

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Shishido, Tânia Keiko; Jokela, Jouni; Humisto, Anu; Suurnäkki, Suvi; Wahlsten, Matti; Alvarenga, Danilo O.; Sivonen, Kaarina; Fewer, David P.

Title: The Biosynthesis of Rare Homo-Amino Acid Containing Variants of Microcystin by a Benthic Cyanobacterium

Year: 2019

Version: Published version

Copyright: © The Authors, 2019.

Rights: CC BY 4.0

Rights url: <https://creativecommons.org/licenses/by/4.0/>

Please cite the original version:

Shishido, T. K., Jokela, J., Humisto, A., Suurnäkki, S., Wahlsten, M., Alvarenga, D. O., Sivonen, K., & Fewer, D. P. (2019). The Biosynthesis of Rare Homo-Amino Acid Containing Variants of Microcystin by a Benthic Cyanobacterium. *Marine Drugs*, 17(5), Article 271.
<https://doi.org/10.3390/md17050271>

Article

The Biosynthesis of Rare Homo-Amino Acid Containing Variants of Microcystin by a Benthic Cyanobacterium

Tânia Keiko Shishido ^{1,2}, Jouni Jokela ¹, Anu Humisto ¹, Suvi Suurnäkki ^{1,3}, Matti Wahlsten ¹, Danillo O. Alvarenga ¹, Kaarina Sivonen ¹ and David P. Fewer ^{1,*}

¹ Department of Microbiology, University of Helsinki, Viikinkaari 9, FI-0014 Helsinki, Finland; tania.shishido@helsinki.fi (T.K.S.); jouni.jokela@helsinki.fi (J.J.); anu.humisto@helsinki.fi (A.H.); matti.wahlsten@helsinki.fi (M.W.); danillo.oliveiradealvarenga@helsinki.fi (D.O.A.); kaarina.sivonen@helsinki.fi (K.S.)

² Institute of Biotechnology, Helsinki Institute of Life Science, University of Helsinki, Viikinkaari 5D, FI-0014 Helsinki, Finland

³ Department of Biological and Environmental Science, University of Jyväskylä, FI-40014 Jyväskylä, Finland

* Correspondence: david.fewer@helsinki.fi; Tel.: +358-9-19159270

Received: 1 April 2019; Accepted: 5 May 2019; Published: 7 May 2019

Abstract: Microcystins are a family of chemically diverse hepatotoxins produced by distantly related cyanobacteria and are potent inhibitors of eukaryotic protein phosphatases 1 and 2A. Here we provide evidence for the biosynthesis of rare variants of microcystin that contain a selection of homo-amino acids by the benthic strain *Phormidium* sp. LP904c. This strain produces at least 16 microcystin chemical variants many of which contain homophenylalanine or homotyrosine. We retrieved the complete 54.2 kb microcystin (*mcy*) gene cluster from a draft genome assembly. Analysis of the substrate specificity of McyB₁ and McyC adenylation domain binding pockets revealed divergent substrate specificity sequences, which could explain the activation of homo-amino acids which were present in 31% of the microcystins detected and included variants such as MC-LHty, MC-HphHty, MC-LHph and MC-HphHph. The *mcy* gene cluster did not encode enzymes for the synthesis of homo-amino acids but may instead activate homo-amino acids produced during the synthesis of anabaenopeptins. We observed the loss of microcystin during cultivation of a closely related strain, *Phormidium* sp. DVL1003c. This study increases the knowledge of benthic cyanobacterial strains that produce microcystin variants and broadens the structural diversity of known microcystins.

Keywords: adenylation domain; hepatotoxin; cyanobacteria; nonribosomal peptide synthetase (NRPS); polyketide synthase (PKS); mass spectrometry

1. Introduction

Microcystins are potent hepatotoxins, which inhibit eukaryotic protein phosphatases type 1 and 2A [1]. Microcystins are frequently detected during blooms of cyanobacteria in freshwater environments and are linked to poisoning of humans and animals [2,3]. Microcystins have an unusual chemical structure with more than 240 reported chemical variants [4]. The microcystin chemical structure can be summarized as cyclo-D-Ala¹-X²-D-MeAsp³-Z⁴-Adda⁵-D-Glu⁶-Mdha⁷ where X and Z are variable L-amino acids. The bulk of microcystin chemical variation can be attributed to the amino acids found at positions two (X) and four (Z).

Microcystins are synthesized by large hybrid multifunctional enzymes known as nonribosomal peptide synthetases (NRPS) and polyketide synthases (PKS) [5–8]. The chemical variability of

microcystins is mainly attributed to a substrate promiscuity of microcystin biosynthetic enzymes, which allows the incorporation of a diverse range of proteinogenic/non-proteinogenic amino acids, non-proteinogenic organic acids and different degrees of methylation [9]. This is due to differences in substrate specificity of the NRPS enzymes, the deletion of enzymatic domains or loss/gain of tailoring enzymes [1,8].

Homo-amino acids have an extra methylene (CH_2) group in the carbon side chain and are frequently found in cyanobacterial natural products [9–16]. Cyanobacteria have been found to produce microcystins that contain homo-amino acids in positions X and Z, such as homotyrosine (Hty), homophenylalanine (Hph), homoisoleucine (Hil), homoarginine (Har), homoserine (Hse) and homocysteine (Hcy) [4,10–13]. However, the biosynthetic origins of such microcystins are unclear.

Phormidium is a genus of cyanobacteria belonging to Oscillatoriales order that includes unbranched filamentous strains. *Phormidium* is not well defined taxonomically and further studies are needed to improve its separation with other genera of Oscillatoriales [17], which complicates the interpretation of reports about *Phormidium* in the scientific literature. Cyanobacteria belonging to *Phormidium* genus have been isolated from mangrove [18], saline-alkaline lakes [19], thermal springs [20], high altitude lakes [21] and mats/biofilms in Antarctic lakes [22] among others. *Phormidium* spp. isolated from Egypt, USA, Brazil and Spain have been reported to produce microcystins [13,23–25]. Putative *Phormidium* spp. producers of microcystin were reported to be associated to the poisoning of a dog in New Zealand [26]. However, analysis based on 16S rRNA sequences indicates a close similarity of these strains with *Planktothrix* [26]. Microcystin-LR was detected from *Phormidium* strains isolated from benthic environments in USA [24] and here we further analyze selected strains from that study. We describe the microcystin biosynthetic gene cluster from *Phormidium* sp. LP904c and show that this strain produces rare homo-amino acid containing microcystins in addition to microcystin-LR. We also report the loss of the microcystin synthesis by *Phormidium* sp. DVL1003c after 11 years of cultivation in our laboratory.

2. Results

2.1. *Phormidium* Strains Producing Unusual Microcystins (MCs)

Phormidium sp. LP904c is a benthic cyanobacterium isolated from the Lake Perris in Riverside County, California (USA) that is reported to produce microcystin-LR. This cyanobacterium was found to produce diverse microcystins varying in the positions X and Z (Figure 1A). Further variations were related to the presence or absence of methylation in the aspartic acid (Aa^3) or Adda (Aa^5) and for the presence of a methylated serine or dehydroalanine instead of the N-methyldehydroalanine (Aa^7) (Figure 2).

MC-LR was the main variant as previously reported [24] and accounted for 55% of the total microcystin produced by this strain. However, the strain produces a range of other microcystins including MC-LHty, MC-HphHty, MC-LHph and MC-HphHph (Figure 2, Table 1). A number of amino acids were found in position X and Z, such as Leu, Arg, Phe, Trp, Tyr, Met and unusual homo-amino acids such as homophenylalanine and homotyrosine (Figure 2, Table 1).

