



# Biosynthesis of microcystin hepatotoxins in the cyanobacterial genus *Fischerella*

Karina Heck<sup>a</sup>, Danilo O. Alvarenga<sup>a, c</sup>, Tania K. Shishido<sup>b</sup>, Alessandro M. Varani<sup>c</sup>, Felipe A. Dörr<sup>d</sup>, Ernani Pinto<sup>d</sup>, Leo Rouhiainen<sup>b</sup>, Jouni Jokela<sup>b</sup>, Kaarina Sivonen<sup>b</sup>, Marli F. Fiore<sup>a, \*</sup>

<sup>a</sup> University of São Paulo (USP), Center for Nuclear Energy in Agriculture, Piracicaba, SP, Brazil

<sup>b</sup> University of Helsinki, Department of Food and Environmental Sciences, Division of Microbiology and Biotechnology, Helsinki, Finland

<sup>c</sup> São Paulo State University (UNESP), College of Agricultural and Veterinary Science, Department of Technology, Jaboticabal, SP, Brazil

<sup>d</sup> University of São Paulo (USP), School of Pharmaceutical Sciences, São Paulo, SP, Brazil

## ARTICLE INFO

### Article history:

Received 12 June 2017

Received in revised form

18 October 2017

Accepted 22 October 2017

Available online 27 November 2017

### Keywords:

Cyanotoxins

Phosphatase inhibitors

Genome mining

Nostocales

## ABSTRACT

Microcystins (MCs) are serine/threonine phosphatase inhibitors synthesized by several members of the phylum Cyanobacteria. Mining the draft genome sequence of the nostocalean MC-producing *Fischerella* sp. strain CENA161 led to the identification of three contigs containing *mcy* genes. Subsequent PCR and Sanger sequencing allowed the assembling of its complete biosynthetic *mcy* gene cluster with 55,016 bases in length. The cluster encoding ten genes (*mcyA-J*) with a central bidirectional promoter was organized in a similar manner as found in other genera of nostocalean cyanobacteria. However, the nucleotide sequence of the *mcy* gene cluster of *Fischerella* sp. CENA161 showed significant differences from all the other MC-producing cyanobacterial genera, sharing only 85.2 to 74.1% identities. Potential MC variants produced by *Fischerella* sp. CENA161 were predicted by the analysis of the adenylation domain binding pockets and further investigated by LC-MS/MS analysis. To our knowledge, this study presents the first complete *mcy* cluster characterization from a strain of the genus *Fischerella*, providing new insight into the distribution and evolution of MCs in the phylum Cyanobacteria.

© 2017 Elsevier Ltd. All rights reserved.

## 1. Introduction

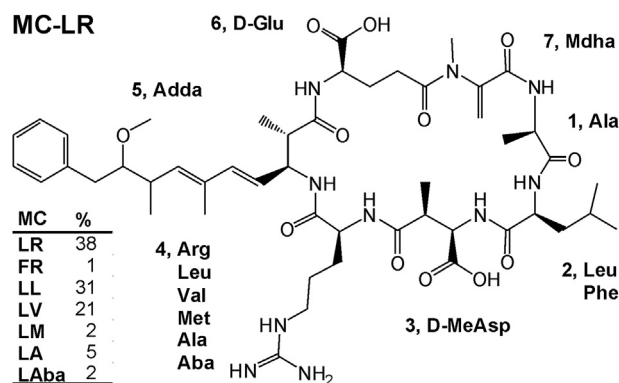
Microcystins (MCs) are small cyclic heptapeptides synthesized by several members of the phylum Cyanobacteria with global significance due to their toxicity to humans and other animals (Jochimsen et al., 1998; Sivonen and Jones, 1999). Their toxicity is exerted through inhibition of members of the protein phosphatase families PP1 and PP2A (MacKintosh et al., 1990; Gullledge et al., 2002). Despite best known for their acute hepatotoxicity, MCs are of interest as possible anti-cancer drug development targets (Niedermeyer et al., 2014; Kounnis et al., 2015). The general structure of MCs can be summarized as cyclo-D-Ala<sup>1</sup>-X<sup>2</sup>-D-MeAsp<sup>3</sup>-Z<sup>4</sup>-Adda<sup>5</sup>-D-Glu<sup>6</sup>-Mdha<sup>7</sup> (see Fig. 1) (Botes et al., 1985), where X and Z are variable L-amino acids, while D-MeAsp corresponds to D-erythro-β-methyl-aspartic acid, Mdha to N-methyl-α-

β-dehydroalanine and Adda to (2S,3S,8S,9S)-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-(4E,6E)-dienoic acid. The latter is exclusive to these toxins and nodularins and contributes to the molecule toxicity (Gullledge et al., 2002; Kounnis et al., 2015).

MCs are synthesized through enzymatic modification of short precursor peptides in the nonribosomal pathway. This process is driven by a multifunctional modular enzyme complex consisted of a combination of nonribosomal peptide synthetases (NRPS), type I polyketide synthases (PKS-I), hybrid NRPS/PKS-I and tailoring enzymes (Nishizawa et al., 2000; Tillett et al., 2000; Christiansen et al., 2003; Rouhiainen et al., 2004; Fewer et al., 2013). The microcystin gene cluster (*mcy*) is composed of nine to ten genes depending on taxa and the involvement of several *mcy* genes in MC biosynthesis was established by gene inactivation studies (Dittmann et al., 1997; Pearson et al., 2004; Christiansen et al., 2008; Fewer et al., 2008). The closely related nodularin (*nda*) synthetase gene cluster from *Nodularia* was also elucidated, indicating that it derived from MC synthetase genes through a deletion event and a change in substrate specificity (Moffitt and Neilan, 2004; Rantala et al., 2004).

\* Corresponding author. University of São Paulo, Center for Nuclear Energy in Agriculture, Av. Centenário 303, 13400-970, Piracicaba, SP, Brazil.

E-mail address: [fiore@cena.usp.br](mailto:fiore@cena.usp.br) (M.F. Fiore).



**Fig. 1.** Structure of MC-LR. The *Fischerella* sp. CENA161 MC variations encountered are shown schematically according to their position. Abbreviations: Adda, (2S,3S,8S,9S)-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-(4E,6E)-dienoic acid; D-Glu, glutamic acid; Mdha, N-methyl- $\alpha$ -dehydroalanine; Ala, alanine; Leu, leucine; Phe, phenylalanine; D-MeAsp, D-erythro- $\beta$ -methyl-aspartic acid; Aba, Aminoisobutyric acid, Arg, arginine, Val, valine and Met, methionine.

The biological role of cyanobacterial MC is not currently understood, but several hypotheses have been suggested such as contributing in photosynthesis, environmental adaptation, protection against oxidative stress, nutrient metabolism and storage, quorum sensing, colony formation, defense against zooplanktonic grazers, iron uptake or transfer and allelopathy (Omidi et al., 2017). These authors stated that conflicting results, unstandardized experimental design, strain-specific behavior and differences between conditions in laboratory and nature hinder generalizations on microcystin functions.

Despite several MC-producing strains have been found in the genera *Microcystis*, *Anabaena*, *Nostoc*, *Fischerella*, *Hapalosiphon*, *Oscillatoria*/*Planktothrix*, and *Phormidium* (Bishop et al., 1959; Botes et al., 1984; Krishnamurthy et al., 1986; Eriksson et al., 1988; Meriluoto et al., 1989; Sivonen et al., 1990; Harada et al., 1991; Prinsep et al., 1992; Izaguirre et al., 2007; Fiore et al., 2009), the MC biosynthetic pathway was only characterized in few strains of the genera *Microcystis*, *Anabaena*, *Planktothrix* and *Nostoc* (Tillett et al., 2000; Christiansen et al., 2003; Rouhiainen et al., 2004; Rounge et al., 2009; Fewer et al., 2013). The *mcy* gene clusters of these distantly related cyanobacterial genera have revealed a highly conserved set of multidomain proteins depicting the same basic reaction steps. Differences among these clusters have been observed in gene arrangements, localization and orientation of promoter regions, and in genes coding for tailoring enzymes. Interestingly, a cyanobacterium containing one *mcy* gene cluster can produce more than one MC variant mainly due to the relaxed specificity of adenylation (A) domains of McyB-A<sub>1</sub> and McyC-A (amino acid positions 2 and 4, Fig. 1). Therefore, the description of novel MC gene clusters from different cyanobacterial taxa offers high potential for isolating variants with unique properties.

