Na agricultura, a ANS poderia ser utilizada como inseticidas, já que sua ação se iguala aos inseticidas organofosforados, com a vantagem dessa toxina ser solúvel em água em meio alcalino.

2.1.2 Espumigina

Inibidores de proteases é uma família de moléculas sintetizadas por cianobactérias e outras bactérias amplamente aplicadas em fármacos antirretrovirais, usados na terapia combinada para o tratamento de infecções pelo vírus HIV, além da aplicação em doenças gástricas e cardíacas (MACHADO et al., 2014). Essas substâncias tipicamente contêm aminoácidos não proteinogênicos e são na maioria das vezes cíclicos. Elas estão organizadas em famílias com base nas suas características estruturais e funcionais (NAMIKOSHI; RINEHART, 1996; BURJA et al., 2001; SIVONEN; BÖRNER, 2008). Muitos desses peptídeos possuem uma grande variação química e geralmente são produtos finais de uma via que envolve peptídeo sintetase não ribossomal (MARAHIEL et al., 1997; SIEBER; MARAHIEL, 2005), com exceção das microviridinas, inibidores de proteases que são sintetizados pela via ribossomal em cianobactérias.

Espumiginas são tetrapeptídeos lineares que contêm na parcela N-terminal um ácido D-hidroxifenilacético, metil prolina e um derivado aldeído de arginina na porção C-terminal (Figura 6) (FEWER et al., 2009). Vinte variantes já foram descritas em *Nodularia spumigena* e *Anabaena compacta* (Tabela 3), variando a porção C-terminal, metil prolina e L-prolina, sendo a variante que apresenta metil prolina um potente inibidor de tripsina (FUJII et al., 1997a; FEWER et al., 2009; MAZUR-MARZEE et al., 2013; LIU et al., 2015). A presença de aminoácidos não proteicos sugere que as espumiginas sejam sintetizadas pela via não ribossomal de biossíntese de metabólitos secundários, ou seja, são produtos de uma NRPS. Os agrupamentos gênicos já descritos na literatura apresentam quatro módulos de NRPS responsáveis pela confecção da molécula (FEWER et al., 2009).



R_1 = H; R_1 = CH_3
R_2 = CHO; R_2 = CH_2OH; R_2 = NH_2; R_2 = OH

Variantes	Subunidades e suas posições				Linhagens produtoras	Referências
	1	**2**	**3**	**4**		
A	Hpla	Hty	MePro	Argol		FUJII et al., 1997a
B1	Hpla	Hty	MePro	Arg		FUJII et al., 1997a
B2	Hpla	Hty	MePro	Arg		FUJII et al., 1997a
C	Hpla	Hty	Pro	Arg		FUJII et al., 1997a
D	Hpla	Hty	Pro	Argol	*Nodularia spumigena*	FEWER et al., 2009
E	Hpla	Hty	MePro	Argal		FEWER et al., 2009
F	Hpla	Hty	Pro	Argal		FEWER et al., 2009
G	Hpla	Hph	MePro	Argal		FEWER et al., 2009
H	Hpla	Hph	Pro	Argal		FEWER et al., 2009
I	Hpla	Leu	Pro	Argol		FEWER et al., 2009
J	Hpla	Hty	MePro	Me-Lys	*Anabaena compacta*	ANAS et al., 2012
1	Hpla(Ac)	Hty	Pro	Arg		MAZUR-MAEZEE et al., 2013
2	Hpla(Ac)	Hty	MePro	Argal		MAZUR-MAEZEE et al., 2013
3	Hpla(Ac)	Hty	Pro	Argol		MAZUR-MAEZEE et al., 2013
4	Hpla(Ac)	Hty	Pro	Argal		MAZUR-MAEZEE et al., 2013
12	Hpla	Tyr	Pro	Argal	*Nodularia spumigena*	MAZUR-MAEZEE et al., 2013
13	Hpla	Hty	MePro	Agm		MAZUR-MAEZEE et al., 2013
15	Hpla(Ac)	Leu	Pro	Argal		MAZUR-MAEZEE et al., 2013
17	Hpla	Hty	MePro	NH_2		MAZUR-MAEZEE et al., 2013
18	Hpla	Hty	Pro	OH		MAZUR-MAEZEE et al., 2013

Figura 6 – Estrutura química geral da espumigina. R1 e R2 ilustram os substituintes das variantes nas respectivas posições. As variações na estrutura são mostradas esquematicamente de acordo com suas posições. Abreviações: Hpla, ácido hidroxifenilacético; Hpla(Ac), ácido acetil hidroxifenilacético; Hty, homotirosina; Leu, leucina; Tyr, tirosina; Me-Pro, metil prolina; Pro, prolina; Argol, argininol; Argal, argininal; Arg, arginina; Me-Lys, metil lisina; Agm, agmatina

2.1.3 Anabenopeptina

Anabenopeptinas (Figura 7) são potentes inibidores de proteases e foram nomeadas pela descoberta em *Anabaena flos-aquae*, sendo isoladas nessa linhagem juntamente com microcistina e anatoxina-a(s) (MATSUNAGA et al., 1989; HARADA et al., 1995). A nomenclatura dessa classe de moléculas é muito diversificada, podendo referir-se à linhagem de origem, localidade geográfica ou o peso molecular. Além disso, 96 variantes de anabenopetina já foram reportadas, sendo 87 isoladas de cianobactérias e 9 de esponjas, o que mostra a grande versatilidade desta substância (SPOOF et al., 2016).

Anabenopeptinas são peptídeos cíclicos que possuem um anel com cinco resíduos de aminoácidos conectados a um resído exocíclico por uma ligação denominada ureído, a qual ainda não teve seu mecanismo de formação totalmente elucidado (CHRISTIANSEN et al., 2011; KEHR et al., 2011). A estrutura geral pode ser representada por X_1-CO-[Lys-X_3-X_4-X_5-X_6], em que os colchetes representam os aminoácidos constituintes do ciclo, e X_1 e X_3-X_6 são resíduos de aminoácidos proteicos ou não proteicos variáveis. O ciclo se origina a partir da ciclização da carboxila da porção C-terminal com a amina primária da porção N-terminal da lisina. A ligação ureído juntamente com a configuração D da lisina se mantêm conservadas, podendo os outros resíduos variar a sua configuração (SPOOF et al., 2016).

Esses inibidores de proteases são sintetizados a partir de uma peptídeo sintetase não ribossômica (NRPS), a qual possui domínios específicos responsáveis pela seleção dos resíduos de aminoácidos e pela elongação da cadeia. A baixa especificidade por substratos da NRPS de anabenopeptina é responsável pela rica diversidade estrutural desses peptídeos. O agrupamento gênico pode conter um único módulo iniciador, como ocorre em *Nostoc* e *Nodularia*, ou dois módulos iniciadores, como ocorre em *Anabaena* sp. 90 (ROUHIAINEN et al., 2010). Ademais, moléculas com massas moleculares mais baixas apresentam maior poder de inibição para determinadas enzimas proteolíticas, sendo a característica da grande maioria das anabenopeptinas a inibição de metaloexopeptidases de Zn, como as carboxipeptidases A e B (SPOOF et al., 2016).

O papel funcional das anabenopetinas ainda não é muito bem compreendido, porém acredita-se que essas substâncias estejam envolvidas em mecanismos de defesas contra organismos parasitas, como fungos e amebas patogênicas, ou no controle da densidade das células cianobacterianas (URRUTIA-CORDEIRO et al., 2013; SEDMAK et al., 2008).

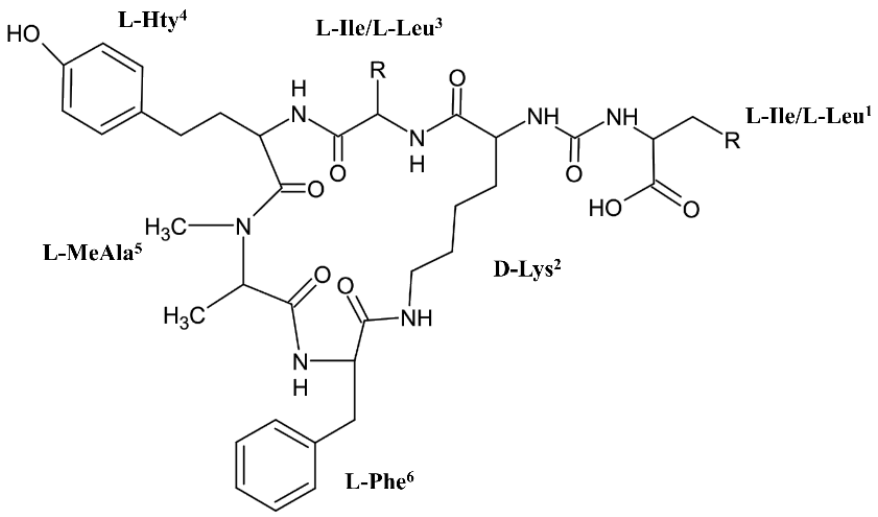


Figura 7 - Estrutura química da anabenopeptina encontrada na linhagem *Sphaerospermopsis torques-reginae* ITEP-024; abreviações: Ile, isoleucina; Leu, leucina; Hty, homotirosina; MeAla, metil alanina; Phe, fenilalanina

2.2 Genômica de cianobactérias

Em 2005, o surgimento das técnicas de sequenciamento massivo de DNA proporcionou um grande impulso na pesquisa de genomas e metagenomas. O rápido avanço dessa técnica de sequenciamento, a redução dos custos deste procedimento e a disponibilidade de softwares e novas ferramentas de bioinformática de uso público e de fácil acesso pela internet têm propiciado um aumento substancial de informações genômicas dos organismos num menor intervalo de tempo. Embora o número atual de genomas (completos e incompletos) cianobacterianos seja baixo (Figura 8), comparado com outros filos, esforços recentes têm objetivado aumentar este número.

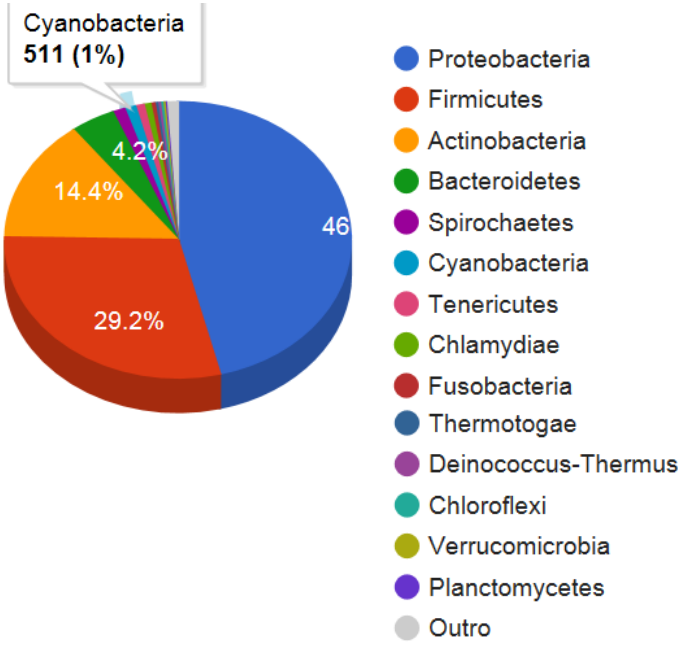


Figura 8 - Porcentagem de projetos de sequenciamento de genomas cianobacterianos em relação aos outros filos bacterianos disponível no "Genomes Online Database" (GOLD) em 15/02/2016. Fonte: http://genomesonline.org/cgi-bin/GOLD/index.cgi

O primeiro genoma de cianobactéria foi sequenciado em 1996 por pesquisadores japoneses e a linhagem escolhida foi *Synechocystis* sp. PCC 6803, isolada de água doce. Essa linhagem foi completamente sequenciada utilizando 107 clones de cosmídeos, 15 clones lambda e 17 produtos de PCR, contendo aproximadamente 3,5 Mpb (KANEKO et al., 1996). As informações geradas permitiram revelar a estrutura do genoma, identificar genes e mapear as posições relativas de cada gene no genoma. Além disso, a função de aproximadamente metade dos genes foi deduzida, baseando-se na similaridade da sequência com genes que apresentam função conhecida. Um dos fatos importantes também foi o estudo da maquinaria fotossintética desta bactéria. Uma grande parte dos genomas cianobacterianos sequenciados é de linhagens provenientes de ambientes marinhos que apresentam baixa diversidade genérica. Esse número ainda reduzido de genomas cianobacterianos pode ser um reflexo da complexidade desse grupo bacteriano. A diversidade formidável das cianobactérias se reflete não só nas várias morfologias, capacidades fisiológicas e seleção de hábitat, mas também ao nível do genoma. Os tamanhos dos genomas conhecidos de cianobactérias variam mais de 10 vezes, de pouco mais de 1 Mpb a mais de 10 Mpb, e sequências codificantes de cerca de

1.200 a mais de 11.000 (DAGAN et al., 2013; LARSSON; NYLANDER; BERKAM, 2011; SHIH et al., 2013). Além disso, vários gêneros apresentam altas proporções de sequências de inserção que dificultam o fechamento de genomas. Durante os bilhões de anos de evolução das cianobactérias, uma série de específicos nichos ambientais e forças competitivas provavelmente atuaram sobre elas e podem ter moldado genomas individuais para refletir exclusivamente as características exibidas. Expansão e diminuição do genoma pela eliminação de genes dispensáveis, aumento do conteúdo G-C e variação da taxa evolutiva são algumas das mudanças observadas. Em contraste a isso, alguns genes são extremamente conservados os quais geralmente estão envolvidos na síntese de estruturas proteicas complexas ou rotas relacionadas ao metabolismo essencial (LARSSON; NYLANDER; BERKAM, 2011).

O sequenciamento de genomas de cianobactérias pertencentes à ordem Nostocales tem sido mais reduzido do que os das outras ordens, provavelmente devido ao esperado maior tamanho de genoma. A Tabela 2 apresenta a lista de genomas completos de nostocacean sequenciados até o momento. Como pode ser verificado, o gênero *Sphaerospermopsis*, pertencente à ordem Nostocales e objeto do presente estudo, não possui linhagens sequenciadas.

Tabela 2 - Relação de genomas de cianobactérias da ordem Nostocales com sequenciamento completo e incompleto (NCBI, janeiro de 2016)

Genomas completos	**Ambiente isolado**	**Tamanho (Mpb)**	**(G+C)%**	**Genes**	**NCBI N° Projeto**	**Referência**
Anabaena/*Nostoc* sp. PCC7120	água doce	7,21	41,3	6.222	244	KANEKO et al. 2001
Anabaena variabilis ATCC29413	água doce	7,11	41,4	5.813	10642	THIEL et al. 2014
Anabaena sp. Wa102	desconhecido	5,70	38,4	4.742	282166	Não publicado
Anabaena sp. 90	água doce	5,31	38,1	4.797	30803	WANG et al., 2012
Anabaena cylindrica PCC 7122	água doce	7,06	38,8	6.258	43355	SHIH et al., 2013
Calothrix sp. 336/3	lago	6,42	41,41	4.946	255071	ISOJARVI et al., 2015
Calothrix sp. PCC 6303	água doce	6,96	39,8	5.841	158041	SHIH et al., 2013
Calothrix sp. PCC 7507	pântano	7,02	42,2	6.250	158683	SHIH et al., 2013
Cylindrospermum stagnale PCC 7417	solo	7,61	42,2	6.738	158809	SHIH et al., 2013
Fischerella sp. NIES-3754	água termal	5,82	41,0	3.982		HIROSE et al., 2016
Nostoc sp. PCC 7107	água doce	6,33	40,4	5.538	158705	SHIH et al., 2013
Nostoc azollae 0708	endossimbionte de *Azolla filiculoides*	5,49	38,4	5.379	30807	RAN et al. 2010
Nostoc punctiforme PCC 73102	simbionte de *Macrozamia riedlei*	9.06	41,3	7.164	216	EKMAN et al. 2013
Nostoc sp. PCC 7524	fonte termal	6.72	41,5	5.687	158707	SHIH et al., 2013
Rivularia sp. PCC 7116	marinha	8,73	37,5	6.946	63147	SHIH et al., 2013
Genomas incompletos						
Anabaena sp. PCC 7108	marinha	5,89	38,8	5.263	158737	SHIH et al., 2013
Aphanizomenon flos-aquae 2012/KM1/D3	laguna	5,74	38,2	5.478	257725	ŠULIUS et al, 2015
Aphanizomenon flos-aquae NIES-81	água doce	5,85	38,5	5.802	232534	CAO et al., 2014
Calothirx sp. PCC 7103	desconhecido	11.58	38,6	10.120	159495	SHIH et al., 2013
Cylindrospermopsis sp. CR12	água doce	3,72	40,08	3.269	300393	Não publicado
Cylindrospermopsis raciborskii CS-505	água doce	3,88	40,2	3.501	40109	STUCKEN et al., 2010
Chlorogloeopsis sp. PCC 9212	água termal	7,65	41,5	6.749	104963	DAGAN et al, 2012
Chlorogloeopsis fritschii PCC 6912	solo	7,75	41,5	6.914	104961	DAGAN et al, 2012
Dolichospermum circinale AWQC131C	água doce	4,44	37,33	4.062	179122	Não publicado
Dolichospermum circinale AWQC310F	água doce	4,40	37,49	3.991	179123	Não publicado
Fischerella muscicola PCC 7414	água termal	6,9	41,2	6.138	104967	DAGAN et al, 2012
Fischerella muscicola PCC 73103	solo	7,36	40,3	6.384	104965	DAGAN et al, 2012
Fischerella thermalis PCC 7521	água termal	5,44	41	4.684	104969	DAGAN et al, 2012
Fischerella sp. JSC-11	água termal	5,38	41,1	4.671	61093	SHIH et al., 2013

continua

Genomas incompletos	Ambiente isolado	Tamanho (Mpb)	(G+C)%	Genes	NCBI N° Projeto	Referência
Fischerella sp. PCC 9431	desconhecido	7,14	40,2	–	158821	SHIH et al., 2013
Fischerella sp. PCC 9605	rocha calcárea	8,2	42,6	–	158819	SHIH et al., 2013
Hapalosiphon sp. MRB220	água	7,42	40,21	6.173	293405	Não publicado
Hassallia byssoidea VB512170	biofilme sobre monumentos	13,0	44,0	10.183	266752	SINGH et al., 2015
Mastigocladopsis repens PCC 10914	desconhecido	6,47	43,5	5.757	158735	SHIH et al., 2013
Mastigocladus laminosus UU774	água termal	8,56	44,7	6.654	272369	Não publicado
Mastigocoleus testarum BC008	marinha	15,87	48,2	–	163055	
Microchaete sp. PCC 7126	água doce	5,74	42,2	5.339	158817	SHIH et al., 2013
Nodularia spumigena CCY 9414	marinha	5,32	41,3	4.904	13447	VOß et al., 2013
Nostoc piscinale CENA21	planície de inundação	7,09	40,54	5.022	288507	Não publicado
Raphidiopsis brookii D9	água doce	3,19	40,1	3.057	40111	STUCKEN et al., 2010
Richelia intracellularis	simbionte da diatomácia *Rhizosolenia*	5,40	39,16	7.768	5207	HILTON, 2014
Richelia intracellularis HH01	simbionte da diatomácia *Rhizosolenia*	3,20	34	2.012	2272692	HILTON et al., 2012
Richelia intracellularis HM01	simbionte da diatomácia *Rhizosolenia*	2,20	33,76	1.742	2272302	HILTON et al., 2012
Scytonema hofmanni PCC 7110	rocha calcárea	12,01	41,5	10.706	157363	DAGAN et al, 2012
Scytonema hofmanni UTEX 2349	rocha	8,18	41,2	–	63425	SHIH et al., 2013
Scytonema tolypothrichoides VB61278	biofilme sobre monumentos	10	47	7.148	271448	DAS et al., 2015
Scytonema millei VB511283	biofilme sobre monumentos	11,63	51	3.011	3200207	SEN et al., 2015
Tolypothrix bouteillei VB521301	biofilme sobre monumentos	11,5	42	7.777	242379	CHANDRABABUNAIDU et al., 2015
Tolypothrix campylonemoides VB511288	biofilme	10,62	48	7.886	271449	DAS et al., 2015
Tolypothrix sp. PCC 7601	lago	9,96	40,6	10.065	72337	YERRAPRAGADA et al., 2015

conclusão

No Brasil, os genomas cianobacterianos até o momento publicados são de espécies unicelulares (FIORE et al., 2013; GUIMARÃES et al., 2015; LIMA et al., 2014; RIGONATO et al., 2015), embora sequenciamentos de genomas de morfotipos filamentosos e filamentosos ramificados de cianobactérias estejam em andamento (M.F. Fiore comunicação pessoal).