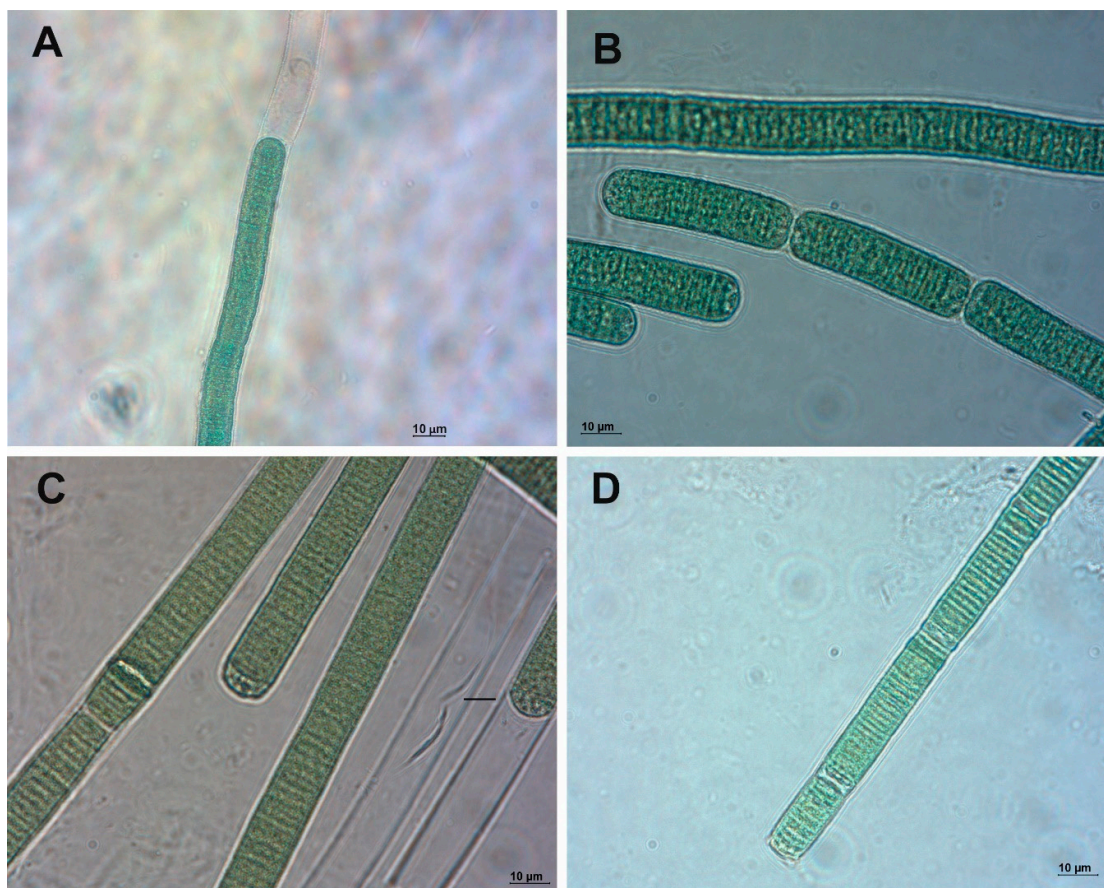


Figure 1. Photomicrography of the studied *Phormidium* strains. (A) LP904c. (B) DVL1003c. (C) LM603d. (D) LS703b.

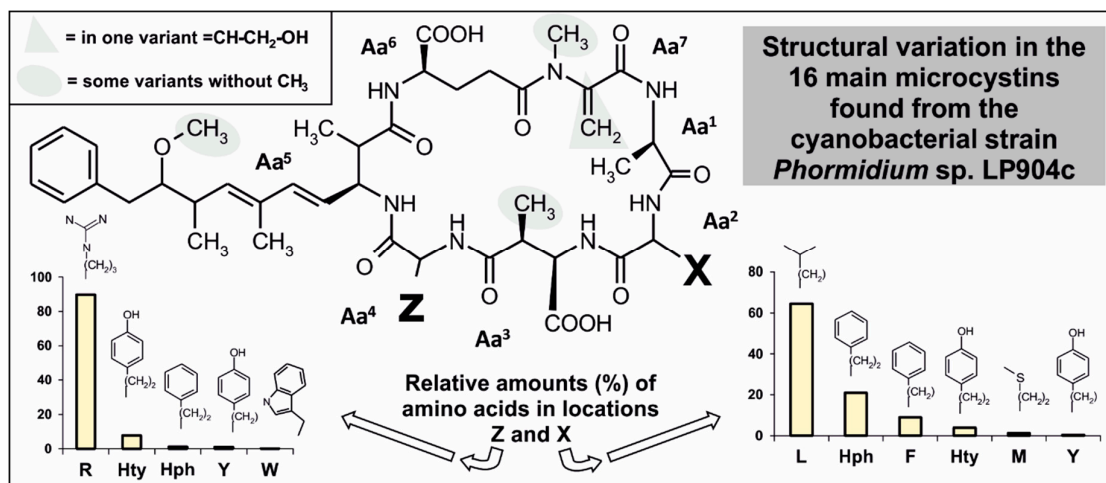


Figure 2. Structural variation in the 16 microcystins detected in the *Phormidium* sp. LP904c. The column charts indicate the relative amounts (%) of amino acids detected in the X and Z positions of microcystins. Hty = homotyrosine, Hph = homophenylalanine.

Phormidium sp. DVL1003c (Figure 1B) was isolated from a freshwater reservoir in Riverside County, California. Strains DVL1003c and LP904c were shown to synthesize similar microcystin variants in 2005 (Table 1, Supplementary Figures S1, S2 and S3). Structural characterizations of microcystins produced by *Phormidium* sp. LP904c were done using high-performance liquid chromatography/electrospray ionization ion-trap mass spectrometry (HPLC-ESI-ITMS) and ultra-

high performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UPLC-QTOF). MCs identified were divided to two groups, ten arginine (Arg) containing MCs (Numbers 1–10) and six non Arg MCs (Numbers 11–16) (Table 1). All Arg MCs eluted first, followed by the more hydrophobic non Arg MCs (Supplementary Figures S2 and S3). These two MC groups fragment differently as protonated molecules resulting in distinct product ion spectra (Supplementary Figure S4). Both spectral groups were highly similar to spectra recorded in similar ion trap conditions [27]. In Arg containing MC structures fragmentation prefers the formation of ion 4a–6c (relative intensities (RI) from 51 to 100%, Arg⁴-Adda⁵-Glu⁶, Supplementary Figures S4 and S5) which have an *m/z* 599 when methyl group 2 is present. When the methyl group is absent, e.g., [DMAdda⁵]MC, an *m/z* 585 is detected. Serine in MC structure generates a base peak of [M + H – H₂O]⁺ which was seen in the spectra of [MeSer⁷]MC-LR and which lowered the relative intensity of 4a–6c ion (Supplementary Figure S4). Ions 4a–6c – CO (RI 26–94%) and 4a–6c – NH₃ (RI 6–28%) were also present (Supplementary Figure S4). The presence of ion 4a–7c (RI 4–20%, Arg⁴-Adda⁵-Glu⁶-Aa⁷, Supplementary Figure S5) shows the identity of the Aa⁷. Other diagnostic intense ions were 4a–2c/7a–5c (RI 8–38%) which showed the identity of amino acids 3 and 6, and 7a–4c (RI 45–94%, Aa⁷-Ala¹-Aa²-MeAsp³-Arg⁴) together with other ions showed the identity of amino acids 4 and 5. Ion 5y–5x–7c was present in all the protonated MC product ion spectra except in the [DMAdda⁵]MC-LR spectra because cleavage at position 5y is not so favorable when there is hydroxyl group instead of the methoxy group vicinal to the 5y bond [28]. In the presence of methyl group 3 the ion mass is *m/z* 375 or 393 when Aa⁷ is MeSer and when it is absent the mass is *m/z* 361, such as in the [Dha⁷]MC-LR. Ion 5y–5x–7c identifies amino acids 5, 6 and 7. Identification of amino acids 1 and 2 using MS is challenging because the intensities of the diagnostic ions are near the noise level but for example ion 4a–1c (*m/z* 753) was clearly present in the spectrum of MC-MR showing the presence of methionine in this microcystin (Supplementary Figures S4 and S5).

Table 1. Protonated ion masses of microcystin (MC) variants from *Phormidium* sp. LP904c and their relative amounts (%) identified from either *Phormidium* sp. LP904c or DVL1003c by HPLC-ESI-ITMS. Peak areas of the extracted ion chromatograms of the protonated microcystins were used in relative amount calculations. Hty = homotyrosine, Hph = homophenylalanine, [M + H]⁺ (*m/z*) = mass per charge value of protonated microcystins, R_t (min) = retention time in minutes, RA = relative amount, Aa = amino acid. Additional information presented in Figure S1.

No.	Microcystin	Aa in Position		R _t (min)	[M + H] ⁺ (<i>m/z</i>)		Error (ppm)	RA (%)	
		X	Z		Calculated	Measured		LP904c	DVL1003c
1	MC-YR	Y	R	15.4	1045.5353	1045.5329	−2.35	<1	<1
2	MC-HtyR	Hty	R	15.8	1059.5510	1059.5510	−0.01	4	4
3	MC-MR	M	R	15.8	1013.5125	1013.5148	2.26	1	1
4	[DMAdda ⁵]MC-LR	L	R	15.8	981.5404	981.5404	−0.04	<1	<1
5	[L-MeSer ⁷]MC-LR	L	R	16.6	1013.5666	1013.5649	−1.73	<1	<1
6	[D-Asp ³]MC-LR	L	R	17	981.5404	981.5404	−0.04	<1	<1
7	MC-LR	L	R	17	995.5560	995.5566	0.51	55	49
8	MC-FR	F	R	17.3	1029.5404	1029.5401	−0.33	9	7
9	[Dha ⁷]MC-LR	L	R	17.5	981.5404	981.5404	−0.04	<1	<1
10	MC-HphR	Hph	R	18.7	1043.5560	1043.5561	0.01	17	14
11	MC-LY	L	Y	22.8	1002.5183	1002.5178	−0.51	1	2
12	MC-LHty	L	Hty	23.2	1016.5339	1016.5347	0.73	5	12
13	MC-HphHty	Hph	Hty	24.3	1064.5339	1064.5337	−0.24	3	8
14	MC-LW	L	W	24.6	1025.5342	1025.5343	0.00	<1	<1
15	MC-LHph	L	Hph	26.7	1000.5390	1000.5369	−2.15	<1	1
16	MC-HphHph	Hph	Hph	27.7	1048.5390	1048.5403	1.20	<1	<1

The base peak is [M + H – C₉H₁₀O]⁺ in the product ion spectra of protonated non Arg microcystins [29]. In the *Phormidium* sp. LP904c microcystins spectra, the most intense peak was [M + H – NH₃]⁺ followed by [M + H – C₉H₁₀O]⁺ (same as 5y), [M + H – (NH₃ + C₉H₁₀O)]⁺ (same as 5y–5x–4c) and [M + H – H₂O]⁺ (Supplementary Table S1). The ion 5y–5x–7c was present in all protonated MC product ion spectra identifying amino acids 5, 6 and 7 (Supplementary Figure S4). Medium intensity ions 5y–5x–1c (*m/z* 446) and 5y–5x–2c, which are present in all non Arg MCs, identify the amino acids 1 and 2.