Although MC production and fragments of biosynthetic genes have already been identified in strains of the genus *Fischerella* (Fiore et al., 2009; Cirés et al., 2014) and even prediction of an incomplete gene cluster has been reported (Shih et al., 2013), the entire gene cluster remains unsolved. Here we used a genomics-based approach to characterize the complete biosynthetic gene cluster in the MC-producing strain *Fischerella* sp. CENA161. Prediction analysis based on the amino acid residues lining the substrate-binding pockets in NRPS A domains were performed and potential structural variants investigated by high performance liquid chromatography coupled to tandem mass spectrometry (LC–MS/MS).

## 2. Method

### 2.1. Cyanobacterial strain

The cyanobacterium *Fischerella* sp. CENA161 was isolated from a water sample collected from a small concrete dam of spring water in the municipality of Piracicaba, São Paulo state, Brazil, as previously described (Fiore et al., 2009). This strain is maintained under culture in CENA/USP, located in Piracicaba, SP, Brazil, in BG–11 (Allen, 1968) liquid medium without inorganic nitrogen (BG–11<sub>0</sub>), at  $25 \pm 1$  °C, with a 14:10 h light/dark photoperiod, and photon flux density of 40  $\mu\text{mol photons/m}^2/\text{s}$ .

### 2.2. DNA extraction, PCR amplification and Sanger sequencing

Cells from the cyanobacterial culture were collected and processed as previously described (Heck et al., 2016). Total genomic DNA was extracted using the AxyPrep™ Bacterial Genomic DNA Miniprep Kit (Axygen Biosciences) according to manufacturer instructions. The integrity of the total genomic DNA extracted was verified using 1% agarose gel electrophoresis. The extracted DNA was purified with AxyPrep™ PCR Clean-up Kit (Axygen) according to manufacturer instructions. Microcystin codifying genes (*mcy*) were amplified by polymerase chain reaction using a combination of primer sets previously described in literature and designed for this work (Table 1). PCR products were ligated to pGEM®-T Easy Vector Systems (Promega) and inserted into chemically competent *Escherichia coli* DH5 $\alpha$  cells. Plasmids that received the PCR products were extracted from cells by alkaline hydrolysis (Birnboim and Doly, 1979). Sequencing reactions were performed using the Big-Dye Terminator Cycle Sequencing Kit (GE Healthcare), with vector primers M13F/M13R in a Techne TC-412 thermocycler (Bibby Scientific Limited) for 25 cycles at 95 °C for 20 s, 52 °C for 15 s, and 60 °C for 1 min. Purified reactions were analyzed in an ABI PRISM 3500 genetic analyzer (Life Technologies). The sequenced reads had their base quality analyzed and consensus sequences were generated with the Phred/Phrap/Consed software package (Ewing and Green, 1998; Ewing et al., 1998; Gordon et al., 1998). Sequences were aligned and compared to other sequences available in NCBI GenBank (<http://www.ncbi.nlm.nih.gov/>) using BLASTn (Altschul et al., 1997).

### 2.3. Whole genome sequencing and assembly

Genomic DNA extracted from the cells was quantified using Qubit dsDNA Broad BR Assay kit and Qubit®2.0 Fluorometer (Thermo Fisher Scientific). Paired-ends libraries were prepared with the Nextera XT Sample Prep Kit (Illumina), which were sequenced in the MiSeq (Illumina) platform using the MiSeq 600 cycle Reagent Kit v3 (Illumina) according to manufacturer instructions. The quality of the raw Illumina sequence reads were initially assessed using FastQC v0.10.1 (Andrews, 2010). Bases with quality indices lower than Phred 20 and sequences shorter than 50 bp were removed using the program SeqClean 1.8.10 (Zhbannikov et al., 2015). Overlapping read pairs were merged with PEAR 0.9.6 (Zhang et al., 2014) and genome assembly was performed using SPAdes 3.1.1 (Bankevich et al., 2012).

The complete nucleotide sequence of the *mcy* gene cluster of *Fischerella* sp. CENA161 was deposited in GenBank under accession number KX891213.

### 2.4. Microcystin gene cluster annotation and phylogenetic analysis

The MC synthetase gene cluster was identified by using BLASTn alignments between the 10 *mcy* gene sequences obtained in Sanger

**Table 1**  
PCR primers and thermal cycling programs used to amplify the microcystin synthetase gene.

Gene	Primers	Program	Reference
<i>mcyA</i>	OMET-F: TTATTCCAAGTTGCTCCCA MSR: TGCAGATAACTCCGACGTTG	94 °C/4 min; 30 × 94 °C/20s; 55 °C/30s; 72 °C/1 min; 72 °C/7 min.	Tillett et al., 2000
<i>mcyB</i>	pB3F: CAAAARCARGCAGAAAMTTTCAGG pB9R: CCAATCVCTATCTAAACAC	95 °C/3 min; 30 × 94 °C/30 s; 52 °C/30s; 72 °C/1 min; 72°C/10 min.	Fewer et al., 2007
<i>mcyC</i>	pC1F: GGMGTTGGATTGTGCASCATAT pC13R: CACTTTCTAACCAGTATTTTGCCA	95 °C/3min; 30 × 94 °C/30s; 52 °C/30s; 72 °C/1 min; 72 °C/10 min.	Fewer et al., 2007
<i>mcyD</i>	<i>mcyDF</i> GATCCGATTGAATTAGAAAG <i>mcyDR</i> GTATTCCCAAGATTGCC	95 °C/3min; 30 × 94 °C/30s; 56 °C/30s; 72 °C/1 min; 72 °C/10 min.	Rantala et al., 2004
<i>mcyE</i>	<i>mcyEF2</i> : GAAATTTGTGTAGAAGGTGC <i>mcyER4</i> : AATTCTAAGCCCAAGACG	95 °C/3min; 30 × 94 °C/30s; 56 °C/30s; 72 °C/1 min; 72 °C/10 min.	Rantala et al., 2004
<i>mcyF</i>	<i>mcyFKF</i> : GAACAAGAAKCMCCYAA <i>mcyFKR</i> : CCAGAAATAWAGGARCGA	95 °C/3min; 30 × 94 °C/30s; 50 °C/30s; 72 °C/1 min; 72°C/10 min.	This work
<i>mcyG</i>	<i>mcyGF</i> : GAAATTTGGTGGGGAAGTGGAG <i>mcyGR</i> : TTTGAGCAACAATGATACTTTGCTG	95 °C/3min; 30 × 94 °C/30s; 56 °C/30s; 72 °C/1 min; 72 °C/10 min.	Fewer et al., 2007
<i>mcyH</i>	<i>mcyHKF</i> : CAAGAATMGAAACCYATTGC <i>mcyHKR</i> : AGGCGTTGTTGTCTCTCTA	95 °C/3min; 30 × 94 °C/30s; 56 °C/30s; 72 °C/1 min; 72 °C/10 min.	This work
<i>mcyI</i>	<i>mcyIdgenF</i> : TGTGCGTTATCCTAMTAA <i>mcyIdgenR</i> : GGCTTCTCDDCCCTGAAGC	95 °C/3 min; 30 × 94 °C/30 s; 56 °C/30s; 72 °C/1 min; 72°C/10 min.	Pearson et al., 2007
<i>mcyJ</i>	<i>mcyJKF</i> : GGYTATTGGMAAGARGAAAC <i>mcyJKR</i> : GCCATAAACCCAMCCYCKATYTC	95 °C/3min; 30 × 94 °C/30s; 56 °C/30s; 72 °C/1 min; 72 °C/10 min.	This work
<i>mcyB</i> (gap)	<i>mcyB161ADF</i> : GGAATTTTAAAGGCAGG <i>mcyB161ADR</i> : CGTTGGTTTTCAGAAGTATC	95 °C/3min; 30 × 94 °C/30s; 52 °C/30s; 72 °C/1 min; 72 °C/10 min.	This work
<i>mcyC</i> (gap)	<i>mcyC161ADF</i> : CCTTAGATCCTAGTTATCC <i>mcyC161ADR</i> : CCCCAATATGTAATTCTCC	95 °C/3min; 30 × 94 °C/30s; 52 °C/30s; 72 °C/1 min; 72 °C/10 min.	This work