A disponibilidade de genomas tem facilitado a identificação e caracterização de agrupamentos gênicos de produtos naturais (DONIA; SCHMIDT, 2011; KERSTEN et al., 2011; LEIKOSKI et al., 2010; LI et al., 2010; WANG et al., 2015; ZIEMERT et al., 2010). As sequências genômicas disponíveis de cianobactérias têm permitido identificar presumíveis genes precursores de produtos naturais conhecidos e também de novas substâncias. Estes estudos utilizam tanto genes codificantes de precursores de peptídeos como alvo para as buscas nos genomas, resultando em análogos de substâncias conhecidas, ou utilizam genes que codificam enzimas biossintéticas chaves, muitas vezes resultando em mais análogos distantes dos produtos naturais conhecidos. Essa abordagem genômica fornece informações do tipo presente de agrupamento gênico biossintético e, correspondentemente, das predições estruturais dos produtos naturais produzidos (CORRE; CHALLIS, 2009; MICALLEF et al., 2015; MÜLLER; WINK, 2014; NETT, 2014). As cianobactérias possuem uma média de cinco agrupamentos gênicos de produtos naturais em seus genomas, algumas podendo chegar até a 23 agrupamentos, no entanto, a grande maioria deles sem conhecimento do produto final (DITTIMANN et al., 2015). Os agrupamentos gênicos desconhecidos são características de quase todos os sequenciamentos de cianobactérias. A diversidade genética vai muito além da diversidade química dos produtos naturais conhecidos, e conhecê-la é parte importante para a descoberta das vias de produção e seus produtos finais. As abordagens genômicas têm possibilitado averiguar mais facilmente a diversidade metabólica dos micro-organismos cuja relevância antes era subestimada pela falta de métodos de mineração genômica (WINTER; BEHNKEN; HERTWECK, 2011). Antes, os metabólitos secundários eram encontrados por meio do rastreamento de grandes bibliotecas de genes, sendo que atualmente são iniciados projetos de sequenciamento de genomas para a busca destas substâncias utilizando o conhecimento da estrutura química da molécula de interesse (DITTIMANN et al., 2015) ou a caracterização das sequências gênicas que evidenciam rotas de produção a partir de sintetases não ribossômicas, policetídeo sintases, terpeno sintases, síntese ribossômica, entre outras (CORRE; CHALLIS, 2009). Esses tipos de estudos são uma tendência instigante que apontam o poder da genômica na descoberta de novos produtos naturais, além de oferecer ferramentas importantes para a engenharia de peptídeos

(DITTIMANN et al., 2015). Com o aumento dos sequenciamentos de genomas de micro-organismos essa área de pesquisa caminha para o crescimento e evolução das técnicas.

2.3 A espécie *Sphaerospermopsis torques-reginae* e sua distribuição no Brasil

A espécie *S. torques-reginae* pertence ao filo Cyanobacteria, ordem Nostocales e família Aphanizomenonaceae (KOMÁREK et al., 2014). Entretanto, essa espécie tem uma história de reclassificação do seu gênero devido à problemática existente na sistemática do filo Cyanobacteria. O rearranjo na classificação das cianobactérias tem sido frequente uma vez que esta, elaborada principalmente com base nas características morfológicas, não é consensual, apresentando problemas em níveis genéricos e infragenéricos (PAN et al., 2008; WILLAME et al., 2006; RAJANIEMI et al., 2005; GUGGER; HOFFMANN, 2004; ITEMAN et al., 2002; RUDI et al., 1998; 1997), de modo que os caracteres morfológicos não refletem relações evolutivas e sequências do gene de RNAr 16S existentes não suportam muitas das morfoespécies descritas.

A *Sphaerospermopsis torques-reginae* foi descrita pela primeira vez como *Anabaena torques-reginae* em 1984 em populações planctônicas ocorridas em ambientes eutrofizados de Cuba (KOMÁREK, 1984). Em 2009, um estudo de filogenia baseado no gene de 16S RNAr 16S reclassificou todas as morfoespécies de *Anabaena* planctônicas como *Dolichospermum* (WACKLIN et al., 2009). Assim, a *Anabaena torques-reginae* passou a se denominar *Dolichospermum torques-reginae*. Entretanto, um novo estudo, baseado na caracterização polifásica das morfoespécies de *Anabaena* planctônicas, reclassificou o grupo formado por *Anabaena reniformis* e *Aphanizomenon aphanizomenoides*, do qual *Anabaena torques-reginae* fazia parte, como *Sphaerospermum* (ZAPOMĚLOVÁ et al., 2009). Porém, logo em seguida, o gênero *Sphaerospermum* foi renomeado como *Sphaerospermopsis*, uma vez que se verificou que o nome *Sphaerospermum* era ilegítimo, pois é um homônimo de *Sphaerospermum* Cleve (Nova Acta Regiae Societatis Scientiarum upsaliensis, Serie 3, 6 (11): 12, 35, 1868), atualmente considerado um sinônimo heterotípico do gênero *Mougeotia* C. Agardh (Systema Algarum xxvi, 83, 1824), *nomen conservandum* (ZAPOMĚLOVÁ et al., 2010).

Os táxons incluídos no gênero *Sphaerospermopsis* caracterizam-se morfologicamente pela forma e posição dos acinetos (esféricos, em um ou ambos os lados adjacentes ao heterócito) e pela presença de tricomas espiralados, sendo que a designação da espécie-tipo foi baseada na *Anabaena reniformis*, posteriormente renomeada como *Sphaerospermopsis reniformis*

(Lemmermann) (ZAPOMĚLOVÁ et al., 2009; 2010). A análise filogenética baseada no gene de RNAr 16S mostrou que sequências da espécie *S. torques-reginae* agruparam com alto suporte de reamostragem com outras espécies deste gênero, tais como *S. oumiana*, *S. reniformis*, *S. kisseleviana* e *S. aphanizomenoides* (WERNER et al., 2012).

A presença da espécie *S. torques-reginae* em ambientes aquáticos brasileiros foi observada em vários estados (Tabela 3). Porém, a sua distribuição pode ser ainda mais ampla, pois além da pesquisa com essa espécie ser reduzida e concentrada principalmente no estado do Rio Grande do Sul, existe a possibilidade de que espécies de *S. torques-reginae* tenham sido identificadas equivocadamente como pertencentes a outros gêneros e/ou espécies.

Tabela 3 - Locais de registro da espécie *S. torques-reginae* no Brasil e algumas características físico-químicas desses locais (adaptado de WERNER et al., 2012).

Locais	**Condições Ambientais**			
	Temperatura da água (ºC)	**pH**	**Condutividade (µS cm^{-1})**	**Salinidade (‰)**
Reservatório Tapacurá, Recife, PE	–	7,3 – 8,9	430 – 460	0
Lago da Estrada da Ponte 45, Parque Pantanal da Nhecolândia, Corumbá, MS	23,7	8,9	254	0
Lago Violão, Torres, RS	24,8 – 32,0	7,5 – 8,5	241 – 320	0
Lagoa dos Patos, Rio Grande, RS	26 - 30	–	–	0 - 4
Lago Mato, Sapucaia do Sul, RS	26	7,1	–	0
Lagoas artificiais da Universidade Federal do Rio Grande e da Fundação Zoobotânica do Rio Grande do Sul, RS	–	–	–	–
Lago Peixe, Ribeirão Preto, SP	–	–	–	–

S. torques-reginae foi relatada como formadora de florações em vários ambientes aquáticos subtropicais do sul do Brasil (WERNER e LAUGHINGHOUSE IV, 2009). Entretanto, essas populações foram identificadas como *Anabaena oumiana*, uma vez que na sua morfologia não foram observados filamentos paralelos característicos da espécie *Anabaena torques-reginae*. Essa espécie tem sido relatada geralmente em água doce, mas no extremo sul do Brasil foi também encontrada em uma lagoa costeira (Lagoa dos Patos) que contém uma concentração de sal de 4% (WERNER et al., 2012). Esses autores notaram que, geralmente, ela é observada em sistemas eutrofizados, usualmente em densas florações, predominantemente ou acompanhando florações de outras espécies de cianobactérias,

tais como *Microcystis aeruginosa*, *Microcystis protocystis*, *Microcystis panniformis*, *Sphaerocavum brasiliense*, *Radiocystis fernandoi*, *Cylindrospermopsis raciborskii*, *Anabaenopsis circularis*, *Pseudanabaena* sp. ou por clorofíceas, especialmente *Scenedesmus* sp. A observação de *S. torques-reginae* em diferentes ambientes aquáticos (lagos naturais e artificiais, lagoa, reservatórios e pântano), tanto em água doce como salobra de zonas tropicais, subtropicais e temperadas sugere que essa espécie tem uma ampla distribuição e plasticidade ecológica.

3 HIPÓTESE

A linhagem *Sphaerospermopsis torques-reginae* ITEP-024 é conhecida por produzir o organofosforado anatoxina-a(s) (ANS). A biossíntese de ANS é codificada por genes dispostos em um agrupamento que contém enzimas com funcionalidades semelhantes às propostas anteriormente por análises químicas. A *S. torques-reginae* ITEP-024 também sintetiza outras moléculas bioativas de interesse biotecnológico que podem ser identificadas pelos seus agrupamentos gênicos presentes no genoma e confirmados usando de técnicas de análises químicas.

4 OBJETIVO

O principal objetivo deste estudo foi realizar análise genômica no isolado brasileiro *Sphaerospermopsis torques-reginae* ITEP-024 visando identificar os genes envolvidos na síntese de anatoxina-a(s) (ANS), a fim de gerar informações relacionadas ao agrupamento gênico e à via biossintética dessa toxina cianobacteriana. Sendo assim, os objetivos específicos deste estudo foram: I) Obter o genoma incompleto da linhagem *S. torques-reginae* ITEP-024; II) Utilizar o genoma incompleto para mapeamentos e buscas por genes relacionados à síntese de ANS e de outros produtos naturais; III) Avaliar, por análises químicas, a produção de ANS e dos produtos naturais com agrupamentos gênicos detectados.

Para tanto, foram utilizadas as seguintes abordagens:

- Análise das funções das enzimas que compõe os agrupamentos gênicos preditos;
- Comparação das predições funcionais para as enzimas codificadas com a proposta da via biossintética de ANS anteriormente presumida por análises químicas;
- Comparação dos agrupamentos gênicos encontrados na linhagem em estudo com agrupamentos gênicos semelhantes encontrados em outras cianobactérias quando possível.

5 MATERIAL E MÉTODOS

5.1 Origem da linhagem, condições de cultivo e produção de biomassa da *Sphaerospermopsis torques-reginae* ITEP-024

A linhagem ITEP-024 foi isolada de uma amostra de floração de *Sphaerospermopsis torques-reginae* coletada em 27 de março de 2002 no reservatório de Tapacurá (8º02'14''S, 35º09'46''W), localizado no município de Recife, Pernambuco, conforme descrito anteriormente (MOLICA et al., 2005). Ela foi gentilmente cedida pela Dra. Vera Regina Werner, do Museu de Ciência Natural, Fundação de Zoobotânica do Rio Grande do Sul, Porto Alegre, RS, e é mantida em cultivo no laboratório de Biologia Celular e Molecular do CENA/USP em meio de cultura líquido ASM-1 (GORHAM et al., 1964), sob exposição à luz fluorescente (55 µmol fótons m^{-2} s^{-1}), com ciclo claro/escuro de 14/10h, a 22±1 ºC.

A produção de biomassa da linhagem *S. torques reginae* ITEP-024 foi realizada em frasco de 125 mL contendo 50 mL de meio de cultura líquido ASM-1, nas mesmas condições de luz e temperatura especificadas acima, durante 30 dias.

5.2 Sequenciamento do genoma da *Sphaerospermopsis torques-reginae* ITEP-024

O sequenciamento do genoma da *S. torques-reginae* ITEP-024 foi realizado utilizando o conjunto de reagentes *MiSeq Reagent Kit v3 600 cycles* e a plataforma MiSeq (Illumina Inc., San Diego, CA, EUA) existente no Centro de Genômica Funcional Aplicada à Agropecuária e à Agroenergia do Campus "Luiz de Queiroz" da Universidade de São Paulo, em Piracicaba-SP.

5.2.1 Extração de DNA total

As células de *S. torques-reginae* ITEP-024 contidas na cultura de 50 mL foram concentradas por centrifugação (9.500 × *g*, 10 min) e lavadas duas vezes com solução de lavagem (NaCl 50mM, Tris-HCl 10 mM pH 7,5, EDTA 2,5 mM pH 8,0 e etanol 50%) a fim de diminuir alguns contaminantes como outras bactérias e arqueias que poderiam estar associados à cianobactéria. Na obtenção de DNA total para o primeiro sequenciamento, realizado em abril de 2014, utilizou-se 1 mL das células lavadas e concentradas e o método descrito por Lin et al. (2010), modificado conforme descrito a seguir. As células concentradas foram ressuspendidas

em 500 µL de tampão de lise (100 mM Tris100 mM NaCl, 1 mM EDTA pH 9,0). Posteriormente, foi adicionada lisozima na concentração final de 1 mg mL^{-1} e a amostra foi incubada a 37 °C por 30 minutos. Após a incubação, adicionou-se proteinase K na concentração final de 0,4 mg mL^{-1} e SDS na concentração final 0,4%. As amostras foram incubadas a 55 °C por 1h. Em seguida, foram adicionados 600 µL de uma solução fenol:clorofórmio:álcool isoamílico (25:24:1 v/v). A amostra foi cuidadosamente homogeneizada por 5 minutos e então centrifugada (13.680 × *g*, 5 min). O sobrenadante foi transferido para um novo tubo, onde adicionou-se um volume de 600 µL de uma solução de clorofórmio:álcool isoamílico (24:1 v/v). A amostra foi novamente homogeneizada por 5 minutos e então centrifugada (13.680 × *g*, 5 min). O sobrenadante foi transferido para um novo tubo e adicionou-se RNAse na concentração final de 100 µg.mL^{-1}. Os tubos foram então incubados a 37 °C por 20 minutos. Em seguida, foram adicionados 100 µL de acetato de sódio 3 M a pH 4,8 e 600 µL de isopropanol. As amostras foram homogeneizadas e colocadas a - 80 °C por 30 minutos e, posteriormente, centrifugadas (13.680 × *g*, 10 min, 4 °C). O sobrenadante foi descartado e o precipitado lavado duas vezes com 500 µL de etanol 70% e centrifugado (13.680 × *g*, 5 min, a 4 °C). As amostras foram colocadas em bloco aquecedor a 37 °C até a secagem do precipitado. O DNA total foi ressuspendido em 50 µL de água ultrapura esterilizada. A qualidade dos DNAs extraídos foi verificada em gel de agarose 1% (m/v) e a quantificação foi feita com o fluorômetro Qubit® 2.0 (Life Technologies, Carlsbad, CA, EUA).

O DNA total obtido para o segundo sequenciamento genômico, realizado em fevereiro de 2015, foi extraído utilizando-se o conjunto de reagentes do *AxyPrep Bacterial Genomic DNA Miniprep Kit* (Axygen Biosciences, Union City, CA, EUA) de acordo com as instruções do fabricante. A qualidade do DNA extraído foi verificada em gel de agarose 1% (m/v) e a quantificação foi feita com o fluorômetro Qubit® 2.0 (Life Technologies).

5.2.2 Construção da biblioteca genômica e sequenciamento

A construção da biblioteca genômica *paired-ends* foi realizada usando 1 ng de DNA genômico da *S. torques-reginae* ITEP-024 e o conjunto de reagentes do *Nextera XT DNA Sample Prep Kit* (Illumina Inc.) de acordo com as instruções do fornecedor, resumidas a seguir. O DNA genômico foi incubado em termociclador à 55 °C por 5 minutos para ser fragmentado por transposases e receber sequências adaptadoras às suas extremidades. O tampão *Neutralize Tagment* foi utilizado para neutralizar as enzimas durante 5 minutos à temperatura ambiente.

Depois disso, procedeu-se a amplificação do DNA com os iniciadores específicos aos adaptadores anteriormente inseridos. A seguinte ciclagem térmica foi usada para a amplificação do DNA: 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 usando 30 µL de microesferas magnéticas AMPure XP 0,5 X e etanol 80 %. Depois da purificação dos produtos da amplificação, a biblioteca genômica foi quantificada por PCR em tempo real usando o conjunto de reagentes *KAPA SYBR FAST® qPCR Kit* (KAPABiosystems, Boston, MA, EUA) e submetida ao sequenciamento.