Ions 5y-5x-3c and 5y-5x-4c identify amino acids 3 and 4. The low intensity of ion 5y-5x-3c in some MCs weakened the certainty of amino acids 2 and 4. Product ion spectra of sodiated non-Arg microcystins are presented in Supplementary Figure S6 and product ion assignments are presented in Supplementary Table S2. These data support the assignments of the protonated product ions but do not clearly enhance the identity of amino acids 2 and 4. Additional proof for the MC structures gave the accurate masses measured from the protonated molecules error being lower than ± 2.4 ppm (Table 1). The immonium and other diagnostic ions of amino acids 2, 4 measured with UPLC-QTOF were also in agreement with the presented MC structures (Supplementary Table S3).

In addition to the mass spectral behavior, the chromatographic retention behavior was consistent with the structures of microcystins produced by *Phormidium* spp. DVL1003c and LP904c (Supplementary Figures S2 and S3). Polar surface areas obtained with the topological method from the microcystin 3D molecular structures correlated well ($R^2 = 0.951$) with the measured retention times obtained from reversed phase chromatography (Figure 3 and Supplementary Figure S7). Large macrocyclic structures with many polar groups did not give the best polar surface areas with a topological method. Despite this, Topological Polar Surface Area (TPSA) values of microcystins, which are macrocyclic and polar groups containing molecules correlated well with the retention times. These results show that there is no discrepancy between the derived microcystin structures and their surface polarities.

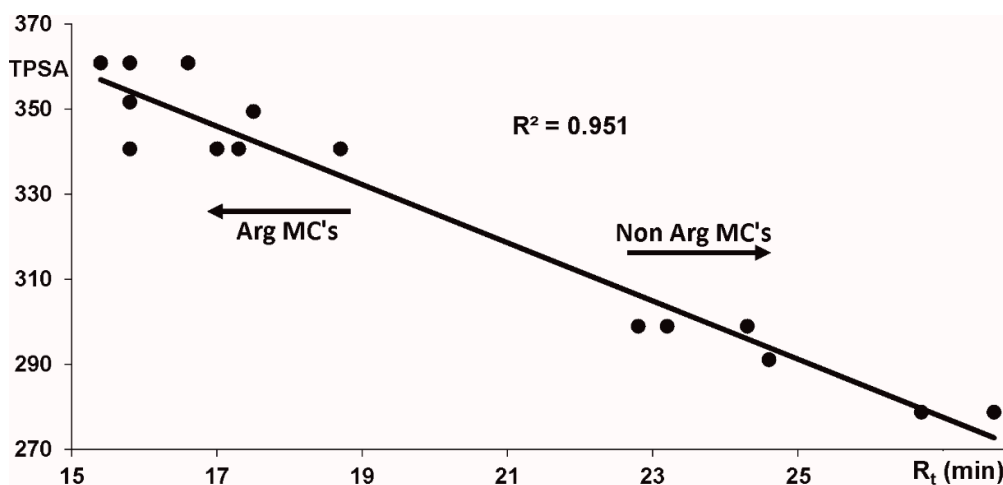


Figure 3. Effect of polar surface areas (PSA) to the retention times (R_t , min) of *Phormidium* sp. LP904c microcystins calculated with a topological polar surface area (TPSA) method [30]. High correlation ($R^2 = 0.951$) shows that the proposed microcystin structures fit well to the measured retention times.

Microcystin contents of *Phormidium* spp. LM603a (5–11b), LM603d (12–29d) (Figure 1C), and LS703b (1–2, Figure 1D) strains were very similar with the *Phormidium* spp. DVL1003c and LP904c (Supplementary Figure S8). However, recent chemical analysis indicates that microcystin is no longer produced by *Phormidium* sp. DVL1003c and we decided to choose strains LP904c for further analysis.

2.2. Microcystin and Homophenylalanine Biosynthetic Genes from *Phormidium*

A draft genome sequence from *Phormidium* sp. LP904c (7.7 Mb, 108 scaffolds, 129,707 N50) was obtained to identify the microcystin biosynthetic gene cluster (*mcy*). The complete 54.2 kb *mcy* gene cluster was identified through tBLASTn searches (Figure 4). The microcystin biosynthetic gene cluster from *Phormidium* sp. LP904c encodes ten genes organized in a bidirectional operon (Figure 4). There are three NRPS genes (*mcyA*, *mcyB* and *mcyC*), one PKS genes (*mcyD*), two hybrid NRPS-PKS gene (*mcyE* and *mcyG*), an aspartate racemase (*mcyF*), an ABC transporter (*mcyH*), a dehydrogenase (*mcyI*) and a methyltransferase (*mcyJ*) (Figure 4 and Supplementary Table S4). Most of the protein sequences from the enzymes of the microcystin biosynthetic pathway in *Phormidium* sp. LP904c are similar to sequences from *Planktothrix* spp. (Supplementary Table S4). The organization of *mcy* genes

is identical to the *Planktothrix agardhii* CYA126 except for the positions of *mcyI* and *mcyJ* (Figure 4). Many of the microcystin variants produced by *Phormidium* sp. LP904c contain homo-amino acids (Figure 2 and Table 1). However, the *mcy* biosynthetic gene cluster from *Phormidium* sp. LP904c lacked obvious enzymes for the biosynthesis of homo-amino acids (Figure 4).

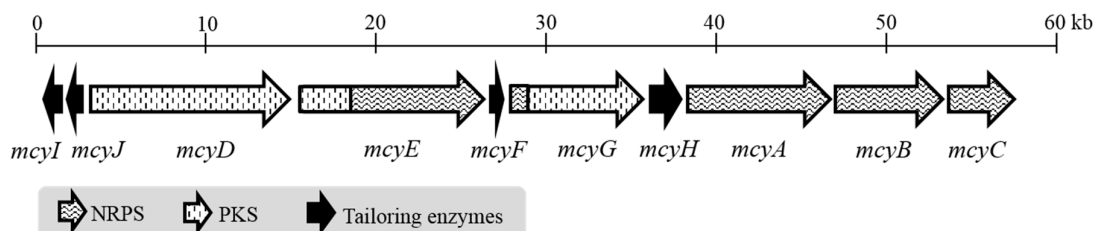


Figure 4. The 54.2 kb microcystin (*mcy*) gene cluster from *Phormidium* sp. LP904c. The respective sequences that will be transcribed to NRPS (nonribosomal peptide synthetase), PKS (polyketides synthase) and tailoring enzymes are indicated in the gene cluster.

2.3. *McyB1* and *McyC* from *Phormidium* sp. LP904c

The adenylation domains of *McyB1* and *McyC* are responsible for the selection and activation of the amino acids in the positions X and Z of the microcystin, respectively. Sequence information based on microcystin biosynthetic genes from *Phormidium* sp. DVL1003c was obtained using PCR products based on specific primers prior to loss of microcystin biosynthesis by this strain [10]. These DNA sequences were translated to amino acid sequences and the predicted adenylation domain binding pockets of *McyB1* and *McyC* from different cyanobacteria were obtained (Tables 2 and 3). Both *McyB1* and *McyC* adenylation domains binding pockets from *Phormidium* sp. LP904c and *Phormidium* sp. DVL1003c differ at the positions 236, 239 and 278 (in addition to position 331 in the *McyB1* and exception in position 278 in *McyC* of DVL1003c) when compared to the same adenylation domains from other strains previously described in the literature (Tables 2 and 3).

Phylogenetic analysis based on the conserved sequences of 16S rRNA genes indicates the relationship of the studied strain to the other cyanobacteria sequences available in the National Center for Biotechnology Information (NCBI) GenBank database (Figure 5). *Phormidium* spp. DVL1003c and LP904c are grouped with other *Phormidium*, *Oscillatoria* and *Lyngbya* strains in the 16S rRNA phylogenetic tree (Figure 5). Other strains producing microcystin belonging to the order Nostocales, Oscillatoriales, Stigonematales and Chroococcales were included in the phylogenetic tree of the 16S rRNA gene and a broad distribution of these strains based on their relatedness to similar strains can be seen (Figure 5).