sequencing and the assembled genome file. Manual annotation was performed using Artemis 15.1.10 (Rutherford et al., 2000). The identification of motifs and adenylation domains was performed using NRPS Predictor2 (Rausch et al., 2005; Röttig et al., 2011). Amino acid sequences from the *McyB*<sub>2</sub> and *McyC* adenylation domains and for the 10 genes found in complete microcystin gene clusters available in the NCBI GenBank database were independently aligned with MUSCLE 3.8.31 (Edgar, 2004) and evolutionary models were estimated with ProtTest 3.2 (Darriba et al., 2011). Phylogenetic trees were reconstructed from alignments by Bayesian inference with MrBayes 3.2.5 (Ronquist and Huelsenbeck, 2003) using 5,000,000 generations, four chains and two independent runs.

### 2.5. Liquid chromatography coupled to tandem mass spectrometry (LC–MS/MS)

The intracellular content of 60 mg of freeze-dried cells from the culture sample was extracted with MeOH:H<sub>2</sub>O 70/30 (v/v) at ultrasound probe (Sonic Ruptor 400, Omni) during 1 min on ice. The supernatant was collected after centrifugation (10,000 × g for 10 min) and the extract diluted to 10% MeOH with ultrapure water. The sample was applied to a solid phase extraction cartridge (Sep-Pak 500 mg, Waters Corp.), previously conditioned by the sequential passage of 5 ml of MeOH and 5 ml of MeOH:H<sub>2</sub>O 10/90 (v/v). After a washing step with 5 ml of MeOH:H<sub>2</sub>O 10/90 (v/v), elution proceeded with 5 ml of MeOH:H<sub>2</sub>O 90/10 (v/v). The solvent was evaporated under a stream of nitrogen and the concentrate reconstituted in 500 µl of MeOH:H<sub>2</sub>O 50/50 (v/v), filtered (0.45 µm, PVDF, Millipore) and transferred to appropriate vials. Chromatography was performed in a Prominence HPLC (Shimadzu) employing a Fusion-RP column (150 × 2 mm, 4 µm; Phenomenex) with a gradient of (A) 2 mM ammonium formate containing 0.1% formic acid and (B) acetonitrile:water 90/10 (v/v) with the same additives, at a flow rate of 0.2 ml/min. Gradient elution proceeded as follows: 35–60% B in 10 min; 60–100% B in 6 min; 100% B in 2 min; 100 to 35% B in 0.5 min and finally kept in 25% B for 6.5 min. Collision-induced dissociation experiments for MC detection and characterization were performed in an Esquire HCT ion trap mass

spectrometer (Bruker Daltonics) equipped with an electrospray ion source.

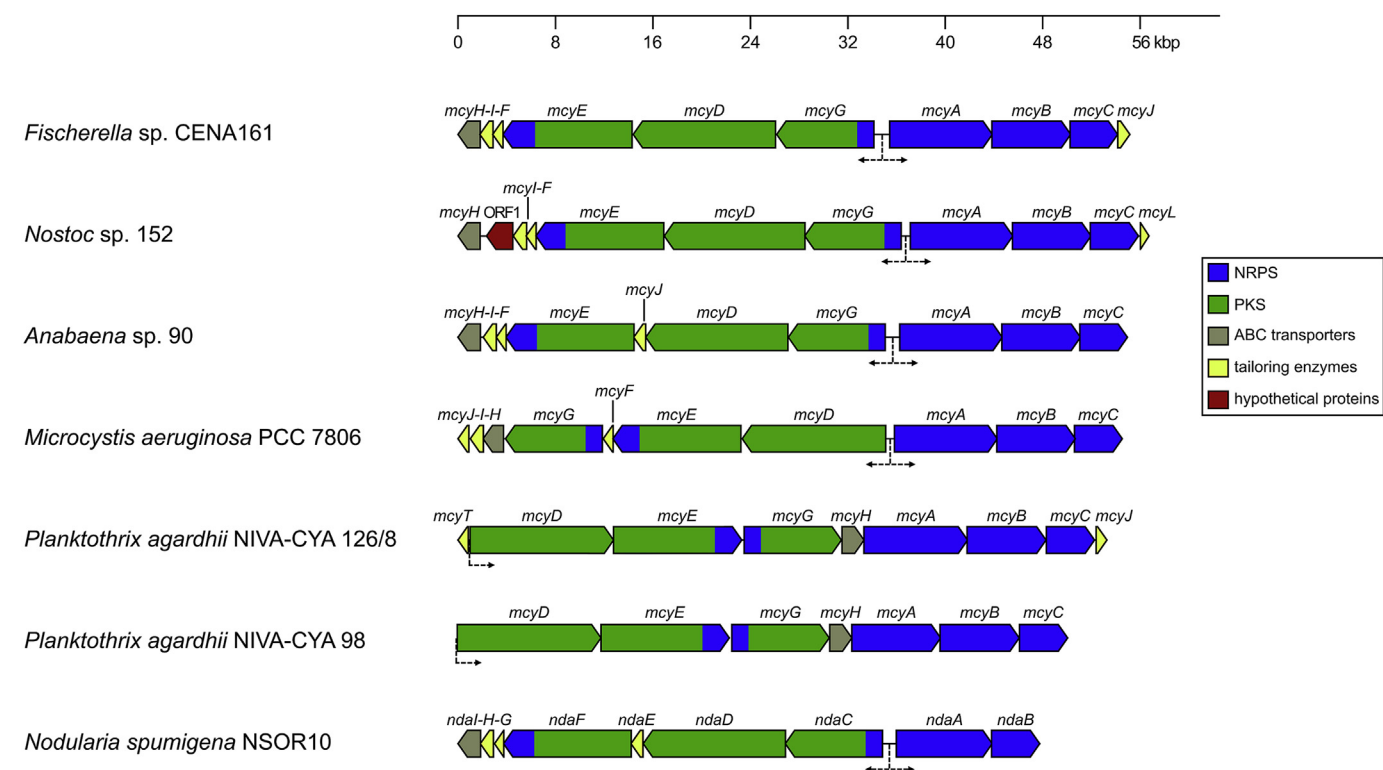
UPLC–QTOF analyses were performed with Acquity I-Class UPLC – Synapt G2-Si HDMS (Waters Corp.) system. Two µl filtered cyanobacterial methanol extract were injected into a Kinetex C8 column (50 × 2.1 mm, 1.7 µm, Phenomenex) which was eluted at 40 °C with a flow rate of 0.3 ml/min using (A) 0.1% formic acid and (B) acetonitrile/isopropanol 50/50 (v/v) containing 0.1% formic acid. Gradient elution proceeded as follows: 25–65% B in 5 min; 65–100% B in 0.01 min; 100% B in 1.99 min; 100 to 20% B in 0.5 min and finally kept in 25% B for 2.5 min. The mass spectrometer was calibrated with sodium formate giving a calibrated mass range from *m/z* 91.055 to 1921.759. Leucine enkephalin was used at 10 s interval as a lock mass reference compound. Mass spectral data was accumulated in positive electrospray ionization at a scan range from *m/z* 50 to 2000.