5.2.3 Processamento das sequências e montagem

Gráficos da qualidade das leituras brutas do sequenciamento pela plataforma MiSeq foram gerados pelo programa FastQC 0.10.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). O programa Focus (SILVA et al., 2014) foi utilizado para estimar a proporção das sequências dos contaminantes, já que a linhagem utilizada não era uma cultura axênica. Os pares das sequências que se complementavam foram unidas utilizando o programa PEAR 0.9.6 (ZHANG et al., 2014). Bases com índice de qualidade menor que Phred 30, bases ambíguas e sequências menores que 50 pb foram filtradas com Seqyclean 1.3.12 (http://cores.ibest.uidaho.edu/software/seqyclean) e PRINSEQ 0.20.4 (SCHMIEDER; EDWARDS, 2011). A montagem do genoma foi feita com os programas SPAdes 3.5.0 (BANKEVICH et al., 2014) e Platanus 1.2.1 (http://platanus.bio.titech.ac.jp/). Sequências com similaridade com cianobactérias foram selecionadas pelo programa Kraken (WOOD; SALZBERG, 2014).

5.2.4 Anotação e detecção dos agrupamentos gênicos de anatoxina-a(s), espumigina e anabenopeptina

A montagem gerada foi anotada usando o programa de anotação automática Prokka 1.9 (SEEMANN, 2014). Predições para potenciais agrupamentos gênicos de metabólitos secundários produzidos pela síntese não ribossômica foram realizadas utilizando o programa antiSMASH 3.0.4 (WEBER et al., 2015). A anotação automática de regiões de interesse foi averiguada e, quando necessário, corrigida por meio de curadoria manual com o programa Artemis 16.0.0 (RUTHERFORD et al., 2000). Análises comparativas de sequências

de proteínas envolvidas com a síntese de arginina foram mineradas usando o programa *Basic Local Alignment Search Tool* (BLAST) do National Center for Biotechnology Information (NCBI), o banco de dados Pfam (FINN et al., 2016) e a busca por domínios conservados (*CD-search*) no banco de dados do NCBI (MARCHLER-BAUER et al., 2015).

5.2.5 Análise filogenética de agrupamentos gênicos de anabenopeptinas

As sequências de nucleotídeos de agrupamentos gênicos de anabenopeptinas de seis linhagens cianobacterianas (*Nodularia spumigena* CCY9414, *Nostoc punctiforme* PCC 73102, *Anabaena* sp. 90, *Planktothrix agardhii* NIVA-CYA 126/8, *Planktothrix rubescens* NIVA-CYA 98 e *Sphaerospermopsis torques-reginae* ITEP-024) foram alinhadas usando o programa ClustalW (THOMPSON et al., 1994). Modelos evolutivos foram selecionados pelo jModelTest 2.1.7 (DARRIBA et al., 2012) e uma árvore filogenética foi construída com os nucleotídeos dos genes concatenados dos agrupamentos de anabenopeptina pelo método de inferência Bayesiana utilizando o MrBayes 3.2.5 (RONQUIST; HUELSENBECK, 2003) usando quatro cadeias de Markov, duas corridas separadas e 5.000.000 de gerações. Árvore filogenética de espumiginas não pode ser construída, pois existe apenas um agrupamento gênico atualmente disponível.

5.3 Análises em cromatografia líquida acoplada a espectrometria de massas (LC-MS/MS)

As análises em espectrometria de massas foram realizadas no Laboratório do Prof. Ernani Pinto do Departamento de Análises Clínicas e Toxicológicas da Faculdade de Ciências Farmacêuticas da Universidade de São Paulo, São Paulo, SP. A biomasa necessária para as análises foram produzidas em doze Erlenmeyers de 500 mL contendo 150 mL de meio de cultura líquido ASM-1, nas mesmas condições de luz e temperatura especificadas acima, durante 30 dias. As células foram, então, concentradas por centrifugação (2.370 × *g*, 5 min, 4°C) e liofilizadas. Nas análises de ANS, o preparo da amostra foi realizado seguindo as recomendações de Henriksen et al. (1997) e Dörr et al. (2012). As células liofilizadas (50 mg) foram ressuspendidas em 5 mL de etanol/ácido acético 0,1 M (20:80 v/v) e as células rompidas por sonicação. Posteriormente, a amostra foi centrifugada (2.370 × *g*, 15 min, 4 °C).

A separação dos pigmentos da amostra foi feita por meio da eluição do extrato em cartucho C_{18} (Sep-Pack, Waters, MA, USA) pré-condicionado com metanol e etanol/ácido acético 0,1 M (20:80 v/v). Os eluatos foram coletados e secados em Speed Vac (Thermo Scientific,

MA, EUA) e, em seguida, ressuspendidos em metanol e filtrados em tubos apropriados para as análises de LC-MS.

As análises de ANS foram realizadas seguindo os procedimentos descritos por Dörr et al. (2010). Experimentos de *Multiple Reaction Monitoring* (MRM) foram realizados em espectrômetro de massas, equipado com uma fonte de ionização por electrospray (ESI) e acoplado a um cromatógrafo líquido. Nitrogênio foi usado como gás de nebulização (15 psi) e secagem (5 L min^{-1}, 350°C) e o hélio como gás tampão (4×10^{-6} mbar). As análises foram realizadas no modo positivo, e as seguintes transições (*m/z* do íon precursor > íon fragmentado) foram selecionadas: *m/z* 253 > 159, *m/z* 253 > 98, *m/z* 253 > 96 e *m/z* 253 > 58. Tempos de permanência de 100 ms foram usados para cada transição. A voltagem do capilar foi de 3,5 KV. A vazão de fluxo foi de 0,1 mL min^{-1}. Separações cromatográficas foram alcançadas com um gradiente de eluição de 0,7 mL min^{-1} sendo o A: água/acetonitrila (20:80 v/v) contendo 5 mM de formiato de amônio e 2 mM de ácido fórmico; e o B: água contendo 10 mM de formiato de amônio e 10 mM de ácido fórmico. Um gradiente linear foi realizado com 20 a 35% de B até 5 min, um segundo gradiente de 40% de B até 10 min, um terceiro gradiente de 65% de B até 11 min e uma eluição isocrática até 14 min com 65% de B. Em seguida, o programa retornou as condições iniciais até 15 min e a coluna foi equilibrada por mais 15 min (tempo total de corrida: 30 min).

Nas análises de anabenopeptina e espumigina, extratos das células liofilizadas foram obtidos pela extração com etanol 70% (v/v), repetida três vezes, em sonicador Omni Soni Disruptor por 1 minuto (Omni International, NW, EUA) e centrifugadas Eppendorf, 5804R Centrifuge (9.500 × *g*, 10 min) (Fisher Scientific, KS, EUA), conforme descrito anteriormente (SANZ et al., 2015). O sobrenadante foi evaporado até a secura com corrente de nitrogênio no concentrador Tecnal, TE-019 (Tecnal, SP, Brasil), ressuspendido em 1 mL de solução metanol:água (1:1 v/v) e submetido a análise de LC-MS/MS.

6 RESULTADOS E DISCUSSÕES

6.1 Sequenciamento, montagem e anotação genômica

Os dados sequenciados com a plataforma MiSeq da Illumina apresentaram resultados satisfatórios e valores de qualidade Phred elevados (Figura 9). O processamento das sequências obtidas gerou um conjunto de dados de aproximadamente 14 GB, dos quais 77% resultaram em sequências de cianobactérias, ou seja, aproximadamente 10,7 GB. O sequenciamento resultou em 52.789.680 leituras, em que o comprimento da sequência mais curta foi de 35 pb e o comprimento da sequência mais longa foi de 301 pb, após o pré-processamento. Sequências mais longas e de maior qualidade são fundamentais para uma montagem de alta confiança, principalmente em casos onde genomas individuais são montados a partir de sequências de DNA de uma cultura uni-cianobacterial, mas não axênica, ou até mesmo quando o genoma apresenta elevado número de sequências repetitivas, como é o caso de algumas cianobactérias (WANG et al., 2012).

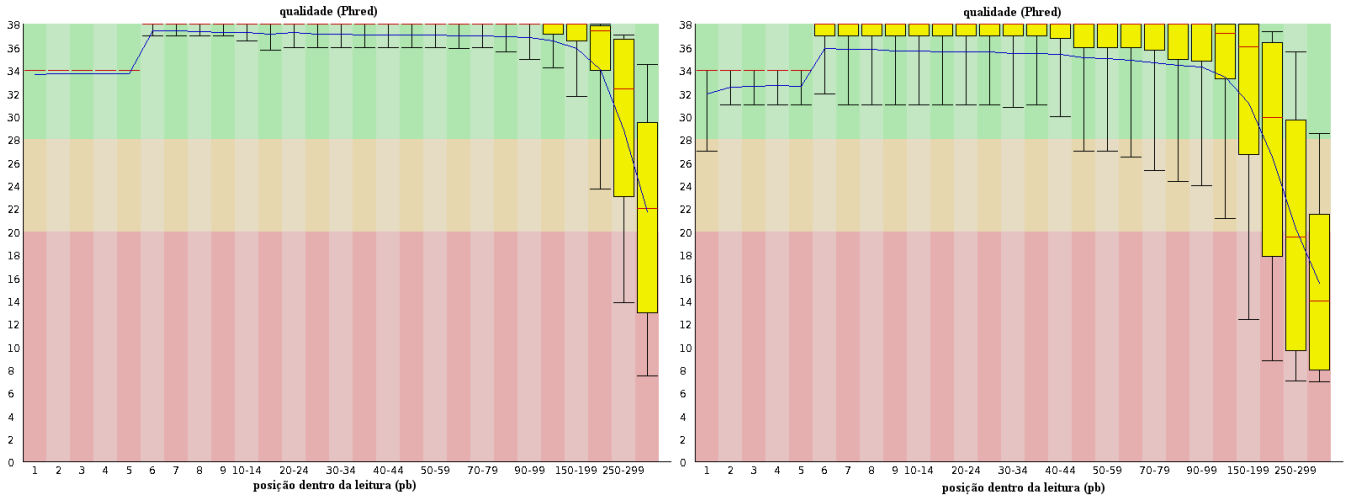


Figura 9 - Qualidade de leituras obtidas em sequenciamento na plafaforma MiSeq (Illumina) do genoma da cianobactéria *S. torques-reginae* ITEP-024. (A) e (B): conjunto dos pares R1 e R2 (respectivamente).

A montagem forneceu 85 sequências contíguas (*contigs*), totalizando 5.211.795 pb, com N50 242.150, maior sequência com 554.390 pb e cobertura média de 843 vezes, resultando, assim, num baixo risco de apresentar quimeras, devido a elevada cobertura (LUO et al., 2011). O conteúdo de GC de 37,42% se mostrou abaixo da média encontrada para táxons da ordem Nostocales, sendo mais próximo da linhagem *Rivularia* sp. PCC 7116, que possui o menor

conteúdo GC dos táxons já sequenciados desta ordem (vide Tabela 2). Como o sequenciamento do genoma da *S. torques-reginae* ITEP-024 é o primeiro dentro do gênero *Sphaerospermopsis*, não foi possível utilizar genoma referência para realizar a comparação. Entretanto, pode-se concluir que as sequências obtidas ao final do processamento de montagem foram satisfatórias, apresentando uma elevada cobertura e confiabilidade, dando suporte e confiança aos resultados obtidos.

A anotação automática dos genes pelo programa Prokka 1.9 (SEEMANN, 2014) resultou na predição de 4.835 genes únicos, 42 tRNA e 1 tmRNA. A anotação manual feita no programa Artemis (RUTHERFORD et al., 2000) possibilitou a exploração dos *contigs* na busca de potenciais agrupamentos gênicos codificadores de metabólitos secundários, evidenciados pelo programa antiSMASH 3.0.4 (WEBER et al., 2015) utilizando o algoritmo-padrão.

6.2 Detecção de potenciais agrupamentos gênicos envolvidos na síntese de metabólitos secundários

O programa antiSMASH (WEBER et al., 2015) previu a existência de nove agrupamentos gênicos no genoma da linhagem ITEP-024 :dois agrupamentos contendo genes que codificam peptídeo sintetases não ribossômicas (NRPSs), um podendo ser responsável pela síntese de hassalidina e outro pela síntese de espumigina; três agrupamentos contendo genes que codificam policetídeo sintases (PKS), sendo uma do tipo I e duas ainda não classificadas; três agrupamentos de genes relacionados a terpenos e um agrupamento que pode estar envolvido na síntese de cianobactinas. Os resultados são apresentados na Tabela 4.

Tabela 4 - Agrupamentos gênicos preditos no genoma de *Sphaerospermopsis torques-reginae* ITEP-024 pelo programa antiSMASH

Agrupamento	Tamanho (pb)	Categoria	Agrupamento gênico mais similar	Genes Similares (%)
1	73.652	NRPS	Hassalidina	34
2	21.904	Terpeno	Não especificado	-
3	25.031	Cianobactina	Anaciclamida	50
4	52.523	PKS tipo I/outra PKS	Glicolipídeo (heterócitos)	85
5	42.243	PKS tipo I	Não especificado	-
6	23.904	Terpeno	Não especificado	-
7	49.902	PKS tipo I/outra PKS	Glicolipídeo (heterócitos)	57
8	7.005	Terpeno	Não especificado	-
9	95.635	NRPS	Espumigina	83

Uma parte importante dos metabólitos secundários de cianobactérias é peptídica ou possui estruturas peptídicas, sendo a maioria desses oligopeptídeos sintetizados por NRPSs ou híbridos de NRPS/PKS (WELKER; VON DÖHREN, 2006). Aproximadamente 70% das cianobactérias possuem essas vias de biossíntese de produtos naturais, cujos genes totalizam por volta de 5% do genoma, com uma média de cinco agrupamentos de NRPS e/ou PKS por genoma (SHIH et al., 2013). No genoma em estudo foram encontrados seis agrupamentos correspondentes a essas mega sintetases, os quais correspondem a aproximadamente 6,55% do genoma montado. Tais valores podem ter sido superestimados devido ao algoritmo utilizado pelo programa antiSMASH; no entanto, ainda estão próximos daqueles descritos pela literatura.

Os metabólitos secundários da classe peptídica podem ser obtidos por modificações pós traducionais de peptídeos sintetizados pela via ribossômica. Essas modificações podem ocorrer pela quebra de ligações responsáveis pela ciclização do peptídeo, por exemplo, pela modificação de aminoácidos da cadeia lateral, ou pela ação de NRPSs (WELKER; VON DÖHREN, 2006).

O sequenciamento do genoma da linhagem *S. torques-reginae* ITEP-024 possibilitou a predição de uma série de agrupamentos gênicos potencialmente envolvidos na síntese de substâncias cujas estruturas ou rotas biossintética são ainda desconhecidas, além de possíveis novas variantes de substâncias já descritas em outras linhagens. As análises com o programa antiSMASH apontou quatro agrupamentos órfãos, os quais podem estar envolvidos na síntese de novas moléculas. A análise com o algoritmo-padrão do programa antiSMASH apontou um grande agrupamento gênico com uma parte dele apresentando alta identidade (83%) de genes da ITEP-024 com genes do agrupamento gênico de espumigina, um inibidor de protease, descrito apenas na *Nodularia spumigena* (FEWER et al., 2009). Nesse grande agrupamento gênico contendo o agrupamento gênico da espumigina (9), foi também detectado pela ferramenta BLAST e pela análise no programa Artemis, o agrupamento gênico da anabenopeptina, também um inibidor de protease, descrito em várias linhagens cianobacterianas, sendo ambos sintetizados por uma NRPS.

6.3 Identificação do agrupamento gênico envolvido na síntese da anatoxina-a(s)

A busca pelo agrupamento gênico codificante da ANS iniciou-se tendo como base os estudos químicos com carbono e nitrogênio marcados que resultaram na proposta de uma rota de biossíntese para esta toxina (MOORE et al., 1992, 1993). Essa via biossintética proposta difere das de outras cianotoxinas já descritas. As toxinas cianobacterianas geralmente são

produzidas pela via não ribossômica, envolvendo um complexo de enzimas com multi-domínios, ou seja, NRPSs, PKSs ou híbridos de NRPS/PKS. Essa via de síntese sempre trabalha para adicionar aminoácidos (NRPS) ou unidades de ácidos carboxílicos (PKS), a partir de uma substância precursora, aumentando a molécula até chegar ao produto final. No entanto, baseado na estrutura química da ANS, não é possível predizer esse tipo de montagem, uma vez que ela não é uma molécula que cresce a partir da adição de aminoácidos ou de unidades carboxílicas, ela é sintetizada a partir da modificação de um único aminoácido, a L-arginina, conforme demonstraram os estudos de Moore et al. (1992; 1993).

A partir dessa premissa, a busca pelo agrupamento gênico de ANS no genoma da *S.torques-reginae* ITEP-024 se iniciou por genes relacionados à síntese de arginina. Um potencial gene da biossíntese de arginina em micro-organismos foi encontrado, *argJ*, que codifica uma proteína denominada ornitina acetiltransferase (OAT). Próximo a este gene (*argJ*) foram encontrados três outros que deram a sustentação necessária para a proposta de biossíntese de ANS apresentada neste trabalho: genes que codificam uma metiltransferase de oxigênio, uma metiltransferase de nitrogênio e uma quinase (família das fosfotransferases). Desta forma, o agrupamento gênico apresentado na Figura 10 foi proposto como potencialmente envolvido na síntese da molécula de ANS, já que possui todas as enzimas necessárias de acordo com as etapas químicas propostas por Moore et al. (1992). Os genes do agrupamento gênico da ANS em *S. torques-reginae* ITEP-024 somam um total de 14,08 kb e foram localizados em um *contig* no genoma desta cianobactéria.