The concatenated phylogenetic tree based on microcystin biosynthetic enzymes *McyD* and *McyE* indicates the relationship of the *Phormidium* sp. LP904c sequences and other cyanobacteria producers of microcystins (Supplementary Figure S9). The studied strains are grouped with close related *Phormidium* and *Planktothrix* strains. However, recombination can be seen analyzing the phylogenetic history based on condensation and adenylation domains from the *McyB1* and *McyC* (Supplementary Figure S10). Condensation domains sequences are divided into two big groups according to each enzyme: *McyB1* and *McyC* (Supplementary Figure S10A). In contrast, the adenylation domain phylogenetic tree shows the recombination events occurred between the gene sequences (Supplementary Figure S10B). Adenylation domain from *Phormidium* spp. LP904c and DVL1003c *McyB1* is grouped with domains from *McyB1* and *McyC* from *Anabaena/Dolichospermum* and *Hapalosiphon* (Supplementary Figure S10B), while the adenylation domain from *McyC* in *Phormidium* sp. LP904c and DVL1003c is grouped with sequences from the *McyB1* domains from *Planktothrix* spp.

Table 2. Specificity codes inferred from the protein sequence of first adenylation domain of McyB₁ according to [31] or NRPSpredictor2 [32,33]. Conserved amino acids are highlighted in light blue and variable amino acids are highlighted in dark green. Homo-amino acids are highlighted in red.

Organism	Strain	McyB ₁										P*	%	Aa #
		235	236	239	278	299	301	322	330	331	517			
<i>Phormidium</i> sp.	DVL1003c	D	I	C	V	F	G	L	V	H	K	Ser-Thr	60	Leu, Hph , Phe, Hty , Tyr, Arg
<i>Planktothrix agardhii</i>	CYA 126/8	D	A	L	F	F	G	L	V	D	K	Leu	100	Arg, Leu
<i>Planktothrix agardhii</i>	213	D	A	L	F	F	G	L	V	D	K	Leu	100	Arg, Leu
<i>Planktothrix agardhii</i>	NIVA-CYA56/3	D	A	L	F	F	G	L	V	D	K	Leu	100	Leu, Arg, Tyr
<i>Planktothrix prolifica</i>	NIVA-CYA 98	D	A	L	F	F	G	L	V	D	K	Leu	100	Leu, Arg
<i>Planktothrix rubescens</i>	NIVA-CYA 407	D	A	L	L	F	G	L	V	D	K	Leu	90	Leu, Arg, Hty
<i>Dolichospermum</i> sp.	90	D	V	W	F	F	G	L	V	D	K	Ser	80	Leu, Arg, Hil
<i>Dolichospermum flos-aquae</i>	18B6	D	V	W	S	F	G	L	V	D	K	Ser	80	Arg, X
<i>Dolichospermum lemmermannii</i>	66 A	D	V	W	S	F	G	L	V	Y	K	Ser	70	Hty , X, Hph , Leu, Tyr, Phe
<i>Nostoc</i> sp.	152	D	A	L	F	F	G	L	I	Y	K	Leu	80	Leu, Hil , X, Val
<i>Nostoc</i> sp.	IO-102-I	D	I	K	N	F	G	A	I	V	K	Gln	50	Leu, X, Phe, Hil , Tyr
<i>Fischerella</i> sp.	PCC9339	D	V	L	I	F	G	L	I	Y	K	Pro	70	Leu
<i>Hapalosiphon hibernicus</i>	BZ-3-1	D	V	W	F	F	G	L	V	D	K	Ser	80	Leu, Arg
<i>Microcystis aeruginosa</i>	PCC 7806	D	A	W	F	L	G	N	V	V	K	Leu	100	Val
<i>Microcystis aeruginosa</i>	FCY-28	D	G	W	T	I	G	A	V	E	K	Arg	90	Leu
<i>Microcystis aeruginosa</i>	FCY-26	D	G	W	T	I	G	A	V	E	K	Arg	90	Ni
<i>Microcystis aeruginosa</i>	UV027	D	V	W	T	I	G	A	V	E	K	Arg	100	Arg
<i>Microcystis aeruginosa</i>	K139	D	A	W	F	L	G	N	V	V	K	Leu	100	Leu
<i>Microcystis aeruginosa</i>	DIANCHI905	D	A	W	F	L	G	N	V	V	K	Leu	100	Ni
<i>Microcystis aeruginosa</i>	PCC 9807	D	A	W	F	L	G	N	V	V	K	Leu	100	Ni
<i>Microcystis aeruginosa</i>	PCC 7941	D	A	W	F	L	G	N	V	V	K	Leu	100	Ni
<i>Microcystis aeruginosa</i>	PCC 9443	D	G	W	T	I	G	A	V	E	K	Arg	90	Ni
<i>Microcystis aeruginosa</i>	NIES-843	D	G	W	T	I	G	A	V	E	K	Arg	90	Arg, Leu, Tyr
<i>Microcystis aeruginosa</i>	SPC777	D	G	W	T	I	G	A	V	E	K	Arg	90	Arg
<i>Microcystis viridis</i>	NIES 102	D	G	W	T	I	G	A	V	E	K	Arg	90	Hil , Leu, Arg, Tyr, Trp, Phe, Hty , X

P* prediction by NRPSpredictor2; Aa # amino acid detected by LC-MS (Supplementary Table S5); X. MC contains an unknown amino acid or the overall amino acid content is not known; Ni. No information.

Table 3. Specificity codes inferred from the protein sequence of the adenylation domain of MeyC according to [31] or NRPS predictor2 [32,33]. Conserved amino acids are highlighted in light blue and variable amino acids are highlighted in dark green. Homo-amino acids are highlighted in red.

Organism	Strain	MeyC										P *	%	Aa #
		235	236	239	278	299	301	322	330	331	517			
<i>Phormidium</i> sp.	DVL1003c	D	A	L	C	F	G	L	V	D	K	Leu	100	Arg, Hty , Hph , Trp
<i>Planktothrix agardhii</i>	CYA 126/8	D	P	W	G	F	G	L	V	D	K	Gln	70	Arg
<i>Planktothrix agardhii</i>	213	D	P	W	C	F	G	L	V	D	K	Gln	70	Arg
<i>Planktothrix agardhii</i>	NIVA-CYA 56/3	D	P	W	G	F	G	L	V	D	K	Gln	70	Arg
<i>Planktothrix prolifica</i>	NIVA-CYA 98	D	P	W	G	F	G	L	V	D	K	Gln	70	Arg
<i>Planktothrix rubescens</i>	NIVA-CYA 407	D	P	W	G	F	G	L	V	D	K	Gln	70	Arg
<i>Dolichospermum</i> sp.	90	D	V	W	C	F	G	L	V	D	K	Ser	80	Arg
<i>Dolichospermum flos-aquae</i>	18B6	D	V	W	S	F	G	L	V	D	K	Ser	80	Arg
<i>Dolichospermum lemmermannii</i>	66 A	D	V	W	S	F	G	L	V	D	K	Ser	80	Arg
<i>Nostoc</i> sp.	152	D	V	W	N	F	G	F	I	D	K	Gln	70	Arg, Har
<i>Nostoc</i> sp.	IO-102-I	D	V	W	N	F	G	F	V	D	K	Glu	70	Arg
<i>Fischerella</i> sp.	PCC9339	D	V	W	F	F	G	L	V	D	-	Ser	70	Arg
<i>Hapalosiphon hibernicus</i>	BZ-3-1	D	V	W	F	F	G	L	V	D	K	Ser	80	Ala, Leu, Val
<i>Microcystis aeruginosa</i>	PCC 7806	D	V	W	T	I	G	A	V	D	K	Arg	100	Arg
<i>Microcystis aeruginosa</i>	FCY-28	D	V	W	T	I	G	A	V	D	K	Arg	100	Ni
<i>Microcystis aeruginosa</i>	FCY-26	D	V	W	T	I	G	A	V	D	K	Arg	100	Ni
<i>Microcystis aeruginosa</i>	UV027	D	V	W	T	I	G	A	V	D	K	Arg	100	Arg
<i>Microcystis aeruginosa</i>	K139	D	V	W	T	I	G	A	V	E	K	Arg	100	Arg
<i>Microcystis aeruginosa</i>	DIANCHI905	D	V	W	T	I	G	A	V	D	K	Arg	100	Ni
<i>Microcystis aeruginosa</i>	PCC 9807	D	V	W	T	I	G	I	V	D	K	Arg	90	Ni
<i>Microcystis aeruginosa</i>	PCC 7941	D	V	W	T	I	G	A	V	D	K	Arg	100	Ni
<i>Microcystis aeruginosa</i>	PCC 9443	D	V	W	T	I	G	I	V	D	K	Arg	90	Ni
<i>Microcystis aeruginosa</i>	NIES-843	D	V	W	T	I	G	A	V	D	K	Arg	100	Arg
<i>Microcystis aeruginosa</i>	SPC777	D	V	W	T	I	G	A	V	D	K	Arg	100	Arg
<i>Microcystis viridis</i>	NIES 102	D	V	W	T	I	G	A	V	D	K	Arg	100	Arg

P * prediction by NRPSpredictor2; Aa # amino acid detected by LC-MS (Supplementary Table S5); X. MC contains an unknown amino acid or the overall amino acid content is not known; Ni. No information.