### 3. Results and discussion

The search for conserved regions from *mcy* genes in the *Fischerella* sp. CENA161 genomic DNA by PCR amplification and Sanger sequencing returned positive for all 10 genes. High-throughput sequencing with Illumina MiSeq resulted in approximately 24 million raw reads. The raw data was assembled into 443 contigs, which constituted a draft genome size of 7,210,502 bp with ca. 220 × coverage and GC content at 40.18%. Both genome size and GC content were in agreement with other recently sequenced cyanobacterial members of the genus *Fischerella* (Dagan et al., 2012; Shih et al., 2013; Hirose et al., 2016).

With the draft genome data in hand, we searched for contig(s) that coharbor *mcy* genes using the *mcy* PCR fragments as *in silico* probes. This effort led to the identification of three contigs (45,504 bp, 5557 bp and 2626 bp). Bioinformatics analyses identified that the two gaps separating the three contigs were located within the *mcyB* and *mcyC* genes, which were closed using PCR amplification and Sanger sequencing. The three contigs were assembled into a 55,137 kb contiguous region that contained the 55,016 bp *mcy* gene cluster.

The CENA161 *mcy* gene cluster encompasses ten genes (*mcyA*–*J*) with a central bidirectional promoter (Fig. 2, Table 2). The genes



**Fig. 2.** Arrangement of gene clusters coding for the biosynthesis of microcystin in *Fischerella* sp. CENA161, *Nostoc* (Fewer et al., 2013), *Anabaena* (Rouhiainen et al., 2004), *Microcystis* (Nishizawa et al., 2000; Tillett et al., 2000), *Planktothrix* (Christiansen et al., 2003; Rounge et al., 2009), and of nodularin in *Nodularia* (Moffitt and Neilan, 2004). Arrows indicate the transcriptional start sites from the putative promoter regions.

**Table 2**  
Functions of proteins encoded in the microcystin biosynthetic gene cluster.

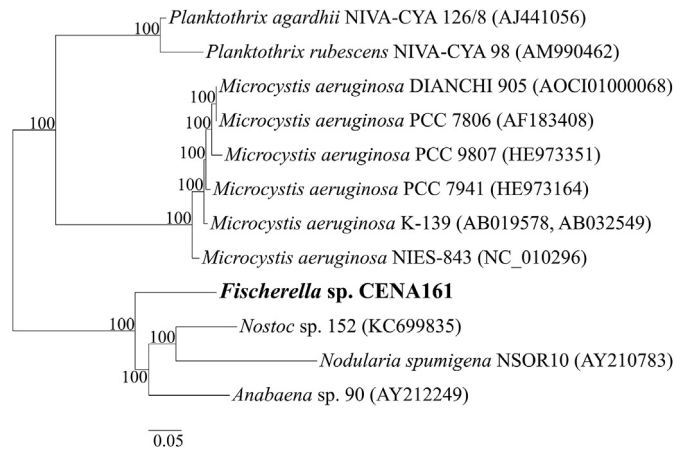
Protein	Lengths (amino acids)	Functions	Top BLAST Hit		
			Organism	Similarity (%)	Accession Number
McyH	590	ABC transporter	<i>Anabaena</i> sp. 90	81.8	AAO62579
McyI	342	putative dehydrogenase	<i>N. spumigena</i> NSOR10	86.1	AAO62580
McyF	262	amino acid racemase	<i>Nostoc</i> sp. 152	81.6	AGZ05271
McyE	3511	NRPS-PKS (KS-AT-PCP-AMT-C-A-PCP-C)	<i>Nostoc</i> sp. 152	80.1	AGZ05272
McyD	3872	PKS (KS-DH-CM-KR-PCP-KS-AT-DH-KR-PCP)	<i>Anabaena</i> sp. 90	79.2	AAO62584
McyG	2639	NRPS-PKS (A-PCP-KS-AT-CM-KR-PCP)	<i>Anabaena</i> sp. 90	79.9	AAO62585
McyA	2787	NRPS (A-NMT-PCP-C-A-PCP-E)	<i>Anabaena</i> sp. 90	81.5	AAO62586
McyB	2136	NRPS (C-A-PCP-C-A-PCP)	<i>Anabaena</i> sp. 90	83.1	AAO62587
McyC	1283	NRPS (C-A-PCP-Te)	<i>Anabaena</i> sp. 90	83.3	AAO62588
McyJ	313	O-acetyltransferase	<i>N. spumigena</i> NSOR10	84.8	AAO64406

encoding the McyG, D, E, A, B and C enzymes are responsible for the stepwise assembly and cyclization of peptide intermediates to form MC, while McyF, I and J are tailoring enzymes and McyH is an ABC transporter hypothetically involved in the efflux of the toxin (Tillett et al., 2000; Christiansen et al., 2003; Rouhiainen et al., 2004; Fewer et al., 2013). MC biosynthesis by *Fischerella* sp. CENA161 follows the collinearity rule, i.e., the order of genes is the same as the order of the single enzymatic steps (Marahiel et al., 1997; von Döhren et al., 1997) as occurs in other nostocalean cyanobacteria, such as *Anabaena* sp. 90 and *Nostoc* sp. 152, and NOD in *Nodularia spumigena* NSOR10. In these cyanobacteria, MC assembly is believed to initiate with the A-PCP domains of the hybrid NRPS/PKS-I enzyme McyG loading phenyllactate (Hicks et al., 2006). Then, the four PKS-I modules of McyG, D and E complete the formation of the Adda skeleton, and the O-methyltransferase McyJ (Christiansen et al., 2003) and the aminotransferase domain of McyE (Tillett et al., 2000) incorporate Adda side chain modifications. The NRPS

modules of McyG, A, B, and C incorporate the six remaining amino acids (Tillett et al., 2000). The enzymes 2-hydroxy acid dehydrogenase McyI (Pearson et al., 2007) and the aspartate racemase McyF (Sielaff et al., 2003) are involved in the biosynthesis of D-erythro-β-methyl-aspartate. Finally, the elongated peptide is released from the enzyme complex by the thioesterase domain of McyC.

The nucleotide sequence of the *Fischerella* sp. CENA161 *mcy* gene cluster showed significant differences from clusters thus far characterized. Hits from the BLAST analyses in descending order of identity were *Anabaena* sp. 90 (coverage 95%, identity 85.2%), *Nostoc* sp. 152 (coverage 89%, identity 83.0%), *Nodularia spumigena* NSOR10 (coverage 78%, identity 80.8%), *Planktothrix rubescens* NIVA-CYA 98 (coverage 71%, identity 78.2%) and *Microcystis aeruginosa* NIES-843 (coverage 71%, identity 74.1%). Similarly, for amino acid sequences, top BLAST hits observed were with other nostocalean strains (Table 2). Comparisons of the *mcy* gene cluster from *Fischerella* sp. CENA161 with other nostocalean genera (*Anabaena*,





**Fig. 3.** Bayesian inference phylogenetic tree reconstructed from concatenated Mcy and Nda amino acid sequences from strains presenting the microcystin gene cluster or nodularin gene cluster. Posterior probabilities are shown in the nodes.

*Nostoc* and *Nodularia*) revealed, in general, the same structural organization with differences in the arrangement of certain genes coding for tailoring enzymes (Fig. 2). On the other hand, major differences were observed in comparisons with distinct cyanobacterial orders (*Planktothrix* and *Microcystis*). These results reinforce previous observations that *mcy* gene arrangements are almost identical among cyanobacteria according to their taxonomic position (Rantala et al., 2004; Jungblut and Neilan, 2006; Kurmayer and Gumpenberger, 2006). In fact, the phylogenetic tree reconstruction based on concatenated amino acid sequences of all complete *mcy* gene clusters so far known (Fig. 3) supported a correlation between

cyanobacterial taxonomy and MC acquisition in these strains. Taking into account the still low number of available sequences, the Bayesian inference sustained the hypothesis that *mcy* genes might have evolved from a common ancestor and their irregular distribution in phylogenetically related taxa is due to repeated loss processes rather than horizontal transfer (Kurmayer et al., 2004; Rantala et al., 2004).