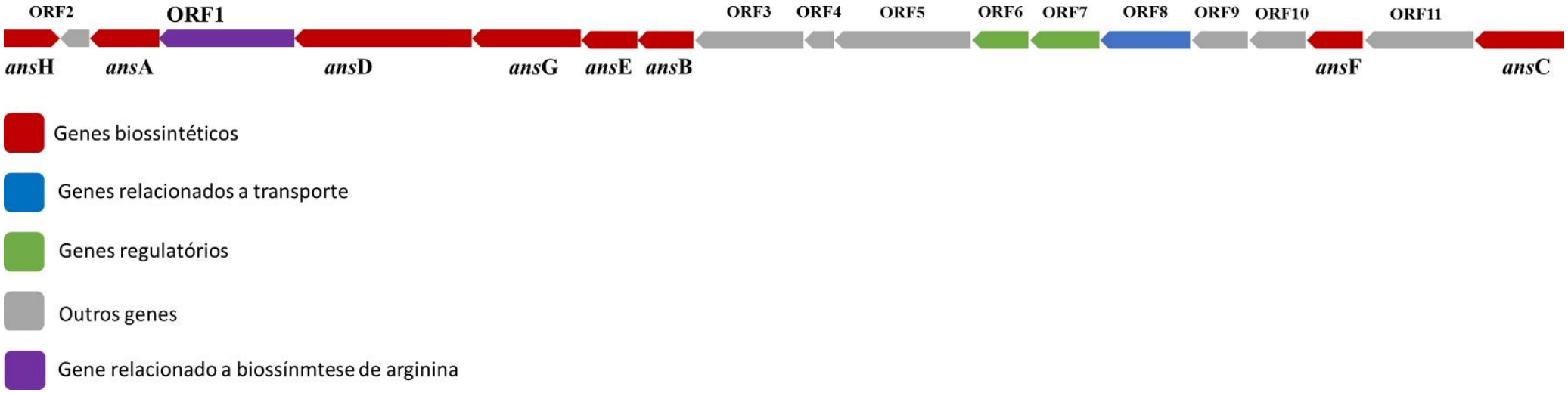


Figura 10 – Representação esquemática proposta do agrupamento gênico da anatoxina-a(s)

O agrupamento gênico é constituído por oito genes biossintéticos e um gene que se acredita estar envolvido com a disponibilização de arginina para a síntese, denominado ORF1. A ORF1 codifica uma ornitina acetiltransferase (OAT), a qual é uma proteína que também participa do quinto passo da síntese da substância denominada ácido clavulânico (ORF4), um potente antibiótico contra infecções bacterianas em humanos que inibe/inativa a enzima β-lactamase, a qual confere resistência às bactérias (TOWNSEND, 2002). A ORF1 de ANS apresentou uma similaridade de sequências de aminoácidos de 35% com a ORF4 da biossíntese de ácido clavulânico em *Streptomyces clavuligerus*. Considerando 20% uma identidade válida para funções homólogas, pode-se afirmar que ambas as proteínas possuem função semelhante. Estudos com o agrupamento gênico do ácido clavulânico sugerem que a baixa identidade da ORF4 (similar a ORF1 de ANS) com a proteína ArgJ (ornitina acetiltransferase – diretamente ligada a síntese de arginina), pode significar que ela esteja relacionada com a modificação de arginina para a produção de ácido clavulânico e não com a síntese de arginina (RODRÍGUES-GARCIA et al., 2000). O mesmo é sugerido para a ORF1 no agrupamento gênico de ANS que estaria envolvida na modificação da arginina e não na sua biossíntese diretamente. Além disso, enzimas semelhantes a OATs, podem estar envolvidas na otimização da produção de arginina, criando uma espécie de “estoque” ou matéria-prima para a biossíntese de ANS (HAMED et al., 2013).

O gene denominado *ansA* codifica uma dioxigenase (*nonheme iron dioxygenase*). Essa proteína (AnsA) apresentou identidade significativa com outras duas proteínas da biossíntese de ácido clavulânico (CS2) e de viomicina (VioC) (Tabela 5), as quais são responsáveis pela hidroxilação na posição beta da L-arginina (THOMAS; CHAN; OZANICK, 2003), exatamente como é proposto para a síntese de ANS. Na síntese de ácido clavulânico a proteína CS2 é uma enzima interativa, atuando em mais de uma etapa da síntese, assim como se propõe na síntese de ANS para a enzima AnsA (TOWNSEND, 2002). A interação entre ORF1 e AnsA dá início ao primeiro passo da biossíntese da ANS – a hidroxilação da L-arginina, formando a 4-hidroxiarginina **(2)** (Figura 11).

Tabela 5 - Comparação entre as sequências de proteínas da AnsA com CS2 e VioC

CS2 (ácido clavulâncio sintase)				
	Streptomyces clavuligerus	*Streptomyces clavuligerus* ATCC 27064	*Actinokineospora spheciospongiae*	*Shigella sonnei*
Identidade (%)	53	53	36	42
Cobertura (%)	10	10	26	18
Número de acesso	Query_119185	Query_80057	Query_194785	Query_12699
VioC (arginina beta-hidroxilase, dependente de Fe(II)/alfa-cetoglutarase)				
	Streptomyces vinaceus ATCC 11861	*Catenulispora acidiphila*	*Saccharothrix mutabilis subsp. capreolus*	
Identidade (%)	50	56	55	
Cobertura	8	6	7	
Número de acesso	Query_75531	Query_25219	Query_204435	

Análises BLAST indicaram que o gene *ansB* codifica uma policetídeo ciclase/desidratase (84% de identidade), além de apresentar identidade com algumas proteínas de membrana. AnsB participa da segunda etapa da síntese, ciclicizando a 4-hidroxiarginina **(2)** e formando o intermediário enduracididina **(3)** (Figura 11). A terceira etapa é a hidroxilação no carbono três (C3) da enduracididina, realizada pela enzima interativa AnsA, formando o intermediário **(4)**. A enduracididina **(3)** e sua versão hidroxilada **(4)** são intermediários encontrados na biossíntese de diversos produtos naturais como mannopeptimicina, viomicina e o antibiótico enduracidina (HATANO et al., 1984; HALTLI et al., 2005; YIN; ZABRISKIE, 2006; BURROUGHS et al., 2013), além de estar presente na biossíntese de ANS. A proteína AnsC codificada pelo gene *ansC* está inclusa na família das treoninas aldolases e participa da quarta etapa da síntese de ANS eliminando uma glicina do intermediário **(4)** e formando o intermediário **(5)**, um aldeído. O intermediário **(5)** é, então, transaminado pela AnsD, que possui identidade com uma amidotransferase (89% de identidade), formando o intermediário **(6)**, o qual é reduzido pela redutase AnsE e forma o intermediário **(7)**. Depois disso, ocorre uma nova hidroxilação no nitrogênio do anel guanidino heterocíclico pela ação da AnsA, formando o intermediário **(8)**. Neste é adicionado o agrupamento fosfato (PO_4^{3-}) pela ação da AnsF, uma quinase (82% de identidade),

e são ainda incorporados os grupos metilas, pela ação da AnsG, metiltransferase de oxigênio (73% de identidade) e AnsH, metiltransferase de nitrogênio (83% de identidade). Ainda não se pode afirmar se o intermediário **(8)** sofre primeiro a metilação no nitrogênio da cadeia lateral ao anel heterocíclico ou se ocorre primeiro a incorporação do grupo fosfato, no entanto, acredita-se que primeiramente é incorporado o fosfato no intermediário **(8)**. Devido ao fato do fosfato ser um grupo muito impedido estericamente, pela presença de quatro oxigênios, seria mais fácil ele ser introduzido numa molécula onde não há dois grupos metilas, os quais também apresentam um impedimento estérico alto, do que ser incorporado quando já estão presentes as duas metilas no nitrogênio da cadeia lateral ao anel. Ao final da ação dessas três enzimas ocorre a formação da molécula **(9)**, a anatoxina-a(s).

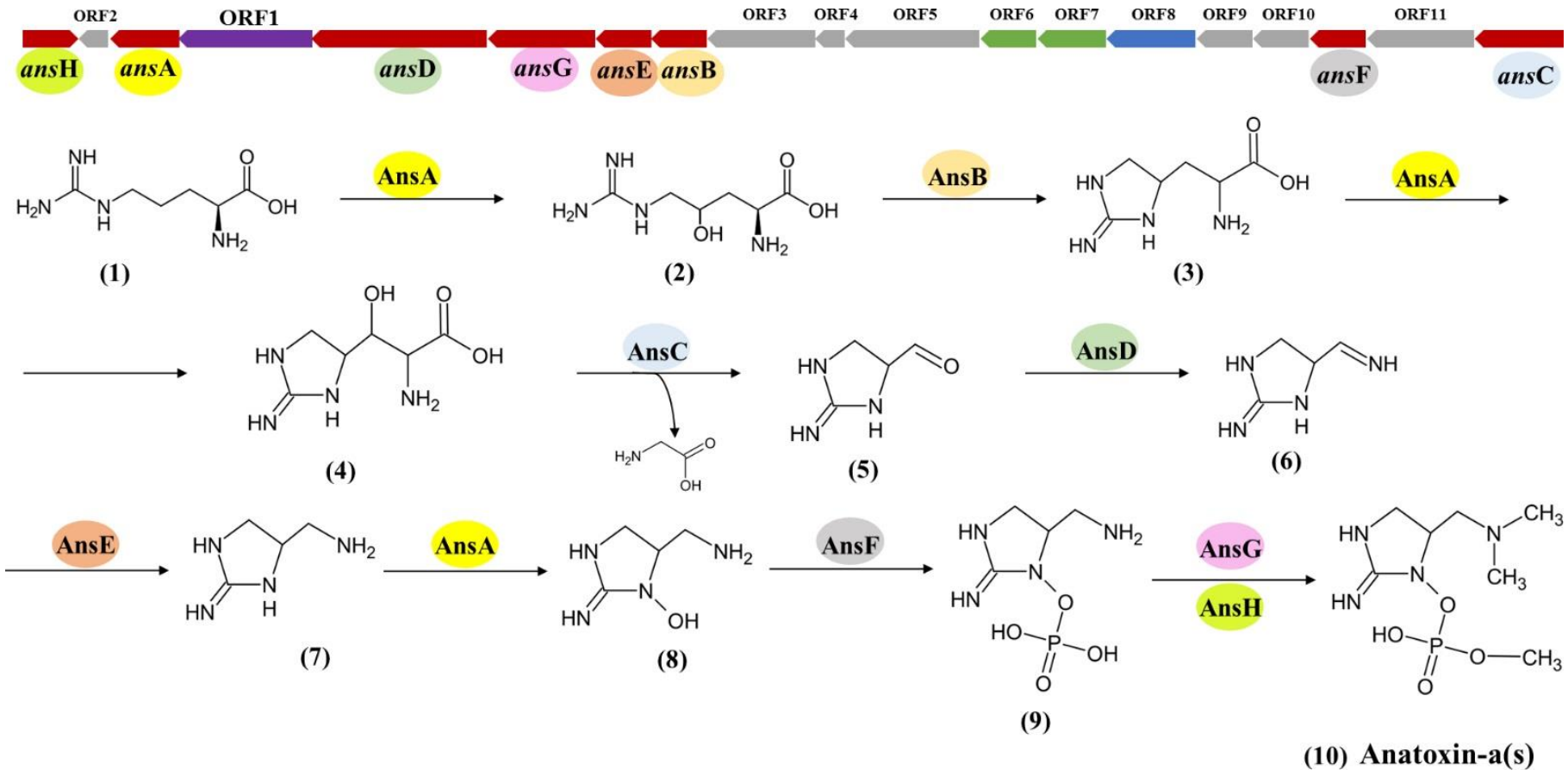


Figura 11 – Proposta de via de biossíntese da anatoxina-a(s) da linhagem *S. torques-reginae* ITEP-024 baseada em enzimas preditas em seu genoma

Além dos genes de biossíntese o agrupamento gênico é composto por doze ORFs, sendo as ORF7 e 8 genes responsáveis pela codificação de um sistema antitoxina e a ORF9 um transportador específico de metabólitos secundários. A Tabela 6 apresenta as funções deduzidas para cada ORF a partir dos resultados da análise BLAST, aliados aos experimentos químicos realizados por Moore et al. (1992) bem como o tamanho das proteínas da síntese, sua homologia com outras cianobactérias. A sequência completa de nucleotídeos do agrupamento gênico de ANS foi comparada com sequências depositadas no banco de dados GenBank do NCBI, não tendo sido encontrado sequências com cobertura e identidade significativa.

Tabela 6 - Funções deduzidas das ORFs do agrupamento gênico de anatoxina-a(s) da linhagem *S. torques-reginae* ITEP-024

Proteína	Aminoácidos	Função Proposta	Proteína homóloga e origem	Similaridade/ Identidade (%)	Número de acesso GenBank
ORF1	413	Ornitina aciltransferase (OAT)	arginine biosynthesis bifunctional protein; ArgJ [*Anabaena* sp. PCC 7108]	82/92	WP_016953551.1
AnsA	148	Hidroxilação da arginina	Dioxygenase [*Anabaena cylindrica*]	84/85	WP_015215026.1
AnsB	149	Ciclização	Cyclase/dehydratase [*Calothrix* sp. PCC 7507]	79/84	WP_015130470.1
AnsC	360	Eliminação de glicina	fructose-bisphosphate aldolase [*Anabaena cylindrica*]	86/89	WP_015215611.1
AnsD	492	Amidotransferase	glutamyl-tRNA amidotransferase subunit B [*Cylindrospermopsis raciborskii*]	83/89	WP_006278766.1
AnsE	145	Redução	Peroxiredoxin [*Anabaena cylindrica*]	85/90	WP_015214435.1
AnsF	182	Transferência de fosfato	adenosylcobinamide kinase/adenosylcobinamide phosphate guanyltransferase [*Anabaena* sp. PCC 7108]	83/82	WP_016952275.1
AnsG	338	Metiltransferase de oxigênio, dependente de SAM	O-methyltransferase [*Microcystis aeruginosa*]	76/73	WP_002735452.1
AnsH	246	Metiltransferase de nitrogênio, dependente de SAM	methyltransferase type 11 [*Nostoc punctiforme*]	76/83	WP_012412600.1

continuação

Proteína	Aminoácidos	Função Proposta	Proteína homóloga e origem	Similaridade/ Identidade (%)	Número de acesso GenBank
ORF2	93	Proteína hipotética	hypothetical protein [*Nostoc azollae*]	88/76	WP_013192879.1
ORF3	483	Dessaturase	zeta-carotene desaturase [*Anabaena cylindrica*]	86/93	WP_015214429.1
ORF4	57	Regulador transicional	CopG-like domain-containing protein DNA-binding [*Anabaena cylindrica* PCC 7122]	89/98	AFZ57792.1
ORF5	544	Proteína hipotética	hypothetical protein [*Anabaena cylindrica*]	82/83	WP_015214427.1
ORF6	140	Sistema toxina-antitoxina	PIN family toxin-antitoxin system [*Mobiluncus mulieris*]	74/41	WP_004017185.1
ORF7	230	Sistema toxina-antitoxina	toxin-antitoxin system antitoxin subunit [*Actinospica robiniae*]	74/28	WP_034267739.1
ORF8	306	Transportador multidroga	multidrug MFS transporter [*Nostoc piscinale* CENA21]	76/76	ALF55933.1
ORF9	49	Proteína hipotética	hypothetical protein Anacy_2330 [*Anabaena cylindrica* PCC 7122]	83/82	AFZ57786.1
ORF10	142	Proteína hipotética	conserved hypothetical protein [*Cylindrospermopsis raciborskii* CS-505]	72/80	EFA71262.1
ORF11	159	Proteína hipotética	hypothetical protein [*Anabaena cylindrica*]	82/76	WP_015214447.1

conclusão

Cianotoxinas como microcistina, nodularina, cilindrospermopsina e saxitoxina, são toxinas muito bem estudadas que já tiveram suas vias elucidadas. Todas essas toxinas foram analisadas por predições funcionais e mutação. A nodularina, por exemplo, foi comparada com a microcistina devido a sua semelhança estrutural, tendo seus genes biossintéticos comparados aos genes de biossíntese de microcistina, para em 2004 ter seu agrupamento gênico sequenciado (MOFFITT; NEILAN, 2001). As atribuições funcionais das enzimas da síntese de nodularina foram baseadas em análises de bioinformática e similaridade com as enzimas de microcistina. Outro exemplo é a saxitoxina, a qual teve um mecanismo de síntese proposto na década de 80 que foi apenas confirmado em 2007, com um estudo realizado *in vitro*, propondo que a síntese se iniciava com uma arginina e uma molécula de acetil-CoA, sendo a primeira via de alcaloide descrita em bactérias (KELLMANN; NEILAN, 2007). A via de biossíntese de ANS na linhagem ITEP-024 segue os mesmos métodos das análises feitas anteriormente para outras toxinas cianobacterianas, o sequenciamento do genoma e a predição funcional de genes e enzimas pela comparação e identidade de sequências homologas presentes em outras vias biossintéticas. A maioria dos agrupamentos gênicos de cianotoxinas e de outros metabólitos secundários de cianobactérias ainda é hipotética.

Os resultados apresentados neste trabalho sobre o agrupamento gênico de ANS propiciam informações inéditas e fundamentais da elucidação da biossíntese da ANS. Os recentes avanços da biologia molecular proporcionam grandes possibilidades de elucidação e comprovação de vias de biossíntese de metabólitos secundário e toxinas, que apontam para a utilização de expressão heteróloga de enzimas e agrupamento gênicos completos, além da purificação, cristalização e estudo dessas proteínas constituintes da síntese, a fim de estudar mais a fundo a relação entre enzima, substrato e produto formado. Dessa forma, as etapas propostas e a comprovação da via de biossíntese da ANS serão realizadas em um esudo futuro. As enzimas chaves da síntese, como, AnsA, AnsF, AnsG e AnsH, bem como a ORF1, entre outras enzimas, terão suas funções e atividades testadas por meio da expressão heteróloga, purificação e cristalização, com o objetivo de estudar a interação entre enzima e substrato e comprovar a formação dos intermediários da rota. A comprovação da formação dos intermediários poderá ser avaliada por análises químicas, como espectrometria de massas, além de testes bioquímicos *in vitro* realizados com as enzimas que comprovam a formação do produto esperado.

6.4 Identificação do agrupamento gênico envolvido na síntese da espumigina

Espumiginas são tetrapeptídeos cíclicos lineares, antes classificados como sendo da família das aeruginosinas pela semelhança estrutural entre as duas moléculas (LIU et al., 2015). Esses peptídeos contêm na parcela N-terminal um ácido D-hidroxifenilacético, uma prolina metilada e um derivado aldeído de arginina na porção C-terminal (FEWER et al., 2009). A presença do agrupamento gênico da espumigina na ITEP-024 foi detectado pelo resultado obtido usando o programa antiSMASH e a alta similaridade (92%) do agrupamento gênico encontrado na ITEP-024 com o agrupamento gênico da espumigina da cianobactéria *Nodularia spumigena* CCY 9414, descrito por Fewer at al. (2009). A tabela 7 compara o tamanho das proteínas codificadas pelos genes do agrupamento gênico de espumigina da ITEP-024 e da CCY 9414, além da similaridade das proteínas das duas linhagens.