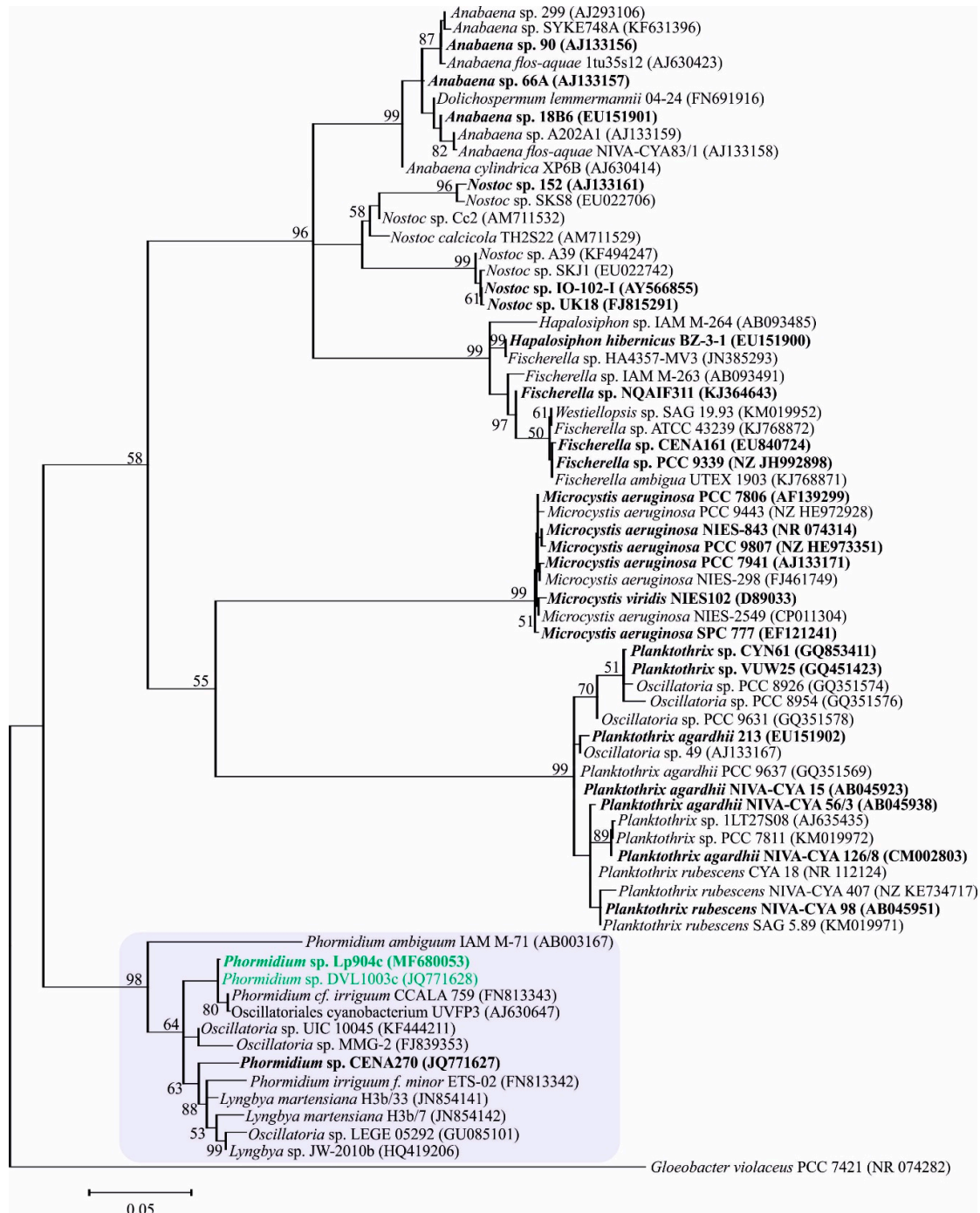


Figure 5. A phylogenetic tree of the 16S rRNA gene sequences constructed using maximum likelihood with 1000 bootstrap. Strains detected producing microcystin are highlighted in bold and the *Phormidium* spp. DVL1003c and LP904c studied here are highlighted in green.

3. Discussion

3.1. Benthic Cyanobacteria Producing Unusual Microcystins

Cyanobacteria are known for the production of toxins implicated in the toxicosis of animals and humans [2]. The hepatotoxin microcystin producing cyanobacteria have been detected mainly from planktonic strains present in freshwater environments distributed worldwide [2]. However, microcystins have been reported from benthic *Fischerella* and *Phormidium* strains [13,23–26,34,35].

Phormidium sp. LP904c was previously described to produce a high amount ($441 \pm 109 \mu\text{g L}^{-1}$) of microcystin and the main variant detected was microcystin-LR [24]. The present study shows the benthic *Phormidium* sp. LP904c producing altogether 16 microcystins. Different isolates from diverse genera of cyanobacteria have been related to the hepatotoxin production, such as *Dolichospermum* spp., *Nostoc* sp., *Plectonema* sp. and *Phormidium* sp. from Egypt [23], *Phormidium* sp. from USA, New Zealand and Brazil [13,24,26] and *Fischerella* sp. from Brazil and Australia [34–36]. Microcystin has been detected in benthic environmental samples from diverse places worldwide, including Switzerland [37], Spain [38,39], Australia [40], Antarctica [41] and the Arctic [42]. These wide distributions of toxic samples and the rise of the number of cyanobacterial isolates from diverse places in the world highlights the need to monitor these benthic communities for toxin production.

3.2. Microcystins Containing Homo-Amino Acids

The results show *Phormidium* sp. LP904c producing, in addition to the main variant microcystin-LR (55%), a variation of amino acids in position X and Z (Figure 2, Table 1). Leucine, arginine, phenylalanine, tryptophan, tyrosine, methionine and rare homo-amino acids homophenylalanine and homotyrosine were detected (Figure 2, Table 1). Microcystins varying in different parts of its chemical structure were also detected, such as D-Asp³, L-MeSer⁷, Dha⁷ and DMAdda⁵ (Figure 2, Table 1). Microcystins are synthesized by nonribosomal peptide synthetases and polyketide synthases [5,9]. These nonribosomal pathways are known for the ability to incorporate proteinogenic and non-proteinogenic amino acids to the growing peptide. *Phormidium* sp. LP904c has been shown in this study to produce diverse variants of microcystins containing homotyrosine and homophenylalanine at positions X and Z (Figure 2 and Table 1). Altogether, 31% of the microcystin variants produced contained homo-amino acids. The homo-amino acids are 14 mass units bigger than the proteinogenic amino acid due the presence of a methylene (CH₂) group in the carbon side chain. *Phormidium* sp. CENA270 isolated from Brazil produces a range of [D-Leu¹] microcystin variants, including homoarginine at the position Z [13]. Homotyrosine has been found in diverse cyanobacterial compounds, such as aeruginosins, cyanopeptolins and microginins [9,43]. Homophenylalanine was detected in pahayokolides [44] and microcystin [45]. Some microcystins, spumigins, lyngbyazothrins, tychonamides, schizotrin A, portoamides and anabaenopeptins contain homophenylalanine and/or homotyrosine in their chemical structure [9,15,46–48]. HphA, HphB and HphCD were shown to be involved in the synthesis of homophenylalanine from L-phenylalanine in *Nostoc punctiforme* PCC73102 [14]. Interestingly, *Escherichia coli* containing the cyanobacterial *hph* genes could also convert L-tyrosine to homotyrosine [14]. The *Phormidium* sp. LP904c has the *hphABCD* genes present in another part of the genome as part of the anabaenopeptin gene cluster [16]. The *mcy* biosynthetic gene cluster of *Phormidium* sp. LP904c did not encode enzymes for the biosynthesis of homo-amino acids and it is plausible that the *mcyB*₁ and *McyC* adenylation domains activate the homo-amino acids produced as part of the biosynthesis of anabaenopeptins.

3.3. Microcystin Gene Cluster from *Phormidium*

The microcystin gene cluster has been described from *Microcystis*, *Planktothrix*, *Dolichospermum*, *Nostoc* and *Fischerella* [5–8,49]. In this study, the microcystin biosynthetic gene cluster from *Phormidium* sp. LP904c is presented. The *Phormidium* microcystin biosynthetic enzymes have high similarity to *Planktothrix* amino acid sequences and it is as well mostly unidirectional, with exception of *mcyI* (dehydrogenase) and *mcyJ* (methyltransferase) (Figure 4 and Supplementary Table S4). However, the thioesterase *McyT* present in the *Planktothrix agardhii* CYA 126 [6] is absent in the *Phormidium* microcystin gene cluster. The new *mcy* gene cluster sequence further expands the known genetic diversity of the microcystin family and will allow methods to detect the toxin based on gene sequences to be further refined.