Multiple alignments of the amino acid residues surrounding the substrate-binding pockets in the NRPS adenylation domains observed in the MC gene cluster of CENA161 with other cyanobacteria are shown in Table 3. Adenylation domains in NRPS modules contain ten highly conserved core motifs, named A1 to A10. Lining the binding pocket, ten conserved amino acid residues are believed to recognize a specific substrate, allowing the prediction of the amino acid likely to be selected for activation (Marahiel et al., 1997; Stachelhaus et al., 1999; Challis et al., 2000; Mikalsen et al., 2003). In the case of *Fischerella* sp. CENA161, the amino acid sequences of the McyG and McyE A domains are highly conserved, containing identical residues as observed in all the MC gene clusters so far known (Table 3). These two modules are responsible for the partial formation of Adda in position 5 and the incorporation of glutamic acid in position 6 of the MC molecule, respectively. In a similar way, the McyA-A<sub>1</sub> and McyA-A<sub>2</sub> modules, responsible for the incorporation of amino acids in positions 7 (serine/MdhA) and 1 (alanine), respectively, are highly conserved. Likewise, McyB-A<sub>2</sub> binding pocket, responsible for the incorporation of amino acids in position 3 (MeAsp), was shown to be conserved. On the other hand, McyB-A<sub>1</sub> and McyC-A sequences showed high diversity in the strain analyzed. These two modules are responsible for incorporating amino acids into positions X<sup>2</sup> and Z<sup>4</sup>, respectively. Indeed, it is well known that positions X<sup>2</sup> and Z<sup>4</sup> show the highest degree of structural variation

**Table 3**

Conservation of the seven adenylation domain binding pockets of Mcy of *Fischerella* sp. CENA161 and other cyanobacteria.

Strain/Binding pocket	McyA <sub>1</sub>	McyA <sub>2</sub>	McyB <sub>1</sub>	McyB <sub>2</sub>	McyC	McyE	McyG
<i>Fischerella</i> sp. CENA161	DVWHISLIDK	DLFNNALTYK	DVLIIFGLIYK	DARHVGIFVK	DVWFFGLVVDK	DPRHSGVVVGK	**LWVAASG*
	Ser 100%	Ala 100%	Leu 70%	Tyr 60%	Ser 80%	Glu 100%	Tcl 50%
<i>Anabaena</i> sp. 90	-----	-----	--WF--VD-	-----	---C-----	DPRHSGVVVGK	**LWVAASGK
	Ser 100%	Ala 100%	Leu 80%	Tyr 60%	Ser 80%	Glu 100%	Tcl 60%
<i>Nostoc</i> sp. 152	-----	-----	-A-F-----	-----	---N-FI---	DPRHSGVVVGK	**LWVAASGK
	Ser 100%	Ala 100%	Leu 80%	Tyr 60%	Gln 70%	Glu 100%	Tcl 60%
<i>Planktothrix agardhii</i> NIVA-CYA 126/8	-----	-----S--	-A-F--VD-	-----	-P-G-----	DPRHSGVVVGK	AILWVAASG*
	Ser 100%	Ala 90%	Leu 70%	Glu 60%	Gln 70%	Glu 100%	Tcl 60%
<i>Planktothrix rubescens</i> NIVA-CYA 98	-F-N-GMVH-	-----	-A-F--VD-	-P-----I-	-P-G-----	DPRHSGVVVGK	AILWVAASG*
	Thr 100%	Ala 100%	Leu 100%	Glu 60%	Gln 70%	Glu 100%	Tcl 60%
<i>Microcystis aeruginosa</i> SPC777	---F-----	-----	-GWTI-AVE-	-----	---TI-A---	#	#
	Ser 100%	Ala 100%	Arg 90%	Tyr 60%	Arg 100%		
<i>Microcystis aeruginosa</i> NIES-843	---F-----	-----	-GWTI-AVE-	-----	---TI-A---	DPRHSGVVVGK	**LWVAASG*
	Ser 100%	Ala 100%	Arg 90%	Tyr 60%	Arg 100%	Glu 100%	Tcl 50%
<i>Microcystis aeruginosa</i> PCC 7806	---F-----	-----	-AWFL-NVV-	-----	---TI-A---	DPRHSGVVVGK	**LWVAASG*
	Ser 100%	Ala 100%	Arg 90%	Tyr 60%	Arg 100%	Glu 100%	Tcl 50%
<i>Microcystis aeruginosa</i> DIANCHI 905	---F-----	-----	-AWFL-NVV-	-----	---TI-A---	DPRHSGVVVGK	**LWVAASG*
	Ser 100%	Ala 100%	Leu 100%	Tyr 60%	Arg 100%	Glu 100%	Tcl 50%
<i>Microcystis aeruginosa</i> K-139	---F-----	-----	-AWFL-NVV-	-----	---TI-A---	DPRHSGVVVGK	**LWVAASG*
	Ser 100%	Ala 100%	Leu 100%	Tyr 60%	Arg 100%	Glu 100%	Tcl 50%
<i>Microcystis aeruginosa</i> PCC7941	---F-----	-----	-AWFL-NVV-	-----	---TI-A---	DPRHSGVVVGK	**LWVAASG*
	Ser 100%	Ala 100%	Leu 100%	Tyr 60%	Arg 100%	Glu 100%	Tcl 50%
<i>Microcystis aeruginosa</i> PCC9807	---F-----	-----	-AWFL-NVV-	-----	---TI-A---	DPRHSGVVVGK	**LWVAASG*
	Ser 100%	Ala 100%	Leu 100%	Tyr 60%	Arg 90%	Glu 100%	Tcl 50%
<i>Nodularia spumigena</i> NSOR10 <sup>1</sup>	-F-N-GMVH-	-	-	-----	---N--F---	DPRHSGVVVGK	--LWVAASGK
	Thr 100%			Tyr 60%	Glu 70%	Glu 100%	Tcl 60%

The adenylation domain is responsible for recognition and activation of amino acids in the microcystins. The probability of incorporation for each amino acid is shown.

# There is no available information about that sequence; \* Unknown amino acid; - There is no homologue gene; <sup>1</sup> Nodularin-producing strain.

when compared to other positions in the molecule (Fewer et al., 2007). Such fluctuations in McyB-A<sub>1</sub> and McyC-A sequences are the major contributors for the diversity in MC biosynthesis in different species. Currently, over a hundred MC structural variants are known, differing in the type of amino acids incorporated or modifications to the peptide backbone (Dittmann et al., 2015).

The LC-MS/MS analysis of the strain CENA161 cell extract allowed the identification of seven MC variants: MC-LR (*m/z* 995, [M+H]<sup>+</sup>), MC-FR (*m/z* 1029), MC-LA (*m/z* 910), MC-LAba (*m/z* 923), MC-LM (*m/z* 970), MC-LV (*m/z* 938) and MC-LL (*m/z* 952) (Supplementary Information Figs. S1 and S2). The identification of MC *m/z* 938 was challenging as the product ion spectrum of *m/z* 938 can be well fitted to three isobaric variants. After thorough comparison of the product ion assignments and intensities it seems that variant MC-LV best explained the high resolution spectrum data (Supplementary Table S1). Exact ion masses, accuracies and intensities of MC isoforms are presented in Supplementary Table S2. MC-LR, MC-LV and MC-LL were found to be the major variants produced by strain CENA161, while the remaining variants were produced in trace amounts. The most studied and common variant MC-LR has been previously reported in the extract of *Fischerella* sp. CENA161 (Fiore et al., 2009) and *Fischerella* sp. NQAF311 (Cirés et al., 2014), whereas the latter also produces the MC-LA and MC-FR variants. The other four variants were previously found in *Microcystis* strains (Craig et al., 1993; Sivonen and Jones, 1999; Diehnelt et al., 2006).