Tabela 7 - Comparação de polipeptídeos codificados pelos agrupamentos gênicos de espumiginas baseada no comprimento da proteína codificada (aa) e similaridade da proteína com *S. torques reginae* ITEP-024

Linhagens	**SpuA**	**SpuB**	**SpuC**	**SpuD**	**SpuE**	**SpuF**	**SpuG**
Sphaeropermopsis torques-reginae ITEP-024	1409	4161	264	367	273	700	204
Nodularia spumigena CCY 9414	1408	4171	264	375	271	683	-
	80%	86%	97%	96%	93%	92%	

O agrupamento gênico da espumigina da *S. torques-reginae* ITEP-024 é composto por sete genes (*spu A – G*), sendo dois genes de biossíntese que codificam um híbrido de NRPS/PKS (Figura 12, Tabela 8). O gene *spuA* é constituído por um módulo composto por três domínios, dos quais o domínio de adenilação (A) apresentou, segundo o antiSMASH, afinidade para três diferentes aminoácidos, não chegando a um consenso. O gene *spuB* apresenta três módulos, cada módulo com um domínio A com afinidade para os substratos tirosina, prolina e arginina, respectivamente. Além dos domínios A, os módulos possuem domínios de condensação (C),

epimerização (E), proteína carreadora de peptidil (PCP), tioéster redutase (TD) e cetorredutase (KR), sendo este último um domínio comumente encontrado em PKS, o que confere a característica híbrida a essa via biossintética.

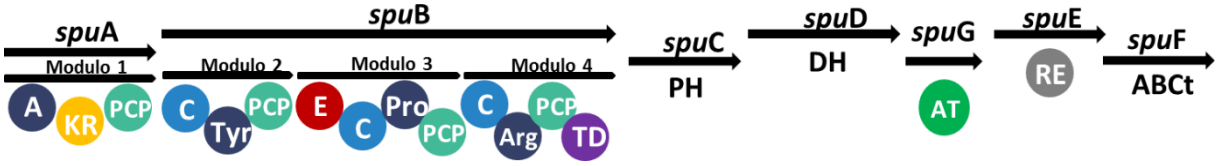


Figura 12 - Representação esquemática do agrupamento gênico da espumigina. Abreviações: A: adenilação; KR: cetorredutase; PCP: proteína carreadora de peptidil; C: condensação; Tyr: tirosina; E: epimerização; Pro: prolina; Arg: arginina; TD: tioéster redutase; PH: proteína hipotética; DH: desidrogenase; AT: aciltransferase; RE: redutase

Tabela 8 - Funções propostas das proteínas codificadas pelo agrupamento gênico de biossíntese de espumigina na linhagem *S. torques-reginae* ITEP-024 e maiores valores de BLAST de proteína

Proteína	**Tamanho (aminoácidos)**	**Funções**	**Organismo**	**Similaridade (%)**	**Número de acesso**
SpuA	1409	NRPS-PKS (A-KR-PCP)	*Aphanizomenon flos-aquae*	85%	WP_051424408.1
SpuB	4161	NRPS (C-A-PCP-E-C-A-PCP-C-A-PCP-TD)	*Nodularia spumigena*	86%	WP_006198652.1
SpuC	264	Desconhecida	*Nodularia spumigena*	97%	WP_006198651.1
SpuD	367	Desidrogenase	*Nodularia spumigena*	96%	WP_006198650.1
SpuE	273	Redutase	*Nodularia spumigena*	93%	WP_006198649.1
SpuF	700	Transportador ABC	*Aphanizomenon flos-aquae*	80%	WP_027403938.1
SpuG	204	Acetiltransferase	*Microcystis aeruginosa*	76%	WP_002796354.1

O gene *spuC* codifica uma proteína hipotética sem função conhecida, *spuE* codifica uma enzima redutase, *spuF*, um transportador do tipo ABC, e *spuG* codifica uma acetiltransferase dependente de coenzima A (CoA). O gene *spuG* não é encontrado no agrupamento gênico de espumigina da *Nodularia spumigena*, sendo um gene exclusivo da linhagem *S. torques-reginae* ITEP-024. A variante de espumigina produzida pela linhagem em estudo também se apresentou diferente das variantes já conhecidas. Estudos de espectrometria de massas realizados por Sanz

et al. (2015) apontaram quatro variantes para espumigina produzidas pela ITEP-024, em que o primeiro aminoácido da linha de montagem seria o Hhpba (*hydroxy-(hydroxyphenyl)butyric acid*) ao invés do Hpla (*4-hydroxyphenylpiruvic acid*) comprovado por Fewer et al. (2009), em ensaios bioquímicos de ATP-PPi (*ATP-pyriphosphate Exchange*).

Pelas informações obtidas nas análises de bioinformática, aliadas às análises químicas por espectrometria de massas da espumigina produzida pela ITEP-024 (SANZ et al., 2015), pode-se presumir que a proteína SpuG, exclusiva da linhagem *S. torques-reginae* ITEP-024, está envolvida na modificação do primeiro substrato do domínio de adenilação do módulo 1, resultando em uma variante nova para a molécula de espumigina da ITEP-024. Sabe-se que as proteínas SpuC, D e E atuam na biossíntese da metilprolina, um aminoácido raro não proteinogênico, encontrado em duas das quatro variantes (LIU et al., 2015; SANZ et al., 2015). A Figura 13 esquematiza a proposta da via de biossíntese da espumigina da ITEP-024, bem como a variante produzida pela linhagem em estudo.

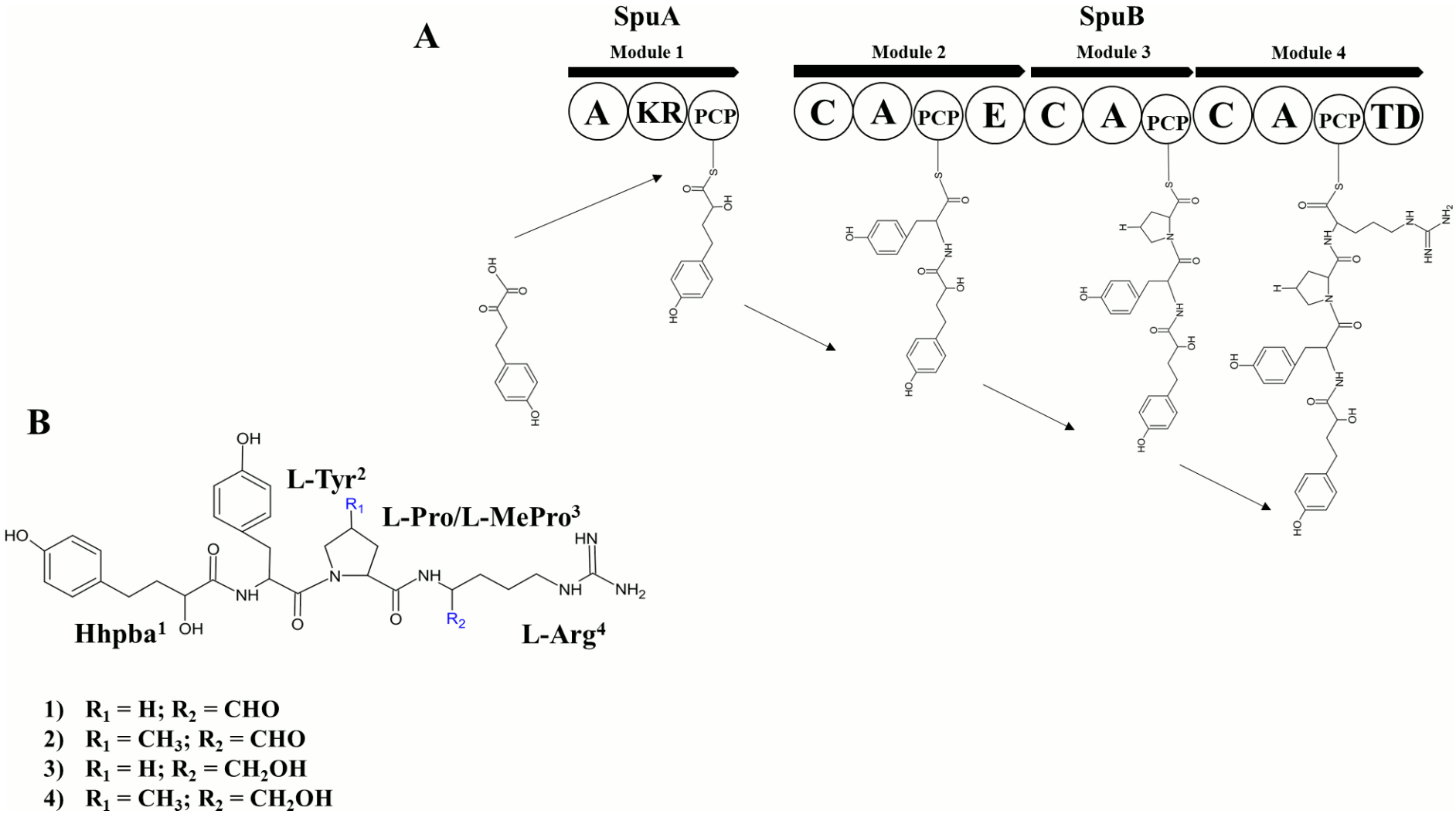


Figura 13 – A: Proposta de biossíntese da espumigina da linhagem *S. torques-reginae* ITEP-024; B: Estrutura da variante produzida pela linhagem *S. torques-reginae* ITEP-024

6.5 Identificação do agrupamento gênico envolvido na síntese da anabenopeptina

Anabenopetinas são hexapeptídeos cíclicos sintetizados por NRPS em cianobactérias. Elas são caracterizadas pela presença de regiões conservadoras como a ligação ureído e a presença do aminoácido D-lisina (CHRISTIANSEN et al., 2011). O agrupamento gênico completo da anabenopeptina foi relatado para linhagens dos gêneros *Anabaena*, *Nodularia*, *Nostoc* e *Planktothrix* (ROUHIAINEN et al., 2000; ROUHIAINEN et al., 2010; ROUNGE et al., 2009), sendo que o presente estudo reporta pela primeira vez este agrupamento gênico no gênero *Sphaerospermopsis*. Esse metabólito é um potente inibidor de protease e algumas espécies podem produzir mais de uma variante devido à capacidade de iniciar a síntese por diferentes módulos de iniciação na NRPS (ROUHIAINEN et al., 2010). As análises *in sílico* e químicas mostraram que a linhagem ITEP-024 produz apenas uma variante desta molécula.

O agrupamento gênico da anabenopeptina possui quatro genes de NRPS (*aptA, aptB, aptC e aptD*) subdivididos em seis módulos, cada um contendo um domínio de adenilação (Figura 14, Tabela 9). Para cada domínio de adenilação foi previsto pelo antiSMASH aminoácidos específicos como substrato: isoleucina, lisina, isoleucina, homotirosina, metilalanina e fenilalanina, respectivamente. Além dos domínios de adenilação (A), cada módulo possui domínios de condensação (C), responsáveis pela formação da ligação peptídica entre o aminoácido já conectado na megasintetase e o aminoácido que está sendo agregado; domínios de proteína carreadora de peptidil (PCP), responsáveis pela formação da ligação tioester que une a cadeia peptídica à NRPS, além de domínios adicionais como, epimerização, metiltransferase e tioesterase, responsáveis pela mudança conformacional da molécula, adição de um grupamento metil e rompimento da ligação tioéster, respectivamente, contribuindo na modificação do produto final. Em adição aos quatro genes de biossíntese, existe um gene (*aptE*) codificador da enzima isopropilmalato sintase e um transportador do tipo ABC (*aptF*). A função da enzima isopropilmalato sintase ainda não foi muito bem definida, porém acredita-se que ela possa estar envolvida na ciclização do peptídeo e na formação da ligação ureído, a qual liga a cadeia lateral de aminoácidos à fração D-lisina (ROUHIAINEN et al., 2010).

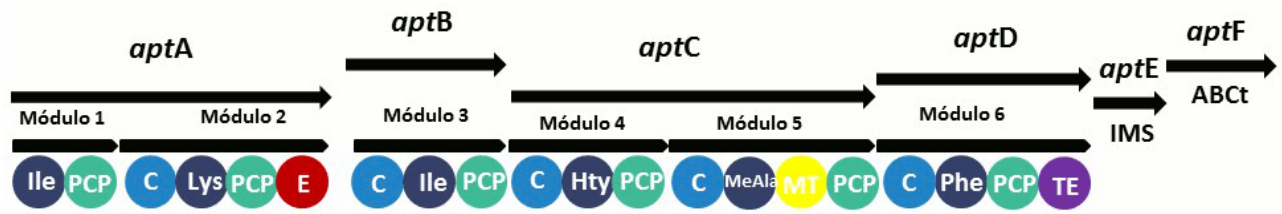


Figura 14 - Representação esquemática do agrupamento gênico de produção de anabenopeptina. Abreviações: Ile: isoleucina; PCP: proteína carreadora de peptidil; C: condensação; Lys: lisina; E: epimerização; Hty: homotirosina; MeAla: metilalanina; MT: metiltranferase; Phe: fenilalanina e TE: tioesterase; IMS: isopropilmalato sintase; ABCt: transportador ABC

Tabela 9. Funções propostas das proteínas codificadas pelo agrupamento gênico da biossíntese da anabenopetina na linhagem *S. torques-reginae* ITEP-024 e maiores valores de BLAST de proteína

Proteína	Tamanho (aminoácidos)	Funções	Organismo	Similaridade (%)	Número de acesso
AptA	2208	NRPS (A-PCP-C-A-PCP-E)	*Nodularia spumigena*	93%	WP_006196339.1
AptB	1069	NRPS (C-A-PCP)	*Aphanizomenon flos-aquae*	89%	WP_027403946.1
AptC	2576	NRPS (C-A-PCP-C-A-MT-PCP)	*Aphanizomenon flos-aquae*	81%	WP_027403947.1
AptD	1415	NRPS (C-A-PCP-TE)	*Aphanizomenon flos-aquae*	80%	WP_027403948.1
AptE	392	-	*Nodularia spumigena*	87%	WP_006196343.1
AptF	803	Transportador ABC	*Nodularia spumigena*	82%	WP_006196345.1

As sequências de nucleotídeos dos agrupamentos gênicos da anabenopeptina da ITEP-024 e de representantes dos gêneros *Anabaena*, *Nodularia*, *Nostoc* e *Planktothrix* foram alinhadas e uma árvore filogenética foi construída (Figura 15). A análise filogenética indicou que a sequência do agrupamento gênico da anabenopeptina da ITEP-024 apresenta um maior grau de parentesco evolutivo com o agrupamento gênico da anabenopeptina da linhagem *Nodularia spumigena* CCY 9414, formando um único clado. Observa-se nessa árvore, a tendência das sequências do agrupamento gênico da anabenopeptina se agruparem de acordo com a descrição taxonômica das cianobactérias, um indicativo de transferência vertical de genes *apt*. A utilização de um número maior de sequências de genes *apt* e também de outros gêneros na análise filogenética poderão comprovar ou não essa tendência no futuro. A tabela 10

compara em termos de tamanho da proteína codificada e identidade dos genes do agrupamento gênico de anabenopeptina das seis linhagens produtoras.

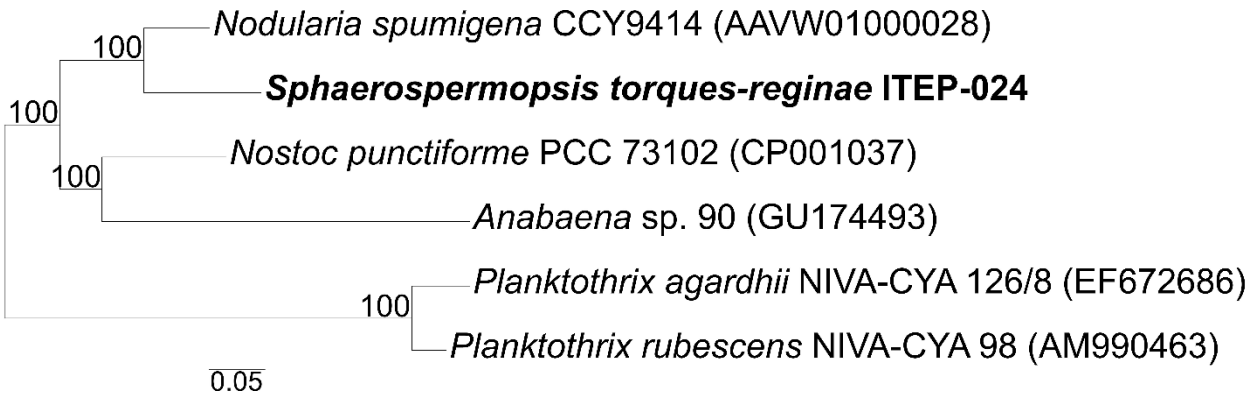


Figura 15 – Árvore filogenética reconstruída por inferência bayesiana com sequências concatenadas de nucleotídeos de seis agrupamentos gênicos de anabenopeptina. A linhagem utilizada neste estudo está em negrito. Probabilidades posteriores estão apresentadas nos nós dos ramos

Tabela 10 - Comparação de polipeptídeos codificados pelos agrupamentos gênicos de anabenopeptinas, incluindo o comprimento da proteína codificada (aa) e similaridade da sequência proteica com a sequência equivalente encontrada em *S. torques reginae* ITEP-024

Linhagens	**$AptA_1$**	**AptA**	**AptB**	**AptC**	**AptD**	**AptE**	**AptF**
Sphaeropermopsis torques-reginae ITEP-024	-	2208	1069	2576	1415	392	803
Anabaena sp. Strain 90	2195 -	2174 68%	1069 78%	2565 77%	1403 77%	392 83%	765 69%
Nodularia spumigena CCY 9414	-	2195 93%	1075 86%	2558 72%	1394 78%	392 87%	738 77%
Nostoc punctiforme PCC 73102	-	2181 66%	1076 78%	2578 70%	1406 72%	392 77%	675 70%
Planktothrix agardhii NIVA CYA 1268	-	2166 54%	1078 64%	2574 61%	1420 58%	678 50%	-
Planktothrix rubescens NIVA CYA 98	-	2166 54%	1078 64%	2575 64%	1403 59%	678 50%	-

Anabenopeptinas foram descritas em cianobactérias isoladas de vários hábitats, como água doce (HARADA et al., 1995), terrestre (RESHEF; CARMELI, 2002), água salobra (FUJII et al., 1997a) e de esponjas marinhas, cujas moléculas receberam o nome de cobamida e ceramida A (KOBAYASHI et al., 1991). Acredita-se que cianobactérias possam ser as produtoras das variantes encontradas em esponjas, já que elas podem viver em simbiose com outros organismos (PELE, 2004). A ampla diversidade estrutural das anabenopeptinas proporciona potencial para sua aplicação terapêutica em fármacos antirretrovirais, por exemplo, em que haverá a inibição de determinadas proteases. O aminoácido localizado na porção C-terminal da ligação ureído confere a especificidade para a inibição das proteases (CHERUKU et al., 2012). A Figura 16 esquematiza a rota biossintética proposta para a variante da ITEP-024 a partir da análise dos domínios de NRPS presentes no agrupamento gênico de biossíntese.