The adenylation domains from the NRPSs *McyB*₁ and *McyC* are responsible for the selection and activation of the amino acid to be incorporated in the positions X and Z of microcystin, respectively [5–8]. In the present study, the gene regions of *mcyB*₁ and *mcyC* from *Phormidium* sp. LP904c were analyzed to address their variation compared with previously described adenylation domains

binding pockets. The adenylation domains have eight to ten amino acids responsible for the substrate specificity in the NRPS [31,50–51]. The binding pocket in the McyB1 and McyC adenylation domains from *Phormidium* sp. LP904c are unusual compared to other adenylation domains from other strains producing microcystins (Tables 2 and 3). The McyC binding pocket is predicted to activate leucine, even though no microcystins with a leucine in the position Z were detected. The adenylation domains of McyA1, McyA2, McyB1 and McyC were found to be affected by recombination events which might have resulted in the high number of chemical variants in the positions 7, 1, 2 and 4 of the microcystin, respectively [10,11,13,52–57]. The adenylation domain sequences of McyB1 and McyC from *Phormidium* sp. LP904c might have gone through recombination events because their sequences group with sequences from different enzymes (Supplementary Figure S10). Adenylation domains from nostopeptolide [58] and cyanopeptolin [59] biosynthetic pathways have been shown to activate more than one amino acid. The synthesis of high numbers of microcystin variants could be attributed to the multispecificity of the adenylation domains, specificity-regulation and gatekeeping function of the condensation domains, tolerance of the condensation domains for recombination and point mutations after recombination events that change the specificity of the adenylation-condensation domains [60]. Further analysis based on biochemical assays of adenylation domains could improve the knowledge of amino acids incorporated by these systems.

3.4. Lack of Microcystin from *Phormidium* sp. DVL1003c

Phormidium sp. DVL1003c was first analyzed by our group for the synthesis of microcystin in 2005 and documented on the microcystin variants and *mcy* synthetase genes. In addition, Izaguirre and collaborators (2007) reported the synthesis of microcystin-LR by this strain [24]. Surprisingly, the culture maintained in the University of Helsinki Culture Collection is no longer able to produce microcystins. Other *Phormidium* strains which have been cultivated together with DVL1003c in identical conditions still produce almost same microcystins variants as DVL1003c was able to produce (Supplementary Figure S8). Thus, this study presents one example of a cyanobacterial strain which loses the ability to synthesize a natural product under laboratorial conditions.

3.5. Diversity of Benthic *Phormidium* Strains

Phormidium spp. isolated from Egypt, USA, Brazil and Spain have been reported to produce microcystins [23–25]. In the phylogenetic tree constructed with 16S rRNA gene sequences, microcystin-producing strains of *Phormidium* are grouped together with other Oscillatoriales such as *Oscillatoria* and *Lyngbya* (Figure 5). The low taxonomic agreement among these strains opens the possibilities for discovery of new benthic cyanobacteria yet to be assigned to some genera and that can be a potential microcystin producer.

4. Materials and Methods

4.1. Cultivation of *Phormidium* Strains

The *Phormidium* strains were isolated from Riverside County, California, USA: DVL1003c was isolated from a freshwater reservoir, strains LM603a (5–11b) and LM603d (12–29d) were isolated from Lake Mathews, strain LP904c (3–7b) from Lake Perris and strain LS703b (1–2) from Lake Skinner [24]. The cultures were grown in 50 mL of Z8 medium at 22 °C under constant light.

4.2. Chemical Analysis

Freeze dried cyanobacterial biomass was extracted with methanol as previously described [13]. Methanol extracts were analyzed with low resolution HPLC-ESI-ITMS (Agilent 1100 Series LC/MSD Ion Trap XCT Plus, Agilent Technologies, Palo Alto, CA, USA) to obtain the chromatographic data and product ion spectra of protonated microcystins. A 10 µL sample was injected into a Zorbax C8 column (4.6 × 150 mm, 5 µm, Agilent technologies, Palo Alto, CA, USA) which was eluted for 7.5 min isocratically with 5% isopropanol (+0.05% trifluoroacetic acid (TFA)) (solvent B) in 0.05% TFA then to

16% of B in 2.5 min, to 46% of B in 40 min, to 100% of B in 10 min and finally isocratically for 5 min at 40 °C with a flow rate of 0.6 mL min⁻¹ so that the total length of the elution program was 65 min. Mass spectral data was accumulated in Ultra Scan positive electrospray ionization mode (26,000 *m/z* s⁻¹) at scan range of *m/z* 100–1200 and by averaging three spectra.

High resolution UPLC-QTOF (Acquity I-Class UPLC - Synapt G2-Si HDMS, Waters Corp., Milford, MA, USA) analyses of microcystins were performed from the methanol extract of *Phormidium* sp. LP904c (3–7b). One µl sample was injected to Cortecst UPLC® C18+ column (2.1 × 50 mm, 1.6 µm, Waters) which was eluted at 40 °C with a flow rate of 0.3 mL min⁻¹ from 20% acetonitrile (+0.1% HCOOH) (solvent B) in 0.1% HCOOH to 95% of B in 5 min, was kept there for 2 min, then back to 20% of B in 0.5 min and finally kept there for 2.5 min before next run. QTOF was calibrated with sodium formate giving a calibrated mass range of *m/z* 91.036–1178.651. Leucine enkephalin was used at 10 s intervals as a lock mass reference compound. Mass spectral data was accumulated in positive electrospray ionization Resolution Mode at scan range of *m/z* 50–1200.

The chemical characterization of certain microcystin variants was analyzed based on HPLC-ESI-ITMS and UPLC-QTOF results. Microcystin structures derived from mass spectral data was matched with chromatographic retention behavior. Polar surface areas of microcystins were calculated with a topological method [30], which gave TPSA (topological polar surface area) values for the microcystins.

4.3. DNA Extraction, PCR and Genome Sequencing and Assembly

The genomic DNA extraction from strain *Phormidium* sp. DVL1003c was obtained as described in [10]. Fragments of the genes 16S rRNA and *mcyB1CDE* were obtained by PCR and sequenced as described in [10]. Alternatively, the genomic DNA extraction from strain *Phormidium* sp. LP904c was obtained as described in [15,61]. An isolated DNA sample from *Phormidium* sp. LP904c was checked using a NanoDrop 1000 spectrophotometer (Thermo Scientific) to measure the concentration and an Agilent TapeStation (Agilent Technologies) to assess the quality. High-molecular DNA was subjected to library (Illumina TruSeq® PCR Free 350bp, Illumina, San Diego, CA, USA) construction and sequenced by Illumina HiSeq2500 platform (Illumina, San Diego, CA, USA) with a paired ends 100 cycles run. The genome data (1 Gb) of *Phormidium* sp. LP904c was firstly checked by Spades (version 3.7.1)[62] for read correction and removal of erroneous ones, and then assembled using Newbler (version 3.0)[63]. Non-cyanobacterial sequences from contaminants were identified by Kraken 1.0 [64] and removed using a custom script. The accession number of complete sequences of the microcystin gene cluster from *Phormidium* sp. LP904c is MK870090.

4.4. Bioinformatics Analysis

These sequences obtained from *mcyB1C* were translated using BioEdit [65]. The predictions of the binding pockets of McyC and McyB1 adenylation domains was performed using the bioinformatics tool PKS/NRPS analysis [66], alignment of the sequences [47] and NRPSpredictor2 [32,33]. The phylogenetic tree based on 16S rRNA gene sequences was constructed in Molecular Evolutionary Genetic Analysis (MEGA) 5 [67] using the maximum likelihood method (K2+G+I model) and 1000 bootstrap replicates. A concatenated phylogenetic tree constructed using McyD and McyE sequences was obtained using the neighbor-joining method (Poisson model + G) and 1000 bootstrap replications. The same parameters were used to construct the phylogenetic tree using adenylation and condensation domains from McyB1 and McyC.