The origin of MC diversity within species has been attributed to recombination events in the *mcy* genes, to the activity of specific tailoring enzymes and to the low substrate specificity of NRPS modules McyB<sub>1</sub> and McyC (Kurmayer and Gumpenberger, 2006; Fewer et al., 2007, 2008, 2013; Tooming-Klunderud et al., 2008; Kaasalainen et al., 2012; Calteau et al., 2014). As a consequence of this relaxed substrate specificity, different amino acids can be incorporated in positions X<sup>2</sup> and Z<sup>4</sup>, allowing a cyanobacterial strain to simultaneously produce several microcystin variants. In this regard, amino acid availability has been associated to the production of different MCs (Tonk et al., 2008; Van de Waal et al., 2010; Liu et al., 2016). These authors suggested that environmental factors affect the intracellular free amino acid levels which, ultimately, result in changes in the biosynthesis of MCs. Whether such correlations apply to the nitrogen-fixing strain CENA161 remains to be further investigated. Nevertheless, the MC production profile of CENA161 cultivated without inorganic nitrogen was dominated by hydrophobic variants with high C:N ratios. Considering the frequencies of amino acids incorporated in position X<sup>2</sup>, leucine was found in six of the seven variants described. These results are in line with the substrate prediction for McyB-A<sub>1</sub> (Table 3). However, McyC-A demonstrated a greater flexibility in terms of substrate selection, loading distantly related amino acids in position Z<sup>4</sup>, from hydrophilic arginine (MC-LR and MC-FR) to hydrophobic leucine (MC-LL).

The toxicity of MCs can be associated with the hydrophobicity of their constitutional amino acids. This probably occurs due to the increased ability of hydrophobic variants to get into cells by OATP-mediated transport or due to membrane interactions (Vestervik and Meriluoto, 2003; Feurstein et al., 2009; Faassen and Lüring, 2013). More toxic variants may harm cell membranes, damage mitochondrial dehydrogenases, and cause lactate dehydrogenase leakage (Monks et al., 2007; Fischer et al., 2010; Vestervik et al., 2012). Therefore, new variants of MC and the elucidation of their respective biosynthetic genes are considered of high interest for public health and pharmaceutical development since peptide structural diversities are reflected in different biological activities (Gupta et al., 2003; Zurawell et al., 2005; Monks et al., 2007; Feurstein et al., 2009; Fischer et al., 2010; Vestervik et al., 2012).

## 4. Conclusions

Identifying and annotating gene clusters responsible by the production of harmful toxic molecules that also have potential pharmacological application is important to the understanding of the process of their synthesis and the rules that govern their evolution, and to exploit their capabilities. The finding that MCs are also produced by subaerophytic cyanobacteria and not only by planktonic species may contribute to bring new insights into the cellular function of these heptapeptides, and as more gene clusters from different cyanobacteria are described it will facilitate the identification of amenable genus to genetic manipulation in order to advance our knowledge of genetic and biology of this toxin.

## Funding sources

This work was supported by grants from the São Paulo Research Foundation (FAPESP, 2013/50425-8) to MFF, the Academy of Finland (1273798) to KS and FAPESP (2014/50420-9) and the National Council for Scientific and Technological Development (CNPq, 311048/2016-1) to EP. KH and DOA were supported by graduate fellowships from the Brazilian Federal Agency for the Support and Evaluation of Graduate Education (CAPES, 99999.008344/2014) and FAPESP (2011/08092-6), respectively. MFF and AMV would also like to thank to the National Council for Scientific and Technological Development (CNPq) for research fellowships (310244/2015-3 and 302599/2016-9, respectively). We acknowledge the Center of Functional Genomics Applied to Agriculture and Agroenergy (University of São Paulo, Campus “Luiz de Queiroz”) for generating Illumina MiSeq data.

## Conflict of interest

The authors declare no conflicts of interest.

## Transparency document

Transparency document related to this article can be found online at <https://doi.org/10.1016/j.toxicon.2017.10.021>.

## Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.toxicon.2017.10.021>.

## References

- Allen, M.B., 1968. Simple conditions for growth of unicellular blue-green algae on plates. *J. Phycol.* 4, 1–4.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI BLAST: a new generation of protein database search programs. *Nucleic Acid. Res.* 25, 3389–3402.
- Andrews, S., 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. (Accessed 12 December 2016).
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Pribelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477.
- Birnboim, H.C., Doly, J., 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7, 1513–1518.
- Bishop, C.T., Anet, E.F.L.J., Gorham, P.R., 1959. Isolation and identification of the past-death factor in *Microcystis aeruginosa* NRC-1. *Can. J. Biochem. Physiol.* 37, 453–471.
- Botes, D.P., Tuinman, A.A., Wessels, P.L., Viljoen, C.C., Kruger, H., Williams, D.H., Santikarn, S., Smith, R.J., Hammond, S.J., 1984. The structure of cyanoginosin-LA, a cyclic heptapeptidotoxin from the cyanobacterium *Microcystis aeruginosa*. *J. Chem. Soc. Perkin Trans. 1*, 2311–2318.