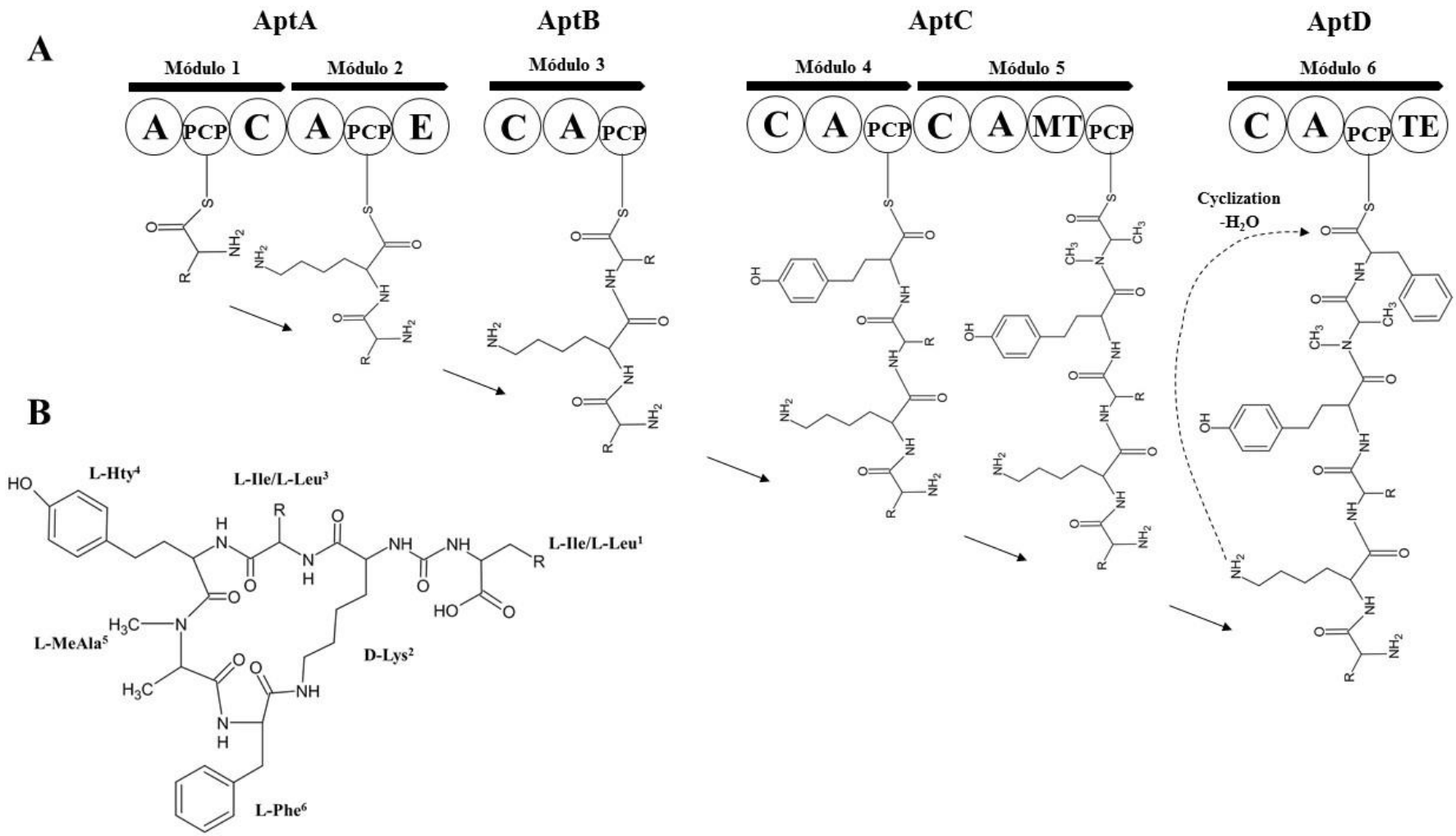


Figura 16 – A: Proposta de biossíntese da anabenopeptina da linhagem *S. torques-reginae* ITEP-024; B: Estrutura da variante produzida pela linhagem *S. torques-reginae* ITEP-024

A análise química por espectrometria de massas corroborou com o resultado da análise de bioinformática, mostrando que a variante produzida pela ITEP-024 (Figura 17) é a mesma variante já descrita para a *Nodularia spumigena* NSOR-02 (MAZUR-MARZEC et al., 2013), muito provavelmente justificando o resultado da proximidade filogenética dos agrupamentos gênicos dessas duas cianobactérias. No entanto, não foi possível prever, apenas por análise de LC-MS/MS, se os aminoácidos das posições 1 e 3 são isoleucina ou leucina, já que são isômeros e possuem a mesma *m/z*. A molécula de anabenopeptina da *S. torques-reginae* ITEP-024 apresentou um valor de *m/z* do íon protonado igual a 808 em um tempo de retenção de 33,7 minutos, em concordância com a variante já descrita por Mazur-Marzec et al. (2013).

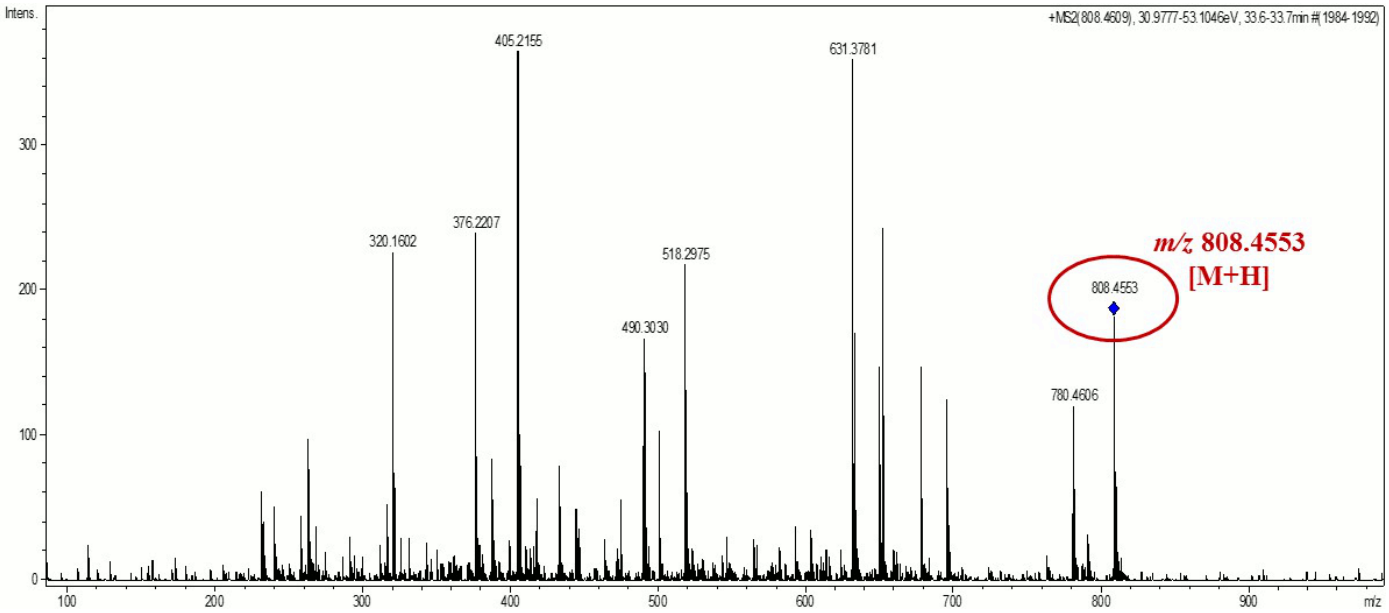


Figura 17 – Análise por LC-MS/MS do extrato da cultura de *S. torques-reginae* ITEP-024. Espectro de íon produzido para anabenopeptina. O primeiro pico à direita representa o íon molecular com sinal de *m/z* de 808,4553 ($[M+H]^+$)

6.6 Interação entre os agrupamentos gênicos de anabenopeptina e espumigina

Na linhagem *S. torques-reginae* ITEP-024 os agrupamentos gênicos da anabenopeptina e espumigina estão localizados no mesmo *contig*, separados numa distância de aproximadamente 12 kb. A proximidade desses dois agrupamentos gênicos sugere algum tipo de interação, a qual deverá ser posteriormente verificada. Acredita-se que ambos os

agrupamentos gênicos estejam compartilhando um domínio de adenilação com especificidade para tirosina, no caso da síntese de espumigina, e de homotirosina, no caso da síntese de anabenopeptina, mostrando uma localização que possivelmente apresenta relevância biológica para esses agrupamentos (Figura 18).

Ao lado do gene *aptA* encontram-se dois genes pertencentes ao agrupamento gênico de biossíntese de homofenilalanina, *hphB*, *hphCD*, e o primeiro gene dessa rota biossintética encontra-se no próprio agrupamento gênico de anabenopeptina, sendo o *hphA* equivalente ao gene *aptE*, exatamente como descrito por Koketsu et al. (2013). A ocorrência do gene envolvido na síntese dessa molécula de homofenilalanina próxima ao agrupamento gênico de espumigina sugere a incorporação dessa molécula no agrupamento *spu*, após sofrer modificações necessárias por enzimas, formando o aminoácido de interesse para o agrupamento gênico *spu*, Hhpba. O compartilhamento de genes entre esses dois agrupamentos gênicos pode sugerir um mecanismo de interação entre os agrupamentos gênicos para formar metabólitos secundários híbridos de interesse para o organismo, semelhante ao descrito por Liu et al. (2015), hipótese essa que será avaliada posteriormente.

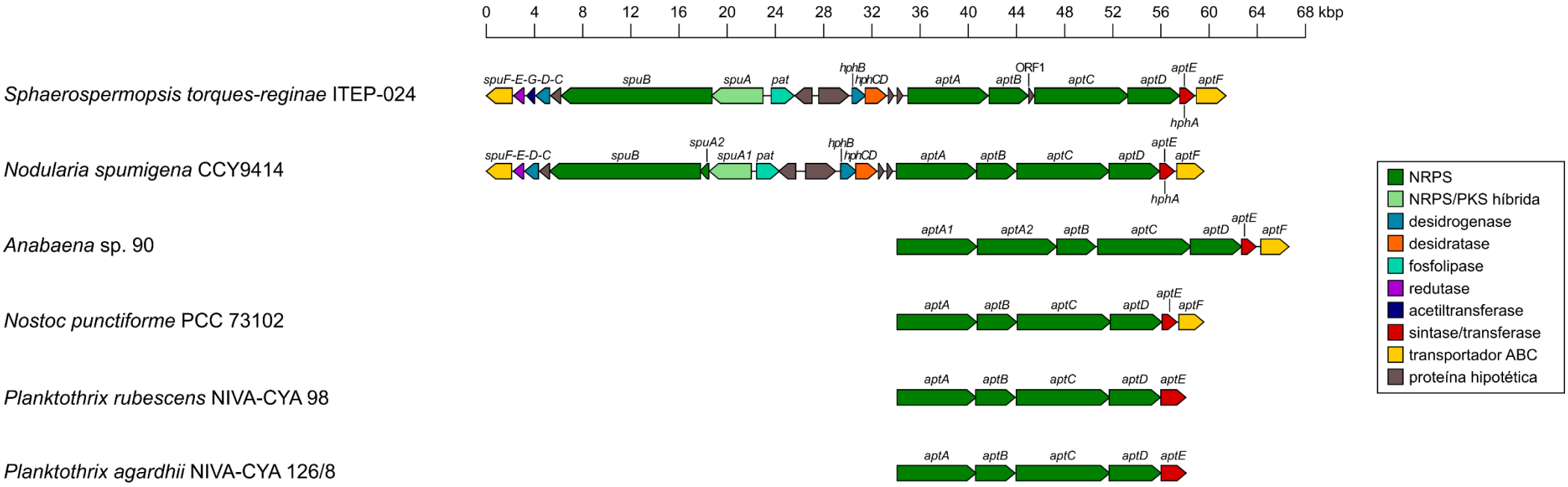


Figura 18 – Estrutura dos agrupamentos gênicos presumidamente relacionados à síntese de espumigina e anabenopeptina em *Sphaerospermopsis torques-reginae* ITEP-024 e em cianobactérias.

7 CONCLUSÕES

- Neste estudo foi gerado o genoma incompleto da cianobactéria *S. torques-reginae* ITEP-024, obtendo no final da montagem 85 sequências contíguas (*contigs*), correspondendo a um genoma incompleto de tamanho de 5.211.795 pb com cerca de 843 × de cobertura, o que possibilitou a mineração de agrupamentos gênicos envolvidos na síntese de produtos naturais bioativos.

- A mineração do genoma por genes envolvidos na biossíntese da ANS resultou na proposta de um agrupamento gênico encontrado dentro de um *contig* que codifica para enzimas cujas funcionalidades atendem a proposta de via biossintética sugerida por Moore e colaboradores. As sequências de aminoácidos das enzimas encontradas mostraram altas similaridades com enzimas de outras vias biossintéticas envolvidas em reações catalíticas semelhantes. A via biossintética de ANS proposta neste estudo envolve a síntese peptídica ribossômica e é diferente de todas as outras vias biossintéticas de cianotoxinas já descritas. Dessa forma, o presente estudo apresenta um mecanismo enzimático inédito e abre caminhos para a mineração de ANS em outros genomas cianobacterianos e a produção de análogos por engenharia molecular.

- A mineração do genoma da *S. torques-reginae* ITEP-024 por outros agrupamentos gênicos envolvidos na síntese de metabólitos secundários identificou dois agrupamentos relacionados com a síntese de duas moléculas conhecidas, a espumigina e a anabenopeptina, que possuem atividade de inibição de protease. O agrupamento gênico da espumigina encontrado mostrou um gene adicional em comparação com o único agrupamento gênico de espumigina já descrito da *Nodularia spumigena* CCY9414, e é responsável pela síntese de quatro novas variantes de espumigina, conforme determinado por espectrometria de massas em um estudo prévio. O agrupamento gênico da anabenopeptina anotado na linhagem ITEP-024 é responsável pela produção de uma variante já conhecida, sendo esta variante predita por análise *in silico* do domínio de adenilação e também por LC-MS/MS. A proximidade desses dois agrupamentos gênicos (*spu* e *apt*) nos genomas da *S. torques-reginae* ITEP-024 e *N. spumigena* CCY9414 foram descobertas neste estudo e sugere algum tipo de interação, a qual deverá ser posteriormente verificada.

- Os três agrupamentos gênicos (*ans*, *spu* e *apt*) encontrados no genoma da *S. torques-reginae* ITEP-024 tiveram seus produtos detectados por LC-MS/MS nos extratos celulares deste organismo, sendo que a produção da anabenopepetina foi determinada no presente estudo e a ANS e espumigina em estudos prévios.

- O presente estudo relata o primeiro sequenciamento genômico de uma linhagem de cianobactéria do gênero *Sphaerospermopsis*, assim como descreve pela primeira vez três agrupamentos gênicos (*ans*, *spu* e *apt*) envolvidos na síntese de metabólitos secundários bioativos neste gênero. O agrupamento gênico *ans* descrito é inédito, sendo que a elucidação da maquinaria molecular envolvida na produção de ANS e o conhecimento dos genes codificantes para essa toxina representa uma importante ferramenta para estudos da presença de cianobactérias potencialmente tóxicas em águas usadas para abastecimento público.

REFERÊNCIAS

ARÁOZ, R.; MOLGÓ, J.; DE MARSAC, N.T. Neurotoxic cyanobacterial toxins. **Toxicon**, Oxford, v. 56, n. 5, p. 813–828, 2010.

ARNISON, P.G.; BIBB, M.J.; BIERBAUM, G.; BOWERS, A.A.; BUGNI, T.S.; BULAJ, G.; CAMARERO, J.A.; CAMPOPIANO, D.J.; CHALLIS, G.L.; CLARDY, J.; COTTER, P.D.; CRAIK, D.J.; DAWSON, M.; DITTMANN, E.; DONADIO, S.; DORRESTEIN, P.C.; ENTIAN, K.D.; FISCHBACH, M.A.; GARAVELLI, J.S.; GÖRANSSON, U.; GRUBER, C.W.; HAFT, D.H.; HEMSCHEIDT, T.K.; HERTWECK, C.; HILL, C.; HORSWILL, A.R.; JASPARS, M.; KELLY, W.L.; KLINMAN, J.P.; KUIPERS, O.P.; LINK, A.J.; LIU, W.; MARAHIEL, M.A.; MITCHELL, D.A.; MOLL, G.N.; MOORE, B.S.; MÜLLER, R.; NAIR, S.K.; NES, I.F.; NORRIS, G.E.; OLIVERA, B.M.; ONAKA, H.; PATCHETT, M.L.; PIEL, J.; REANEY, M.J.; REBUFFAT, S.; ROSS, R.P.; SAHL, H.G.; SCHMIDT, E.W.; SELSTED, M.E.; SEVERINOV, K.; SHEN, B.; SIVONEN, K.; SMITH, L.; STEIN, T.; SÜSSMUTH, R.D.; TAGG, J.R.; TANG, G.L.; TRUMAN, A.W.; VEDERAS, J.C.; WALSH, C.T.; WALTON, J.D.; WENZEL, S.C.; WILLEY, J.M.; VAN DER DONK, W.A. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. **Natural Product Report**, Cambridge, v. 30, p. 108–160, 2012.

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.

BARRIOS-LLERENA, M.E.; BURJA, A.M.; WRIGHT, P.C. Genetic analysis of polyketide synthase and peptide synthetase genes in cyanobacteria as a mining tool for secondary metabolites. **Journal of Industrial Microbiology and Biotechnology**, Heidelberg, v. 34, p. 443-456, 2007.

BECKER, V.; IHARA, P.; YUNES, J.S.; HUSZAR, V.L.M. Occurrence of anatoxin-a(s) during a bloom of Anabaena crassa in a water-supply reservoir in southern Brazil. **Journal of Applied Phycology**, Dordrecht, v. 22, p. 235-241, 2010.

WELKER, M.; DÖHREN, H. Cyanobacterial peptides: nature's own combinatorial biosynthesis. **FEMS Microbiology Reviews**, Amsterdam, v. 30, n. 4, p. 530-563, 2006.