5. Conclusions

Benthic cyanobacteria belonging to *Phormidium* genus isolated from freshwater environments were shown to produce microcystin containing homo-amino acids. This finding increases the knowledge of microcystin variants produced by *Phormidium*. Furthermore, we describe the microcystin biosynthetic genes from *Phormidium* sp. LP904c, which resembles the ones found in *Planktothrix* strains. Interestingly, *Phormidium* sp. DVL1003c ceased the synthesis of microcystins

during the laboratorial maintenance in our laboratory. Further work is necessary to unveil why two strains kept under the same cultivation conditions had different fates.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1: Figure S1: Relative amounts (%) of microcystin variants in *Phormidium* sp. LP904c (yellow) and DVL1003c (green); Figure S2: Ultraviolet (UV), total ion current (TICC) and extracted ion (EIC) chromatograms obtained with HPLC-ITMS of protonated microcystins found from *Phormidium* sp. LP904c; Figure S3: Ultraviolet (UV), total ion current (TICC) and extracted ion (EIC) chromatograms obtained with HPLC-ITMS of protonated microcystins found from *Phormidium* sp. DVL1003c; Figure S4: Product ion spectra of protonated microcystins from *Phormidium* sp. LP904c obtained with HPLC-ITMS; Figure S5: Coding for the product ions generated from the protonated and sodiated microcystins; Figure S6: Product ion spectra of sodiated non Arg microcystins from *Phormidium* sp. LP904c obtained with HPLC-ITMS. Mass/charge values of protonated microcystins and their retention times (min) are marked to the spectra; Figure S7: Effect of polar surface areas (PSA) to the retention times (R_t , min) of *Phormidium* sp. DVL1003c microcystins calculated with a topological polar surface area (TPSA) method (Ertl et al., 2000); Figure S8: Microcystin variants produced by studied *Phormidium* strains and relative amount produced; Figure S9: Concatenated phylogenetic tree of the McyD and McyE amino acid sequences constructed using neighbor-joining with 1000 bootstrap replications; Figure S10: Phylogenetic tree constructed using condensation (A) and adenylation (B) domains from McyB1 (blue) and McyC (pink) amino acid sequences; Table S1: Assignments, ion masses (m/z) and intensities (%) of the protonated non Arg microcystins of the most important product ions; Table S2: Assignments, ion masses (m/z) and intensities (%) of the sodiated non Arg microcystins of the most important product ions; Table S3: Microcystin variants, retention times (R_t), relative amounts (RA) and small diagnostic ions (from protonated MCs) with corresponding amino acids from *Phormidium* sp. LP904c by HPLC-ITMS and UPLC-QTOF. RAs were calculated from the sum of the peak areas of the extracted ion chromatograms of different ion species (H, Na, K and 2H) of microcystins; Table S4: Sequence similarity of the microcystin gene cluster from *Phormidium* sp. LP904c obtained by BLASTp.

Author Contributions: D.P.F. and K.S. designed the study. T.K.S., D.P.F. and J.J. wrote the article. D.P.F., T.K.S., J.J., M.W., A.H., S.S. performed the experimental work. T.K.S., D.P.F., H.W. and D.O.A. performed bioinformatics analysis. D.P.F. and K.S. contributed for the reagents and equipment.

Funding: The work is supported by grants from the Academy of Finland grant (1273798) to Kaarina Sivonen and (259505) to David P. Fewer, Jane and Aatos Erkkö Foundation to Kaarina Sivonen, São Paulo Research Foundation to Danillo O. Alvarenga (2018/01563-2), and Norforsk grant NordAqua to Kaarina Sivonen and David P. Fewer. Tania Keiko Shishido is supported by Academy of Finland (308266).

Acknowledgments: We would like to thank George Izaguirre for providing the strains for this study and Lyudmila Saari for maintaining the strains.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dittmann, E.; Fewer, D.P.; Neilan, B.A. Cyanobacterial toxins: Biosynthetic routes and evolutionary roots. *FEMS Microbiol. Rev.* **2013**, *37*, 23–43.
2. Sivonen, K. Cyanobacterial toxins. In *Encyclopedia of Microbiology*; Moselio, S., Ed.; Elsevier: Oxford, UK, 2009; pp. 290–307.
3. Merel, S.; Walker, D.; Chicana, R.; Snyder, S.; Baurès, E.; Thomas, O. State of knowledge and concerns on cyanobacterial blooms and cyanotoxins. *Environ. Int.* **2013**, *59*, 303–327.
4. Meriluoto, J.; Spoof, L.; Codd, G. *Handbook of Cyanobacterial Monitoring and Cyanotoxin Analysis*; John Wiley & Sons, Ltd.: Chichester, UK, 2017; ISBN: 978-1-119-06868-6.
5. 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. *Chem. Biol.* **2000**, *7*, 753–764.
6. Christiansen, G.; Fastner, J.; Erhard, M.; Börner, T.; Dittmann, E. Microcystin biosynthesis in *Planktothrix*: Genes, Evolution, and Manipulation. *J. Bacteriol.* **2003**, *185*, 564–572.
7. Rouhiainen, L.; Vakkilainen, T.; Siemer, B.L.; Buikema, W.; Haselkorn, R.; Sivonen, K. Genes coding for hepatotoxic heptapeptides (microcystins) in the cyanobacterium *Anabaena* strain 90. *Appl. Environ. Microbiol.* **2004**, *70*, 686–692.

8. Fewer, D.P.; Wahlsten, M.; Österholm, J.; Jokela, J.; Rouhiainen, L.; Kaasalainen, U.; Rikkinen, J.; Sivonen, K. The genetic basis for *O*-acetylation of the microcystin toxin in cyanobacteria. *Chem. Biol.* **2013**, *20*, 861–869.
9. Welker, M.; von Döhren, H. Cyanobacterial peptides - nature's own combinatorial biosynthesis. *FEMS Microbiol. Rev.* **2006**, *30*, 530–563.
10. Fewer, D.P.; Rouhiainen, L.; Jokela, J.; Wahlsten, M.; Laakso, K.; Wang, H.; Sivonen, K. Recurrent adenylation domain replacement in the microcystin synthetase gene cluster. *BMC Evol. Biol.* **2007**, *7*, 183.
11. Tooming-Klunderud, A.; Fewer, D.P.; Rohrlack, T.; Jokela, J.; Rouhiainen, L.; Sivonen, K.; Kristensen, T.; Jakobsen, K.S. Evidence for positive selection acting on microcystin synthetase adenylation domains in three cyanobacterial genera. *BMC Evol. Biol.* **2008a**, *8*, 256.
12. Alexova, R.; Haynes, P.A.; Ferrari, B.C.; Neilan, B.A. Comparative protein expression in different strains of the bloom-forming cyanobacterium *Microcystis aeruginosa*. *Mol. Cell. Proteomics* **2011**, *10*, M110.003749.
13. Shishido, T.K.; Kaasalainen, U.; Fewer, D.P.; Rouhiainen, L.; Jokela, J.; Wahlsten, M.; Fiore, M.F.; Yunes, J.S.; Rikkinen, J.; Sivonen, K. Convergent evolution of [D-Leucine(1)] microcystin-LR in taxonomically disparate cyanobacteria. *BMC Evol. Biol.* **2013**, *13*, 86.
14. Koketsu, K.; Mitsushashi, S.; Tabata, K. Identification of homophenylalanine biosynthetic genes from the cyanobacterium *Nostoc punctiforme* PCC73102 and application to its microbial production by *Escherichia coli*. *Appl. Environ. Microbiol.* **2013**, *79*, 2201–2208.
15. 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). *Chem. Biol.* **2010**, *17*, 265–273.
16. Shishido, T.K.; Jokela, J.; Fewer, D.P.; Wahlsten, M.; Fiore, M.F.; Sivonen, K. Simultaneous production of anabaenopeptins and namalides by the cyanobacterium *Nostoc* sp. CENA543. *ACS Chem. Biol.* **2017**, *12*, 2746–2755.
17. Komárek, J.; Anagnostidis, K. Cyanoprokaryota 2. Teil: Oscillatoriales. In Süßwasserflora von Mitteleuropa; Büdel, B., Gärtner, G., Krienitz, L., Schagerl, M., Eds.; Elsevier/Spektrum Akademischer Verlag: Heidelberg, 2005; pp. 390–391.
18. Silva, C.S.; Genuário, D.B.; Vaz, M.G.; Fiore, M.F. Phylogeny of culturable cyanobacteria from Brazilian mangroves. *Syst. Appl. Microbiol.* **2014**, *37*, 100–112.
19. Andreote, A.P.; Vaz, M.G.; Genuário, D.B.; Barbiero, L.; Rezende-Filho, A.T.; Fiore, M.F. Nonheterocytous cyanobacteria from Brazilian saline-alkaline lakes. *J. Phycol.* **2014**, *50*, 675–684.
20. Bravakos, P.; Kotoulas, G.; Skaraki, K.; Pantazidou, A.; Economou-Amilli, A. A polyphasic taxonomic approach in isolated strains of Cyanobacteria from thermal springs of Greece. *Mol. Phylogenet. Evol.* **2016**, *98*, 147–160.
21. Singh, Y.; Khattar, J.; Singh, D.P.; Rahi, P.; Gulati, A. Limnology and cyanobacterial diversity of high altitude lakes of Lahaul-Spiti in Himachal Pradesh, India. *J. Biosci.* **2014**, *39*, 643–657.
22. Hawes, I.; Sumner, D.Y.; Andersen, D.T.; Mackey, T.J. Legacies of recent environmental change in the benthic communities of Lake Joyce, a perennially ice-covered Antarctic lake. *Geobiology* **2011**, *9*, 394–410.
23. Mohamed, Z.A.; El-Sharouny, H.M.; Ali, W.S.M. Microcystin production in benthic mats of cyanobacteria in the Nile River and irrigation canals, Egypt. *Toxicon* **2006**, *47*, 584–590.
24. Izaguirre, G.; Jungblut, A.D.; Neilan, B.A. Benthic cyanobacteria (*Oscillatoriaceae*) that produce microcystin-LR, isolated from four reservoirs in southern California. *Water Res.* **2007**, *41*, 492–498.
25. Cantoral Uriza, E.A.; Asencio, A.D.; Aboal, M. Are we underestimating benthic cyanotoxins? Extensive sampling results from Spain. *Toxins (Basel)* **2017**, *9*, doi:10.3390/toxins9120385.
26. Wood, S.A.; Heath, M.W.; Holland, P.T.; Munday, R.; McGregor, G.B.; Ryan, K.G. Identification of a benthic microcystin-producing filamentous cyanobacterium (*Oscillatoriales*) associated with a dog poisoning in New Zealand. *Toxicon* **2010**, *55*, 897–903.
27. Mayumi, T.; Kato, H.; Imanishi, S.; Kawasaki, Y.; Hasegawa, M.; Harada, K. Structural characterization of microcystins by LC/MS/MS under Ion Trap conditions. *J. Antibiot.* **2006**, *59*, 710–719.
28. Yuan, M.; Namikoshi, M.; Otsuki, A.; Sivonen, K. Effect of amino acid side-chain on fragmentation of cyclic peptide ions: Differences of electrospray ionization collision-induced decomposition mass spectra of toxic heptapeptide microcystins containing ADMAdda instead of Adda. *Eur. Mass. Spectrom* **1998**, *4*, 287–298.
29. Diehnelt, C.W.; Dugan, N.R.; Peterman, S.M.; Budde, W.L. 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.* **2006**, *78*, 501–512.