- Botes, D.P., Wessels, P.L., Kruger, H., Runnegar, M.T.C., Santikarn, S., Smith, R.J., Barna, J.C.J., Williams, D.H., 1985. Structural studies on cyanoginosins-LR, -YR, -YA, and -YM, peptide toxins from *Microcystis aeruginosa*. J. Chem. Soc. Perkin Trans. 1, 2747–2748.
- Calteau, A., Fewer, D.P., Latifi, A., Coursin, T., Laurent, T., Jokela, J., Kerfeld, C.A., Sivonen, K., Piel, J., Gugger, M., 2014. Phylum-wide comparative genomics unravel the diversity of secondary metabolism in Cyanobacteria. BMC Genomics 15, 977.
- Challis, G.L., Ravel, J., Townsend, C.A., 2000. Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem. Biol. 7, 211–224.
- Christiansen, G., Fastner, J., Erhard, M., Börner, T., Dittmann, E., 2003. Microcystin biosynthesis in *Planktothrix*: genes, evolution, and manipulation. J. Bacteriol. 185, 564–572.
- Christiansen, G., Molitor, C., Philmus, B., Kurmayer, R., 2008. Nontoxic strains of cyanobacteria are the result of major gene deletion events induced by a transposable element. Mol. Biol. Evol. 25, 1695–1704.
- Cirés, S., Alvarez-Roa, C., Wood, S.A., Puddick, J., Loza, V., Heimann, K., 2014. First report of microcystin-producing *Fischerella* sp. (Stigonematales, Cyanobacteria) in tropical Australia. Toxicon 88, 62–66.
- Craig, M., McCready, T.L., Luu, H.A., Smillie, M.A., Dubord, P., Holmes, C.F.B., 1993. Identification and characterization of hydrophobic microcystins in Canadian freshwater cyanobacteria. Toxicon 31, 1541–1549.
- Dagan, T., Roettger, M., Stucken, K., Landan, G., Koch, R., Major, P., Gould, S.B., Goremykin, V.V., Rippka, R., de Marsac, N.T., Gugger, M., Lockhart, P.J., Allen, J.F., Brune, I., Maus, I., Pühler, A., Martin, W.F., 2012. Genomes of stigonematalean Cyanobacteria (Subsection V) and the evolution of psycogenic photosynthesis from prokaryotes to plastids. Genome Biol. Evol. 5, 31–44.
- Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2011. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27, 1164–1165.
- Diehnelt, C.W., Dugan, N.R., Peterman, S.M., Budde, W.L., 2006. Identification of microcystin toxins from a strain of *Microcystis aeruginosa* by liquid chromatography introduction into a hybrid linear ion trap Fourier transform ion cyclotron resonance mass spectrometer. Anal. Chem. 78, 501–512.
- Dittmann, E., Neilan, B.A., Erhard, M., von Döhren, H., Börner, T., 1997. Insertional mutagenesis of a peptide synthetase gene that is responsible for hepatotoxin production in the cyanobacterium *Microcystis aeruginosa* PCC 7806. Mol. Microbiol. 26, 779–787.
- Dittmann, E., Gugger, M., Sivonen, K., Fewer, D.P., 2015. Natural product biosynthetic diversity and comparative genomics of the cyanobacteria. Trends Microbiol. 23, 642–652.
- Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinforma. 5, 113.
- Eriksson, J.E., Meriluoto, J.A.L., Kujari, H.P., Skulberg, O.M., 1988. A comparison of toxins isolated from the cyanobacteria *Oscillatoria agardhii* and *Microcystis aeruginosa*. Comp. Biochem. Physiol. 89, 207–210.
- Ewing, B., Hillier, L., Wendt, M.C., Green, P., 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 8 (3), 175–185.
- Ewing, B., Green, P., 1998. Base-calling of automated sequencer traces using Phred. II. Error probabilities. Genome Res. 8, 186–194.
- Faassen, E.J., Lüring, M., 2013. Occurrence of the microcystins MC-LW and MC-LF in Dutch surface waters and their contribution to total microcystin toxicity. Mar. Drugs 11, 2643–2654.
- Feurstein, D., Holst, K., Fischer, A., Dietrich, D.R., 2009. Oatp-associated uptake and toxicity of microcystins in primary murine whole brain cells. Toxicol. Appl. Pharmacol. 234, 247–255.
- Fewer, D.P., Rouhiainen, L., Jokela, J., Wahlsten, M., Laakso, K., Wang, H., Sivonen, K., 2007. Recurrent adenylation domain replacement in the microcystin synthetase gene cluster. BMC Evol. Biol. 7, 183.
- Fewer, D.P., Tooming-Klunderud, A., Jokela, J., Wahlsten, M., Rouhiainen, L., Kristensen, T., Rohrlack, T., Jakobsen, K.S., Sivonen, K., 2008. Natural occurrence of microcystin synthetase deletion mutants capable of producing microcystins in strains of the genus *Anabaena* (Cyanobacteria). Microbiology 154, 1007–1014.
- Fewer, D.P., Wahlsten, M., Österholm, J., Jokela, J., Rouhiainen, L., Kaasalainen, U., Rikkinen, J., Sivonen, K., 2013. The genetic basis for O-acetylation of the microcystin toxin in cyanobacteria. Chem. Biol. 20, 861–869.
- Fiore, M.F., Genuário, D.B., Silva, C.S.P., Shishido, T.K., Moraes, L.A.B., Cantúcio-Neto, R., Silva-Stenico, M.E., 2009. Microcystin production by a freshwater spring cyanobacterium of the genus *Fischerella*. Toxicon 53, 754–761.
- Fischer, A., Hoeger, S.J., Stemmer, K., Feurstein, D.J., Knobloch, D., Nussler, A., Dietrich, D.R., 2010. The role of organic anion transporting polypeptides (OATPs/SLCOs) in the toxicity of different microcystin congeners in vitro: a comparison of primary human hepatocytes and OATP-transfected HEK293 cells. Toxicol. Appl. Pharmacol. 245, 9–20.
- Gordon, D., Abajian, C., Green, P., 1998. Consed: a graphical tool for sequence finishing. Genome Res. 8, 195–202.
- Gulledge, B.M., Aggen, J.B., Huang, H.B., Nairn, A.C., Chamberlin, A.R., 2002. The microcystins and nodularins: cyclic polypeptide inhibitors of PP1 and PP2A. Curr. Med. Chem. 9, 1991–2003.
- Gupta, N., Pant, S.C., Vijayaraghavan, R., Rao, P.V., 2003. Comparative toxicity evaluation of cyanobacterial cyclic peptide toxin microcystin variants (LR, RR, YR) in mice. Toxicology 188, 285–296.
- Harada, K.-I., Ogawa, K., Kimura, Y., Murata, H., Suzuki, M., Thorn, P.M., Evans, W.R., Carmichael, W.W., 1991. Microcystins from *Anabaena flos-aquae* NRC 525-17. Chem. Res. Toxicol. 4, 535–540.
- Heck, K., Machineski, G.S., Alvarenga, D.O., Vaz, M.G.M.V., Varani, A.M., Fiore, M.F., 2016. Evaluating methods for purifying cyanobacterial cultures by qPCR and high-throughput Illumina sequencing. J. Microbiol. Meth. 129, 55–60.
- Hicks, L.M., Moffitt, M.C., Beer, L.L., Moore, B.S., Kelleher, N.L., 2006. Structural characterization of in vitro and in vivo intermediates on the loading module of microcystin synthetase. ACS Chem. Biol. 1, 93–102.
- Hirose, Y., Fujisawa, T., Ohtsubo, Y., Katayama, M., Misawa, N., Wakazuki, S., Shimura, Y., Nakamura, Y., Kawachi, M., Yoshikawa, H., Eki, T., Kanesaki, Y., 2016. Complete genome sequence of cyanobacterium *Fischerella* sp. NIES-3754, providing thermoresistant optogenetic tools. J. Biotechnol. 220, 45–46.
- Izaguirre, G., Jungblut, A.D., Neilan, B.A., 2007. Benthic cyanobacteria (Oscillatoriaceae) that produce microcystin-LR, isolated from four reservoirs in southern California. Water Res. 41, 492–498.
- Jochimsen, E.M., Carmichael, W.W., An, J., Cardo, D., Cookson, S.T., Holmes, C.E.M., Antunes, M.B.C., Melo Filho, D.A., Lyra, T.M., Barreto, V., Azevedo, S.M.F.O., Jarvis, W.R., 1998. Liver failure and death following exposure to microcystin toxins at a hemodialysis center in Brazil. N. Engl. J. Med. 13, 873–878.
- Jungblut, A.D., Neilan, B.A., 2006. Molecular identification and evolution of the cyclic peptide hepatotoxins, microcystin and nodularin, synthetase genes in three orders of cyanobacteria. Arch. Microbiol. 185, 107–114.
- Kaasalainen, U., Fewer, D.P., Jokela, J., Wahlsten, M., Sivonen, K., Rikkinen, J., 2012. Cyanobacteria produce a high variety of hepatotoxic peptides in lichen symbiosis. Proc. Natl. Acad. Sci. U. S. A. 109, 5886–5891.
- Kounnis, V., Chondrogiannis, G., Mantzaris, M.D., Tzakos, A.G., Fokas, D., Papanikolaou, N.A., Galani, V., Sainis, I., Briassoulis, E., 2015. Microcystin LR shows cytotoxic activity against pancreatic cancer cells expressing the membrane OATP1B1 and OATP1B3 transporters. Anticancer Res. 35, 5857–5865.
- Krishnamurthy, T., Carmichael, W.W., Sarver, E.W., 1986. Investigation of freshwater cyanobacteria (blue-green algae) toxic peptide. I. Isolation, purification and characterization of peptides from *Microcystis aeruginosa* and *Anabaena flos-aquae*. Toxicon 24, 865–873.
- Kurmayer, R., Christiansen, G., Fastner, J., Börner, T., 2004. Abundance of active and inactive microcystin genotypes in populations of the toxic cyanobacterium *Planktothrix* spp. Environ. Microbiol. 6, 831–841.
- Kurmayer, R., Gumpenberger, M., 2006. Diversity of microcystin genotypes among populations of the filamentous cyanobacteria *Planktothrix rubescens* and *Planktothrix agardhii*. Mol. Ecol. 15, 3849–3861.
- Liu, J., Van Oosterhout, E., Faassen, E.J., Lüring, M., Helmsing, N.R., Van de Waal, D.B., 2016. Elevated pCO<sub>2</sub> causes a shift towards more toxic microcystin variants in nitrogen-limited *Microcystis aeruginosa*. FEMS Microbiol. Ecol. 92, fiv159.
- MacKintosh, C., Beattie, K.A., Klumpp, S., Cohen, P., Codd, G.A., 1990. Cyanobacterial microcystin-LR is a potent and specific inhibitor of protein phosphatases 1 and 2A from both mammals and higher plants. FEBS Lett. 264, 187–192.
- Marahiel, M.A., Stachelhaus, T., Mootz, H.D., 1997. Modular peptide synthetases involved in nonribosomal peptide synthesis. Chem. Rev. 97, 2651–2674.
- Meriluoto, J.A.O., Sandstrom, A., Eriksson, J.E., Remaud, G., Craig, A.G., Chattopadhyaya, J., 1989. Structure and toxicity of a peptide hepatotoxin from the cyanobacterium *Oscillatoria agardhii*. Toxicon 27, 1021–1034.
- Mikalsen, B., Boison, G., Skulberg, O.M., Fastner, J., Davies, W., Gabrielsen, T.M., Rudi, K., Jakobsen, K.S., 2003. Natural variation in the microcystin synthetase operon *mcyABC* and impact on microcystin production in *Microcystis* strains. J. Bacteriol. 185, 2774–2785.
- Moffitt, M.C., Neilan, B.A., 2004. Characterization of the nodularin synthetase gene cluster and proposed theory of the evolution of cyanobacterial hepatotoxins. Appl. Environ. Microbiol. 70, 6353–6362.
- Monks, N.R., Liu, S., Xu, Y., Yu, H., Bendelow, A.S., Moscow, J.A., 2007. Potent cytotoxicity of the phosphatase inhibitor microcystin LR and microcystin analogues in OATP1B1- and OATP1B3-expressing HeLa cells. Mol. Cancer Ther. 6, 587–598.
- Niedermeyer, T.H.J., Daily, A., Swiatecka-Hagenbruch, M., Moscow, J.A., 2014. Selectivity and potency of microcystin congeners against OATP1B1 and OATP1B3 expressing cancer cells. PLoS One 9 (3), e91476.
- Nishizawa, T., Ueda, A., Asayama, M., Fujii, K., Harada, K., Ochi, K., Shirai, M., 2000. Polyketide synthase gene coupled to the peptide synthetase module involved in the biosynthesis of the cyclic heptapeptide microcystin. J. Biochem. 127, 779–789.
- Omid, A., Esterhuizen-Londt, M., Pflugmacher, S., 2017. Still challenging: the ecological function of the cyanobacterial toxin microcystin – what we know so far. Toxin Rev. <https://doi.org/10.1080/15569543.2017.1326059>.
- Pearson, L.A., Hisbergues, M., Börner, T., Dittmann, E., Neilan, B.A., 2004. Inactivation of an ABC transporter gene, *mcyH*, results in loss of microcystin production in the cyanobacterium *Microcystis aeruginosa* PCC 7806. Appl. Environ. Microbiol. 70, 6370–6378.
- Pearson, L.A., Barrow, K.D., Neilan, B.A., 2007. Characterization of the 2-hydroxy-acid dehydrogenase *McyI*, encoded within the microcystin biosynthesis gene cluster of *Microcystis aeruginosa* PCC7806. J. Biol. Chem. 282, 4681–4692.
- Prinsep, M.R., Caplan, F.R., Moore, R.E., Patterson, G.M.L., Honkanen, R.E., Boynton, A.L., 1992. Microcystin-LA from a blue-green alga belonging to the stigonematales. Phytochemistry 31, 1247–1248.
- Rantala, A., Fewer, D.P., Hisbergues, M., Rouhiainen, L., Vaitomaa, J., Börner, T., Sivonen, K., 2004. Phylogenetic evidence for the early evolution of microcystin synthesis. Proc. Natl. Acad. Sci. U. S. A. 101, 568–573.
- Rausch, C., Weber, T., Kohlbacher, O., Wohlleben, W., Hudson, D.H., 2005. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). Nucleic Acids Res. 33, 5799–5808.

- Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- Röttig, M., Medema, M.H., Bin, K., Weber, T., Rausch, C., Kohbacher, O., 2011. NRPSpredictor2 — a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* 39, 362–367.
- Rouhiainen, L., Vakkilainen, T., Lumbye-Siemer, B., Buikema, W., Haselkorn, R., Sivonen, K., 2004. Genes coding for hepatotoxic heptapeptides (microcystins) in the cyanobacterium *Anabaena* strain 90. *Appl. Environ. Microbiol.* 70, 686–692.
- Rounge, T.B., Rohrlack, T., Nederbragt, A.J., Kristensen, T., Jakobsen, K.S., 2009. A genome-wide analysis of nonribosomal peptide synthetase gene clusters and their peptides in a *Planktothrix rubescens* strain. *BMC Genomics* 10, 396.
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., Barrell, B., 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16 (10), 944–945.
- Shih, P.M., Wu, D., Latifid, A., Axen, S.D., Fewer, D.P., Talla, E., Calteau, A., Cai, F., De Marsac, N.T., 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., 2013. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc. Natl. Acad. Sci. U. S. A.* 110, 1053–1058.
- Sielaff, H., Dittmann, E., de Marsac, N.T., Bouchier, C., von Döhren, H., Börner, T., Schwede, T., 2003. The *mcyF* gene of the microcystin biosynthetic gene cluster from *Microcystis aeruginosa* encodes an aspartate racemase. *Biochem. J.* 373, 909–916.
- Sivonen, K., Carmichael, W.W., Namikoshi, M., Rinehart, K.L., Dahlem, A.M., Niemela, S.I., 1990. Isolation and characterization of hepatotoxic microcystin homologs from the filamentous freshwater cyanobacterium *Nostoc* sp. strain 152. *Appl. Environ. Microbiol.* 56, 2650–2657.
- Sivonen, K., Jones, G., 1999. Cyanobacterial toxins. In: Chorus, I., Bartram, J. (Eds.), *Toxic Cyanobacteria in Water: a Guide to Their Public Health Significance, Monitoring and Management*. E & FN Spon, London, pp. 41–111.
- Stachelhaus, T., Mootz, H.D., Marahiel, M.A., 1999. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* 6, 493–505.
- Tillett, D., Dittman, E., Erhard, M., von Döhren, H., Börner, T., Neilan, B., 2000. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC 7806: as integrated peptide-polyketide synthetase system. *Chem. Biol.* 7, 753–764.
- Tonk, L., Van de Waal, D.B., Slot, P., Huisman, J., Matthijs, H.C.P., Visser, P.M., 2008. Amino acid availability determines the ratio of microcystin variants in the cyanobacterium *Planktothrix agardhii*. *FEMS Microbiol. Ecol.* 65, 383–390.
- Tooming-Klunderud, A., Mikalsen, B., Kristensen, T., Jakobsen, K.S., 2008. The mosaic structure of the *mcvABC* operon in *Microcystis*. *Microbiology* 154, 1886–1899.
- Vestervik, P.S.M., Meriluoto, J.A.O., 2003. Interaction between microcystins of different hydrophobicities and lipid monolayers. *Toxicon* 41, 349–355.
- Vestervik, P.S.M., Misiorek, J.O., Spoof, L.E.M., Toivola, D.M., Meriluoto, J.A.O., 2012. Comparative cellular toxicity of hydrophilic and hydrophobic microcystins on Caco-2 cells. *Toxins* 4, 1008–1023.
- Van de Waal, D.B., Ferreruela, G., Tonk, L., Van Donk, E., Huisman, J., Visser, P.M., Matthijs, H.C.P., 2010. Pulsed nitrogen supply induces dynamic changes in the amino acid composition and microcystin production of the harmful cyanobacterium *Planktothrix agardhii*. *FEMS Microbiol. Ecol.* 74, 430–438.
- von Döhren, H., Keller, U., Vater, J., Zocher, R., 1997. Multifunctional peptide synthetases. *Chem. Rev.* 29, 2675–2706.
- Zhang, J., Kobert, K., Flouri, T., Stamatakis, A., 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30 (5), 614–620.
- Zhbannikov, I.Y., Hunter, S.S., Foster, J., Settles, M.L., 2015. SeqyClean: a software tool for comprehensive preprocessing of sequence data. <https://github.com/ibest/seqyclean>. (Accessed 15 February 2017).
- Zurawell, R.W., Chen, H., Burke, J.M., Prepas, B.R., 2005. Hepatotoxic cyanobacteria: a review of the biological importance of microcystins in freshwater environment. *J. Toxicol. Environ. Health B* 8, 1–37.