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

BURROUGHS, A.M.; HOPPE, R.W.; GOEBEL, N.C.; SAYYED, B.H.; VOEGTLINE, T.J.; SCHWABACHER, A.W.; ZABRISKIE, T.M.; SILVAGGI, N.R. Structural and functional characterization of MppR, an enduracididine biosynthetic enzyme from *Streptomyces hygroscopicus*: functional diversity in the acetoacetate decarboxylase-like superfamily. **Biochemistry**, Washington, DC, v.52, p. 4492-4506, 2013.

CAO, H.; SHIMURA, Y.; MASANOBU, K.; YIN, Y. Draft genome sequence of the toxic bloom-forming cyanobacterium *Aphanizomenon flos-aquae* NIES-81. **Genome Announcements**, Washington, DC, v. 2, e00044-14, 2014.

CARMICHAEL, W.W.; GORHAM, P.R. Anatoxins from clones of Anabaena flos-aquae isolated from lakes of Western Canada. **Mitteilungen Internationale Vereinigung für Theoretische und Angewandte Limnologie**, Heidelberg, v. 21, p. 285-295, 1978.

CHERUKU, P.; PLAZA, A.; LAURO, G.; KEFFER, J.; LLOYD, J.R.; BIFULCO, G.; BEWLEY, C.A. Discovery and synthesis of namalide reveals a new anabaenopeptin scaffold and peptidase inhibitor. **Journal of Medical Chemistry**, Washington, DC, v. 55; n. 2, p. 735-742.

CHORUS, I.; BARTRAM, J. **Toxic cyanobacteria in water, a guide to their public health consequences, monitoring and management**. London: E & FN Spon, 1999. 416 p.

CHRISTIANSEN, G.; PHILMUS, B.; HEMSCHEIDT, T.; KURMAYER, R. Genetic variation of adenylation domains of the anabaenopeptin synthesis operon and evolution of substrate promiscuity. **Journal of Bacteriology**, Washington, DC, v. 193, n. 15, p. 3822-3831, 2011.

ČOLOVIĆ, M.B.; KRSTIĆ, D.Z.; LAZAREVIĆ-PAŠTI, T.D.; BONDŽIĆ, A.M.; VASIĆ, V.M. Acetylcholinesterase Inhibitors: Pharmacology and Toxicology. **Current Neuropharmacology**, London, v. 11, p. 315-335, 2013.

COOK, W.O.; BEASLEY, V.R.; DAHLEM, A.M.; DELLINGER, J.A.; HARLIN, K.S.; CARMICHAEL, W.W. Comparison of effects of anatoxin-a(s) and paraoxon, physostigmine and pyridostigmine on mouse brain cholinesterase activity. **Toxicon**, Oxford, v. 26, n. 8, p. 750–753, 1988.

COOK, W.O.; BEASLEY, V.R.; LOVELL, R.A.; DAHLEM, A.M.; HOOSER, S.B.; MAHMOOD, N.A.; CARMICHAEL, W.W. Consistent inhibition of peripheral cholinesterases by neurotoxins from the freshwater cyanobacterium *Anabaena flos-aquae*: Studies of ducks, swine, mice and a steer. **Environmental Toxicology and Chemistry**, New York, v. 8, p. 915-922, 1989.

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

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

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.

DAS, A.; PANDA, A.; SINGH, D.; CHANDRABABUNAIDU, M.M.; MISHRA, G.P.; BHAN, S.; ADHIKARY, S.P.; TRIPATHY, S. Deciphering the genome sequences of the

hydrophobic cyanobacterium *Scytonema tolypothrichoides* VB-61278. **Genome Announcements**, Washington, DC, v. 3, p. 1-2, 2015.

DITTMANN, E.; GUGGER, M.; SIVONEN, K.; FEWER, D.P. Natural product biosynthetic diversity and comparative genomics of the cyanobacteria. **Trends in Microbiology**, Cambridge, v. 23, n. 10, p. 642-652, 2015.

DITTMANN, E.; NEILAN, B.A.; BÖRNER, T. Molecular biology of peptide and polyketide biosynthesis in cyanobacteria. **Applied Microbiology and Biotechnology**, Heidelberg, v. 57, n. 4, p. 467–473, 2001.

DITTMANN, E.; WIEGAND, C. Cyanobacterial toxins: occurrence, biosynthesis and impact on human affairs. **Molecular Nutrition & Food Research**, Weinheim, v. 50, n. 1, p. 7–17, 2006.

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

DÖRR, F.A.; RODRÍGUEZ, V.; MOLICA, R.; HENRIKSEN, P.; KROCK, B.; PINTO, E. Methods for detection of anatoxin-a(s) by liquid chromatography coupled to electrospray ionization-tandem mass spectrometry. **Toxicon**, Oxford, v. 55, n. 1, p. 92–99, 2010.

EKMAN, M.; PICOSSI, 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.

FEWER, D.P; JOKELA, J.; ROUHIAINEN, L.; WAHLSTEN, M.; KOSKENNIEMI, K.; STAL, L.J.; SIVONEN, K. The non-ribosomal assembly and frequent occurrence of the protease inhibitors spumigins in the bloom-forming cyanobacterium *Nodularia spumigena*. **Molecular Microbiology**, Oxford, v. 73, n. 5, p. 924-937, 2009.

FINKING, R.; MARAHIEL, M.A. Biosynthesis of nonribosomal peptides. **Annual Review of Microbiology**, Palo Alto, v. 58, p. 453-488, 2004.

FINN, R.D.; COGGILL, P.; EBERHARDT, R.Y.; EDDY, S.R.; MISTRY, J.; MITCHELL, A.L.; POTTER, S.C.; PUNTA, M.; QURESHI, M.; SANGRADOR-VEGAS, A.; SALAZAR, G.A.; TATE, J. The Pfam protein families database: towards a more sustainable future. **Nucleic Acids Research**, London, v. 44, p. D279-D285, 2016.

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

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

FUJII, K.; SIVONEN, K.; ADACHI, K.; NOGUCHI, K.; HIROSHI SANO, H.; HIRAYAMA, K.; SUZUKI, M.; HARADA, K. Comparative study of toxic and non-toxic cyanobacterial products: novel peptides from toxic *Nodularia spumigena* AV1. **Tetrahedron Letters**, Oxford, v. 39, n. 31, p. 5525-5528, 1997a.

GORHAM, P.R.; MCLACHLAN, J.R.; HAMMER, V.T.; KIM, W.K. Isolation and culture of toxic strains of *Anabaena flos-aquae* (Lyngb.) Breb. Verhandlungen der Internationalen Vereiningung fuer Theoretische und Angewandte. **Limnologie**, Tolouse, v. 15, p. 796-804. 1964.

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.

GUIMARÃES, P.I.; LEÃO, T.F.; MELO, A.G.C.; RAMOS, R.T.J.; SILVA, A.; FIORE, M.F.; SCHNEIDER, M.P.C. Draft genome of the picocyanobacterium *Synechococcus* sp. strain GFB01, isolated from a freshwater lagoon in the Brazilian Amazon. **Genome Announcements**, Washington, DC, v. 3, n. 4, e00876-15, 2015.

HALTLI, B.; TAN, Y.; MAGARVEY, N.A.; WAGENAAR, M.; YIN, X.; GREENSTEIN, M.; HUCUL, J.A.; ZABRISKIE, T.M. Investigating β-hydroxyenduracididine formation in the biosynthesis of the mannopeptimycins. **Chemistry & Biology**, New York, v. 12, p. 1163-1168, 2005.

HAMED, R.B.; GOMEZ-CASTELLANOS, J.R.; HENRY, L.; DUCHO, C.; MCDONOUGHA, M.A.; SCHOFIELD, C.J. The enzymes of b-lactam biosynthesis. **Natural Product Report**, Cambridge, v. 30, p. 21-107, 2013.

HARADA, K.; FUJII, K.; SHIMADA, T.; SUZUKI, M. Two Cyclic Peptides, Anabaenopeptins, a third group of bioactive compounds from the cyanobacterium *Anabaena flos.aquae* NRC 525-17. **Tetrahedron Letters**, Oxford, v. 36, n. 9, p. 1511-1514, 1995.

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.

HATANO, K.; NOGAMI, I.; HIGASHIDE, E.; KISHI, T. Biosynthesis of enduracidin: origin of enduracididine and other amino acids. **Agricultural and Biological Chemistry**, Tokyo, v. 48, n. 6, p. 1503-1508, 1984.

HEMSCHEIDT, T.; BURGOYNE, D. L.; MOORE, R.E. Biosynthesis of anatoxin-a(s). (2S,4S)-4-hydroxyarginine as an intermediate. **Journal of the Chemical Society, Chemical Communications**, London, p. 205-206, 1995.

HEMSCHEIDT, T.; RAPALA J.; SIVONEN, K.; SKULBERG, O.M. Biosynthesis of anatoxin in *Anabaena flos-aquae* and homoanatoxin-a in *Oscillatoria formosa*. **Journal of Chemical Society, Chemical Communication**, London, v. 13, p. 1361–1362, 1995.

HENRIKSEN, P.; CARMICHAEL, W.W.; NA, J.; MOESTRUP, O. Detection of an anatoxin-a(s)-like anticholinesterase in natural blooms and cultures of cyanobacteria/bluegreen algae from Danish lakes and in the stomach contents of poisoned birds. **Toxicon**, Oxford, v. 35, n. 6, p. 901-913, 1997.

HILTON, J.A. ; FOSTER, R.A. ; TRIPP, H.J. ; CARTER, B.J. ; ZEHR, J.P. ; TRACY A. VILLAREAL, T.A. Genomic deletions disrupt nitrogen metabolism pathways of a cyanobacterial diatom symbiont. **Nature Communications**, London, v. 4, p. 1-7, 2012.

HILTON,J.A. **The genome sequence of a Rhizosolenia-associated heterocyst-forming cyanobacterium**. 2014. 153 f. Thesis (PhD) - University of California, Santa Cruz, 2014.

HONDA, R.Y. **Caracterização morfológica e molecular de cianobactérias do gênero *Anabaena* isoladas de corpos-d'água brasileiros**. 2009. 154 f. Tese (Doutorado em Agronomia) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2009.

ISOJÄRVI, J.; SHUNMUGAM, S.; SIVONEN, K.; ALLAHVERDIYEVA, Y.; ARO, E.M.; BATTCHIKOVA, N. Draft genome sequence of *Calothrix* strain 336/3, a novel H_2-producing cyanobacterium isolated from a finnish lake. **Genome Announcements**, Washington, DC, v. 3, n. 1, e01474-14, 2015.

ITEMAN, I.; RIPPKA, R.; MARSAC, N.T.; HERDMAN, M. rDNA analyses of planktonic heterocystous cyanobacteria, including members of the genera *Anabaenopsis* and *Cyanospira*. **Microbiology**, Reading, v. 148, p. 481–496, 2002.

JONES, A.C.; MONROE, E.A.; EISMAN, E.B.; GERWICK, L.; SHERMAN, D.H.; GERWICK, W.H. The unique mechanistic transformations involved in the biosynthesis of modular natural products from marine cyanobacteria. **Natural Product Reports**, Cambridge, v. 7, p. 1048–1065, 2010.

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. AND TABATA, S. Complete genomic sequence of the filamentous nitrogenfixing cyanobacterium *Anabaena* sp. strain PCC 7120. **DNA Research**, Tokyo, v. 8, n. 5, p. 205–213, 2001.

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

KAO, C.Y. Paralytic shellfish poisoning. In: FALCONER, I.R. (Ed.). **Algal toxins in seafood and drinkingwater**. New York: Academic Press, 1993. p. 75-86.

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.

KELLMANN, R.; MIHALI, T.K.; JEON, Y.J.; PICKFORD, R.; POMATI, F.; NEILAN, B.A. Biosythetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria. **Applied and Environmental Microbiology**, Washington, DC, v. 74, n. 3, p. 4044–4053, 2008.

KELLMANN, R.; MIHALI, T.K.; NEILAN, B.A. Identification of a saxitoxin biosynthesis gene with a history of frequent horizontal gene transfers. **Journal of Molecular Evolution**, New York, v. 67, n. 5, p. 526-538, 2008.

KERSTEN, R.D.; YANG, Y.L.; XU, Y.; CIMERMANCIC, P.; NAM, S.J.; FENICAL, W.; FISCHBACH, M.A.; MOORE, B.S.; DORRESTEIN, P.C. A mass spectrometry-guided genome mining approach for natural product peptidogenomics. **Nature Chemical Biology**, New York, v. 7, p. 794, 2011.

KOBAYASHI, M.; KAKIZONO, T.; NAGAI, S. Astaxanthin production by a green alga, *Haematococcus pluvialis*, accompanied with morphological changes in acetate media. **Journal of Fermentation and Bioengineering,** Osaka, v. 71, n. 5, p. 335-339.

KOKETSU, K.; MITSUHASHI, S.; TABATA, K. Identification of homophenylalanine biosynthetic genes from the cyanobacterium nostoc punctiforme PCC73102 and application to its microbial production by *Escherichia coli*. **Applied and Environmental Microbiology**, Washington, DC, v. 79, p. 2201-2208, 2013.

KOMÁREK J. 1984. Sobre las cyanoficeas de Cuba: (3) especies planctónicas que forman florecimientos de las aguas. **Acta Botanica Cubana**, La Habana, v. 19, p. 1-33.

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.

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.

LARSSON, J; NYLANDER, J.A.A; BERKAM, B. Genome fluctuations in cyanobacteria reflect evolutionary, developmental and adaptive traits. **BCM Evolutionary Biology**, London, v. 11, p. 1-21, 2011.

LEIKOSKI, N.; FEWER, D.P.; JOKELA, J.; WAHLSTEN, M.; ROUHIAINEN, L.; SIVONEN, K. Highly diverse cyanobactins in strains of the genus Anabaena. **Applied and Environmental Microbiology**, Washington, DC, v. 76, p. 701-709, 2010.

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

LIMA, A.R.J.; SIQUEIRA, A.S.; DOS SANTOS, B.G.S.; DA SILVA, F.D.F.; LIMA, C.P.; CARDOSO, J.F.; VIANEZ JUNIOR, J.L.D.S.G.; DALL'AGNOL, L.T.; MCCULLOCH, J.A.; NUNES, M.R.T.; GONCALVES, E.C. Draft genome sequence of the Brazilian *Cyanobium* sp. strain CACIAM 14. **Genome Announcements**, Washington, DC, v. 2, n. 4, e00669-14, 2014.

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.

LIU, L.; BUDNJO, A.; JOKELA, J.; HAUG, B.E.; FEWER, D.P.; MATTI WAHLSTEN, M.; ROUHIAINEN, L.; PERMI, P.; FOSSEN, T.; SIVONEN, K. Pseudoaeruginosins, nonribosomal peptides in *Nodularia spumigena*. **ACS Chemical Biology**, Washington, DC, v. 10, p. 725-733.

LIU, L.; BUDNJO, A.; JOKELA, J.; HAUG, B.E.; FEWER, D.P; WAHLSTEN, M.; ROUHIAINEN, L.; PERMI, P.; FOSSEN, T.; SIVONEN, K. Pseudoaeruginosins, nonribosomal peptides in Nodularia spumigena. **Chemical Biology**, New York, v. 10, p. 725-733, 2015.

MACHADO, K.C.; MACHADO, K.; GOMES JÚNIOR, A.L.; FREITAS, R.M. Aplicação farmacêutica de inibidores de proteases: uma prospecção tecnológica. **Innovation, Technology and Management Journal,** Santiago, v. 4., n.2, p. 780-787, 2014.

MAHMOOD, N.A.; CARMICHAEL, W.W. The pharmacology of anatoxin-a(s), a neurotoxin produced by the freshwater cyanobacterium Anabaena flos-aquae NRC 525-17. **Toxicon**, Oxford, v. 24, n. 5, p. 425–434, 1986.

MAHMOOD, N.A.; CARMICHAEL, W.W.; Anatoxin-a(s), an anticholinesterase from the cyanobacterium Anabaena flos-aquae NRC-525–17. **Toxicon**, Oxford, v. 25, n. 11, p. 1221-1227, 1987.

MAHMOOD, N.A.; CARMICHAEL, W.W.; PFAHLER, D. Anticholinesterase poisoning in dogs from cyanobacterial (blue green algae) bloom dominated by *Anabaena flos-aquae*. **American Journal of Veterinary Research**, Chicago, v. 49, n. 4, p. 500–503, 1998.

MARCHLER-BAUER, A.; DERBYSHIRE, M.K.; GONZALES, N.R.; LU, S.; CHITSAZ, F.; GEER, L.Y.; GEER, R.C.; HE, R.C.; GWADZ, M.; HURWITZ, D.I.; LANCZYCKI, C.J.; LU, F.; MARCHLER, G.H.; SONG, J.S.; THANKI, N.; WANG, Z.; YAMASHITA, R.A.; ZHANG, D.; ZHENG, C.; BRYANT, S.H. CDD:NCBI's conserved domain database. **Nucleic Acids Research**, London, 2015, v. 43, p. D222-D226, 2015.

MATSUNAGA, S.; MOORE, R.E.; NIEMCZURA, W.P.; CARMICHAEL, W.W. Anatoxina(s), a potent anticholinesterase from Anabaena flos-aquae. **Journal of the American Chemical Society**, Washington, DC, v. 111, n. 20, p. 8021–8023, 1989.

MAZUR-MARZEC, H.; KACZKOWSKA, M.J.; BLASZCZYK, A.; AKCAALAN, R.; SPOOF, L.; MERILUOTO, J. Diversity of peptides produced by Nodularia spumigena from various geographical regions. **Marine Drugs**, Basel, v. 11, p. 1-19, 2013.

MCGLEENON, B.M.; DYNAN, K.B.; PASSMORE, A.P**.** Acetylcholinesterase inhibitors in Alzheimer's disease. **Clinical Pharmacology**, Stamford, v. 48, p. 471-480, 1999.

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.

MÉJEAN, A.; MANN, S.; MALDINEY, T.; VASSILIADIS, G.; LEQUIN, O.; PLOUX, O. Evidence that biosynthesis of the neurotoxic alkaloids anatoxin-a and homoanatoxin-a in the cyanobacterium Oscillatoria PCC 6506 occurs on a modular polyketide synthase initiated by L-proline. **Journal of the American Chemical Society**, Washington, DC, v. 131, n. 22, p. 7512-7513, 2009.

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.; KELLMANN, R.; MUENCHHOFF, J.; BARROW, K.D.; NEILAN, B.A. Characterization of the gene cluster responsible for cylindrospermopsin biosynthesis. **Applied and Environmental Microbiology**, Washington, DC, v. 74, n. 3, p. 716–722, 2008.

MOFFITT, M.C.; NEILAN, B.A. Characterization of the nodularin synthetase gene cluster and proposed theory of the evolution of cyanobacterial hepatotoxins. **Applied and Environmental Microbiology**, Washington, DC, v. 70, n. 11, p. 6353-6362, 2004.

MOLICA, R.J.R.; OLIVEIRA, E.J.A.; CARVALHO, P.V.V.C.; COSTA, A.N.S.F.; CUNHA, M.C.C.; MELO, G.L.; AZEVEDO, S.M.F.O. Occurrence of saxitoxins and an anatoxin-a(s)-like anticholinesterase in a Brazilian drinking water supply. **Harmful Algae**, Amsterdam, v. 4, n. 4, p. 743–753, 2005.

MONSERRAT, J.M.; YUNES, J.S.; BIANCHINI, A. Effects of Anabaena spiroides (cyanobacteria) aqueous extracts on the acetylcholinesterase activity of aquatic species. **Environmental Toxicology and Chemistry**, New York, v. 20, p. 1228-1235, 2001.

MOORE, B.S. Biosynthesis of marine natural products: microorganisms and macroalgae. **Natural Product Reports**, Cambridge, v. 16, p. 653-674, 1999.

MOORE, B.S.; OHTANI, I.; DE KONING, C.B.; MOORE, R.E.; CARMICHAEL, W.W. Biosynthesis of anatoxin-a(s). Origin of the carbons. **Tetrahedron Letters**, Oxford, v. 33, p. 6595-6598, 1992.

MOORE, R.E.; OHTANI, I.; MOORE, B.S. DE KONING. C.B.; YOSHIDA, W.Y.; RUNNEGAR, M.T.C.; CARMICHAEL, W.W. Cyanobacterial toxins. **Cazzetta Chimica Italiana**, Rome, v. 123, p. 329-336, 1993.

MÜLLER, R.; WINK, J. Future potential for anti-infectives from bacteria: how to exploit biodiversity and genomic potential. **International Journal of Medical Microbiology**, Jena, v. 304, p. 3−13, 2014.

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

NEILAN, B.A.; PEARSON, L.A.; MUENCHHOFF, J.; MOFFITT, M.C.; DITTMANN, E. Environmental conditions that influence toxin biosynthesis in cyanobacteria. **Environmental Microbiology**, Oxford, v. 15, n. 5, p. 1239-1253, 2013.

NETT, M. Genome mining: concept and strategies for natural product discovery. **Progress in the Chemistry of Organic Natural Products**, Wien, v. 99, p. 199−245, 2014.

OMAN, T.J.; VAN DER DONK, W.A. Follow the leader: the use of leader peptides to guide natural product biosynthesis. **Nature Chemical Biology**, New York, v. 6, p. 9–18, 2010.

ONODERA, H.; OSHIMA, Y.; HENRIKSEN, P.; YASUMOTO, T. Confirmation of anatoxin-a(s), in the cyanobacterium *Anabaena lemmermannii*, as the cause of bird kills in Danish lakes. **Toxicon**, Oxford, v. 35, n. 11, p. 1645–1648, 1997.

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.

PAN, X.; CHANG, F.; KANG, L.; LI, G.; LI, D.; LIU, Y.; SHEN, Y.; WEI, Z. Morphological characteristics and phylogenetic relationship of *Anabaena* species from Lakes Dianchi and Erhai, China. **Hydrobiologia**, Dordrecht, v. 614, n. 1, p. 353-362, 2008.

PATTERSON, G.M.L.; LARSEN, L.K.; MOORE, R.E. Bioactive natural products from bluegreen algae. **Journal Applied Phycology**, Amsterdam, v. 6, p. 151–157, 1994.

PETRIKOVICS, I.; PAPAHADJOPOULOS, D.; HONG, K.; CHENG, T.C.; BASKIN, S.I.; JIANG, J.; JASZBERENYI, J.C.; LOGUE, B.A.; SZILASI, M.; MCGUINN, W.D.; WAY, J.L. Comparing therapeutic and prophylactic protection against the lethal effect of paraoxon. **Toxicological Science**, San Diego, v. 77, p. 258-262, 2004.

PITA, R.; ANADON, A.; MARTINEZ-LARRANAGA, M.R. Neurotoxins with anticholinesterase activity and their possible use as warfare agents. **Medicina Clinica**, Barcelona, v. 121, p. 511-517, 2003.

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, n. 1, p. 11 26, 2005a.

RAN, L.; LARSSON, J.; VIGIL-STENMAN, T.; NYLANDER, J.A.; ININBERGS, K.; ZHENG, W.W.; LAPIDUS, A.; LOWRY, S.; HASELKORN, R.; BERGMAN, B. 71 Genome erosion in a nitrogen-fixing vertically transmitted endosymbiotic multicellular cyanobacterium. **PLoS One**, San Francisco, v. 8, n. 5, e11486, 2010.

RESHEF, V.; CARMELI, S. Schizopeptin 791, a new anabeanopeptin-like cyclic peptide from the *cyanobacterium Schizothrix* sp. **Journal of Natural Products,** Cincinnati**,** v. 65, n. 8, p. 1187-1189, 2002.

RIGONATO, J.; ALVARENGA, D.O.; BRANCO, L.H.Z.; VARANI, A.M.; BRANDINI, F.P.; FIORE, M.F. Draft genome of a novel culturable marine *chroococcalean* cyanobacterium from the South Atlantic ocean. **Genome Announcements**, Washington, DC, v. 3, n. 2, e00384-15, 2015.

RODRÍGUEZ-GARCÍA, A.; FUENTE, A.; PÉREZ-REDONDO, R.; MARTÍN, J.F.; LIRAS, P. Characterization and expression of the arginine biosynthesis gene cluster of *Streptomyces clavuligerus*. **Journal of Molecular Microbiology and Biotechnology**, Basel, v. 2, n. 4, p. 543-550, 2000.

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

ROUHIAINEN, L.; PAULIN, L.; SUOMALAINEN, S.; HYYTIÄINEN, H.; BUIKEMA, W.J.; HASELKORN, R.; SIVONEN, K. Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in Anabaena strain 90. **Molecular Microbiology**, Oxford, v. 37, p. 156–167, 2000.

ROUHIAINEN, L.; JOKELA, J.; FEWER, D.P.; URMANN, M.; SIVONEN, K. Two alternative starter modules for the non-ribosomal biosynthesis of specific anabaenopeptin variants in *Anabaena* (cyanobacteria). **Chemistry & Biology**, New York, v. 17, p. 265-273, 2010.

ROUNGE, T.B.; ROHRLACK, T.; NEDERBRAGT, A.J.; KRISTENSEN, T.; AKOBSEN, K.S. A genome-wide analysis of nonribosomal peptide synthetase gene clusters and their peptides in a *Planktothrix rubescens* strain. **BMC Genomics**, London, v. 10, p. 1-11, 2009.

RUDI, K.; SKULBERG, O.M.; JAKOBSEN, K.S. Evolution of cyanobacteria by exchange of genetic material among phyletically related strains. **Journal of Bacteriology**, Washington, DC, v. 180, n. 13, p. 3453-3461, 1998.

RUDI, K.; SKULBERG, O.M.; LARSEN, F.; JAKOBSEN, K.S. Strain characterization and classification of oxyphotobacteria in clone cultures on the basis of 16S rRNA sequences from the variable regions V6, V7, and V8. **Applied and Environmental Microbiology**, Washington, DC, v. 63, n. 7, p. 2593-2599, 1997.

RUTHERFORD, K.; PARKHILL, J.; CROOK, J.; HORSNELL, T.; RICE, P.; RAJANDREAM, M.A.; BARRELL, B. Artemis: sequence visualization and annotation. **Bioinformatics**, Oxford, v. 16, n. 10, p. 944-945, 2000.

SANZ, M.; DÖRR, F.A.; PINTO, E. First report of spumigin production by the toxic Sphaerospermopsis torques-reginae cyanobacterium. **Toxicon,** Oxford, v. 108, p. 15-18, 2015.

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.

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

SEDMAK, B.; CARMELI, S.; ELERŠEK, T. "Non-toxic" cyclic peptides induce lysis of cyanobacteria: an effective cell population density control mechanism in cyanobacterial blooms. **Microbiology Ecology**, Amsterdam, v. 56, p, 201-209, 2008.

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

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

SHIMIZU, Y. Microalgal metabolites: a new perspective. **Annual Review of Microbiology**, Palo Alto, v. 50, p. 43–465, 1996.

SIEBER, S.A.; MARAHIEL, M.A. Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. **Chemical Reviews**, Easton, v. 105, p. 715-738, 2005.

SIGITAS ŠULIUS, S.; ALZBUTAS, G.; KVEDERAVIIUTE˙,K.; KOREIVIENE˙, J.; ZAKRYS, L.; LUBYS, A.; PAŠKAUSKASA, R. Draft genome sequence of the cyanobacterium *Aphanizomenon flosaquae* strain 2012/KM1/D3, isolated from the Curonian lagoon (Baltic Sea). **Genome Announcements**, Washington, DC, v. 3, p. 1-2, 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.

SINGH, D.; CHANDRABABUNAIDU, M.M.; PANDA, A.; SEN, D.; BHATTACHARYYA, S.; ADHIKARY, S.P.; TRIPATHY, S. Draft genome sequence of cyanobacterium *Hassallia byssoidea* strain VB512170, isolated from monuments in India. **Genome Announcements**, Washington, DC, v. 3, p. 1-2, 2015.

SIVONEN, K.; BÖRNER, T. Bioactive compounds produced by cyanobacteria. In: Herrero, A.; Flores, E. (Ed.). **The cyanobacteria: molecular biology, genomics and evolution**. Linton, UK: Caister Academic Press, 2008. p. 159-197.

SPOOF, L.; BŁASZCZYK, A.; MERILUOTO, J.; CEGŁOWSKA, M.; MAZUR-MARZEC, H. Structures and activity of new anabaenopeptins produced by baltic sea cyanobacteria. **Marine Drugs**, Basel, 2016. doi:10.3390/md14010008.

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

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

THOMAS, M.G.; CHAN, Y.A.; OZANICK, S.G. Deciphering tuberactinomycin biosynthesis: isolation, sequencing, and annotation of the viomycin biosynthetic gene cluster. **Antimicrobial Agents and Chemotherapy**, Washington, DC, v. 47, n. 9, p. 2823–2830, 2003.

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.

TILLETT, D.; DITTMANN, E.; ERHARD, M.; VON DÖHREN, H.; BÖRNER, T.; NEILAN, B.A. Structural organization of microcystin biosynthesis in Microcystis aeruginosa PCC7806: an integrated peptide-polyketide synthetase system. **Chemical Biology**, Cambridge, v. 7, n. 10, p. 753-764, 2000.

TOWNSEND, C.A. New reactions in clavulanic acid biosynthesis. **Current Opinion in Chemical Biology**, London, v. 6, p. 583-589, 2002.

TUOVINEN, K.; KALISTE-KORHONEN, E.; RAUSHEL, F.M.; HÄNNINEN, O. Success of pyridostigmine, physostigmine, eptastigmine and phosphotriesterase treatments in acute sarin intoxication. **Toxicology**, Amsterdam, v. 134, p. 169-178, 1999.

URRUTIA-CORDERO, P.; AGHA, R.; CIRÉS, S.; ÁNGELES LEZCANO, M.; SÁNCHEZ-CONTRERAS, M.; WAARA, K-O.; UTKILEN, H.; QUESADA, A. Effects of harmful cyanobacteria on the freshwater pathogenic free-living amoeba *Acanthamoeba castellanii*. **Aquatic Toxicology**, Amsterdam, v. 130-131, p. 9-17, 2013

VALÉRIO, E.; CHAVES, S.; TENREIRO, R. Diversity and impact of prokaryotic toxins on aquatic environments: a review. **Toxins**, Basel, v. 2, p. 2359-2410, 2010.

VAN APELDOORN, M.E.; VAN EGMOND, H.P.; SPEIJERS, G.J.A.; BAKKER, G.J.I. Toxins of cyanobacteria. **Molecular Nutrition and Food Research**, Weinheim, v. 51, p. 7-60, 2007.

VON DÖHREN, H.; KELLER, U.; VATER, J.; ZOCHER, R. Multifunctional peptide synthetases. **Chemical Reviews**, Easton, v. 29, n. 7, p. 2675-2706, 1997.

VOß, B.; BOLHUIS, H.; FEWER, D.P. KOPF, M.; MÖKE, F.; HAAS, F.; ELSHEHAWY, R.; HAYES, P.; BERGMAN, B.; SIVONEN, K.; DITTMANN, E.; SCANLAN, D.J.; HAGEMANN, M.; STAL, L.J.; HESS, W.R. Insights into the physiology and ecology of the brackish-water-adapted Cyanobacterium Nodularia spumigena CCY9414 based on a genome-transcriptome analysis. **PLoS ONE**, San Francisco, v. 8, n. 3, e60224, 2013.

WACKLIN, P.; HOFFMANN, L.; KOMÁREK, J. Nomenclatural validation of the genetically revised cyanobacterial genus *Dolichospermum* (Ralfs ex Bornet et Flahault) comb. nova. **Fottea**, Praha, v. 9, n. 1, p. 59–64, 2009.

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.

WANG, H.; SIVONEN, K; FEWER, D.P. Genomic insights into the distribution, genetic diversity and evolution of polyketide synthases and nonribosomal peptide synthetases. **Current Opinion in Genetics & Development**, London, v. 35, p. 79-85, 2015.

WEBER, T.; BLIN, K.; DUDDELA, S.; KRUG, D.; KIM, H.U.; BRUCCOLERI, R.; LEE, S.Y.; FISCHBACH, M.A.; MÜLLER, 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. 237-243. 2015.

WELKER, M.; VON DÖHREN, H. Cyanobacterial peptides: nature's own combinatorial biosynthesis. **FEMS Microbiology Reviews**, Amsterdam, v. 30, n. 4, p. 530-563, 2006.

WERNER, V.R.; LAUGHINGHOUSE IV, D.H. Bloom forming and other planktonic Anabaena (Cyanobacteria) morphospecies with twisted trichomes from Rio Grande do Sul State, Brazil. **Nova Hedwigia**, Weinheim, v. 89, n.1-2, p. 17-47, 2009

WERNER. V.R.; LAUGHINGHOUSE IV, H.D.; SANT'ANNA, C.L.; FIORE, M.F.; HOFF, C.; SANTOS, K.R.S.; NEUHAUS, E.B.; MOLICA, R.J.R.; HONDA, R.Y.; ECHENIQUE, R.O. Morphological and molecular studies of *Sphaerospermopsis torques-reginae* (Cyanobacteria, Nostocales) from South American water blooms. **Phycologia**, Berkeley, v. 51, n. 2, p. 228-238, 2012.

WHITTON, B.A.; POTTS, M. Introduction to the Cyanobacteria. In: Whitton, B.A.; Potts, M. (Eds.). The ecology of cyanobacteria: their diversity in time and space. Dordrecht: Kluwer Academic, 2000. p. 1-11.

WILLAME, R.; BOUTTE, C.; GRUBISIC, S.; WILMOTTE, A.; KOMÁREK, J.; HOFFMANN, L. Morphological and molecular characterization of planktonic cyanobacteria from Belgium and Luxembourg. **Journal of Phycology**, Oxford, v. 42, n. 6, p. 1312-1332, 2006.

WINTER, J.M.; BEHNKEN, S.; HERTWECK, C. Genomics-inspired discovery of natural products. **Current Opinion in Chemical Biology**, London, v. 15, n. 1, p. 22-31, 2011.

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

YERRAPRAGADA, S.; SHUKLA, A.; HALLSWORTH-PEPIN, K.; CHOI, K.; WOLLAM, A.; CLIFTON, S.; QIN, X.; MUZNY, D.; RAGHURAMAN, S.; ASHKI, H.; UZMAN, A.; HIGHLANDER, S.K.; FRYSZCZYN, B.G.; FOX, G.E.; TIRUMALAI, M.R.; LIU, Y.; KIM, S.; KEHOE, D.M.; WEINSTOCK, G.M. Extreme sensory complexity encoded in the 10-megabase draft genome sequence of the chromatically acclimating cyanobacterium *Tolypothrix* sp. PCC 7601. **Genome Announcements**, Washington, DC, v. 3, n. 3, p. 1-2, 2015.

YIN, X.; ZABRISKIES, M. The enduracidin biosynthetic gene cluster from Streptomyces fungicidicus. **Microbiology**, Reading, v. 152, p. 2969–2983, 2006.

YU, D.; XU, F.; ZENG, J.; ZHAN, J. Type III polyketide synthases in natural product biosynthesis. **IUBMB Life**, London, v. 64, n. 4, p. 285-295, 2012.

YUU HIROSE, Y.; FUJISAWAC, T.; OHTSUBOD, Y.; KATAYAMAE, M.; MISAWAA, N.; WAKAZUKIB, S.; SHIMURAF, Y.; NAKAMURAC, Y.; KAWACHIF, M.; YOSHIKAWAG, H.; EKIA, T.; YU KANESAKIH, Y. Complete genome sequence of cyanobacterium *Fischerella* sp. NIES-3754, providing thermoresistant optogenetic tools. **Journal of Biotechnology**, Amsterdam, 2016. doi:10.1016/j.jbiotec.2016.01.011.

ZAPOMELOVÁ, E.; JEZBEROVÁ, J.; HROUZEK, P.; HISEM, D.; REHÁKOVÁ, K.; KOMÁRKOVÁ, J. Polyphasic characterization of three strains of *Anabaena reniformis* and *Aphanizomenon aphanizomenoides* (Cyanobacteria) and their reclassification to *Sphaerospermum* gen. nov. (incl. *Anabaena kisseleviana*). **Journal of Phycology**, Oxford, v. 45, n. 6, p. 1363-1373, 2009.

ZAPOMELOVÁ, E.; JEZBEROVÁ, J.; HROUZEK, P.; HISEM, D.; REHÁKOVÁ, K.; KOMÁRKOVÁ, J. Nomenclatural note: polyphasic characterization of three strains of *Anabaena reniformis* and *Aphanizomenon aphanizomenoides* (Cyanobacteria) and their reclassification to *Sphaerospermum* gen. nov. (incl. *Anabaena kisseleviana*). **Journal of Phycology**, Oxford, v. 45, n. 2, p. 415, 2010

ZEHR, J.P.; BENCH, S.R.; CARTER, B.J.; HEWSON, I.; NIAZI, F.; SHI,T.; TRIPP, H.J.; AFFOURTIT, J.P. Globally distributed uncultivated oceanic N_2-fixing cyanobacteria lack oxygenic photosystem II. **Science**, Washington, DC, v. 322, p. 1110-1112, 2008.

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.

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