30. Ertl, P.; Rohde, B.; Selzer, P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J. Med. Chem.* **2000**, *43*, 3714–3717.
31. Stachelhaus, T.; Mootz, H.D.; Marahiel, M.A. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* **1999**, *6*, 493–505.
32. Rausch, C.; Weber, T.; Kohlbacher, O.; Wohlleben, W.; Hudson, D.H. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Res.* **2005**, *33*, 5799–5808.
33. Röttig, M.; Medema, M.H.; Bin, K.; Weber, T.; Rausch, C.; Kohlbacher, O. NRPSpredictor2—A web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* **2011**, *39*, W362–W367.
34. Fiore, M.F.; Genuário, D.B.; da Silva, C.S.; Shishido, T.K.; Moraes, L.A.; Cantúcio Neto, R.; Silva-Stenico, M.E. Microcystin production by a freshwater spring cyanobacterium of the genus *Fischerella*. *Toxicon* **2009**, *53*, 754–761.
35. Cirés, S.; Alvarez-Roa, C.; Wood, S.A.; Puddick, J.; Loza, V.; Heimann, K. First report of microcystin-producing *Fischerella* sp. (Stigonematales, Cyanobacteria) in tropical Australia. *Toxicon* **2014**, *88*, 62–66.
36. Gaget, V.; Humpage, A.R.; Huang, Q.; Monis, P.; Brookes, J.D. Benthic cyanobacteria: A source of cylindrospermopsin and microcystin in Australian drinking water reservoirs. *Water Res.* **2017**, *124*, 454–464.
37. Mez, K.; Beattie, K.; Codd, G.; Hanselmann, K.; Hauser, B.; Naegeli, H.; Preisig, H. Identification of a microcystin in benthic cyanobacteria linked to cattle deaths on alpine pastures in Switzerland. *Eur. J. Phycol.* **1997**, *32*, 111–117.
38. Aboal, M.; Puig, M.A. Intracellular and dissolved microcystin in reservoirs of the river Segura basin, Murcia, SE Spain. *Toxicon* **2005**, *45*, 509–518.
39. Hurtado, I.; Aboal, M.; Zafra, E.; Campillo, D. Significance of microcystin production by benthic communities in water treatment systems of arid zones. *Water Res.* **2008**, *42*, 1245–1253.
40. Dasey, M.; Ryan, N.; Wilson, J.; McGregor, G.; Fabbro, L.; Neilan, B.A.; Burns, B.P.; Kankaanpää, H.; Morrison, L.F.; Codd, G.A.; et al. Investigations into the taxonomy, toxicity and ecology of benthic cyanobacterial accumulations in Myall Lake, Australia. *Mar. Freshw. Res.* **2005**, *56*, 45–55.
41. Hitzfeld, B.C.; Lampert, C.S.; Spaeth, N.; Mountfort, D.; Kaspar, H.; Dietrich, D.R. Toxin production in cyanobacterial mats from ponds on the McMurdo Ice Shelf, Antarctica. *Toxicon* **2000**, *38*, 1731–1748.
42. Kleinteich, J.; Wood, S.A.; Puddick, J.; Schleheck, D.; Küpper, F.C.; Dietrich, D. Potent toxins in Arctic environments—presence of saxitoxins and an unusual microcystin variant in Arctic freshwater ecosystems. *Chem. Biol. Interact.* **2013**, *206*, 423–431.
43. Mehner, C.; Müller, D.; Krick, A.; Kehraus, S.; Löser, R.; Gütschow, M.; Maier, A.; Fiebig, H.-H.; Brun, R.; König, G.M. A novel beta-amino acid in cytotoxic peptides from the cyanobacterium *Tychonema* sp. *Eur. J. Org. Chem.* **2008**, *2008*, 1732–1739.
44. An, T.; Kumar, T.K.; Wang, M.; Liu, L.; Lay, J.O., Jr.; Liyanage, R.; Berry, J.; Gantar, M.; Marks, V.; Gawley, R.E.; et al. Structures of pahayokolides A and B, cyclic peptides from a *Lyngbya* sp. *J. Nat. Prod.* **2007**, *70*, 730–735.
45. Namikoshi, M.; Sivonen, K.; Evans, W.R.; Carmichael, W.W.; Rouhiainen, L.; Luukkainen, R.; Rinehart, K.L. Structures of three new homotyrosine-containing microcystins and a new homophenylalanine variant from *Anabaena* sp. strain 66. *Chem. Res. Toxicol.* **1992**, *5*, 661–666.
46. 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*. *Mol. Microbiol.* **2009**, *73*, 924–937.
47. Zainuddin, E.N.; Jansen, R.; Nimtz, M.; Wray, V.; Preisitsch, M.; Lalk, M.; Mundt, S. Lyngbyazothrins A–D, antimicrobial cyclic undecapeptides from the cultured cyanobacterium *Lyngbya* sp. *J. Nat. Prod.* **2009**, *72*, 1373–1378.
48. Leão, P.N.; Pereira, A.R.; Liu, W.T.; Ng, J.; Pevzner, P.A.; Dorrestein, P.C.; König, G.M.; Vasconcelos, V.M.; Gerwick, W.H. Synergistic allelochemicals from a freshwater cyanobacterium. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 11183–11188.
49. Heck, K.; Alvarenga, D.O.; Shishido, T.K.; Varani, A.M.; Dörr, F.A.; Pinto, E.; Rouhiainen, L.; Jokela, J.; Sivonen, K.; Fiore, M.F. Biosynthesis of microcystin hepatotoxins in the cyanobacterial genus *Fischerella*. *Toxicon* **2018**, *141*, 43–50.
50. Conti, E.; Stachelhaus, T.; Marahiel, M.A.; Brick, P. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.* **1997**, *16*, 4174–4183.

51. Challis, G.L.; Ravel, J.; Townsend, C.A. Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem. Biol.* **2000**, *7*, 211–224.
52. Mikalsen, B.; Boison, G.; Skulberg, O.M.; Fastner, J.; Davies, W.; Gabrielsen, T.M.; Rudi, K.; Jacobsen, K.S. Natural variation in the microcystin synthetase operon *mcyABC* and impact on microcystin production in *Microcystis* strains. *J. Bacteriol.* **2003**, *185*, 2774–2785.
53. Tanabe, Y.; Kaya, K.T.; Watanabe, M.M. Evidence for recombination in the microcystin synthetase (*mcy*) genes of toxic cyanobacteria *Microcystis* spp. *J. Mol. Evol.* **2004**, *58*, 633–641.
54. Kurmayer, R.; Christiansen, G.; Gumpenberger, M.; Fastner, J. Genetic identification of microcystin ecotypes in toxic cyanobacteria of the genus *Planktothrix*. *Microbiology* **2005**, *151*, 1525–1533.
55. Kurmayer, R.; Gumpenberger, M. Diversity of microcystin genotypes among populations of the filamentous cyanobacteria *Planktothrix rubescens* and *Planktothrix agardhii*. *Mol. Ecol.* **2006**, *15*, 3849–3861.
56. Tooming-Klunderud, A.; Mikalsen, B.; Kristensen, T.; Jakobsen, K.S. The mosaic structure of the *mcyABC* operon in *Microcystis*. *Microbiology* **2008b**, *154*, 1886–1899.
57. Christiansen, G.; Yoshida, W.Y.; Blom, J.F.; Portmann, C.; Gademann, K.; Hemscheidt, T.; Kurmayer, R. Isolation and structure determination of two microcystins and sequence comparison of the *McyABC* adenylation domains in *Planktothrix* species. *J. Nat. Prod.* **2008**, *71*, 1881–1886.
58. Hoffmann, D.; Hevel, J.M.; Moore, R.E.; Moore, B.S. Sequence analysis and biochemical characterization of the nostopeptolide A biosynthetic gene cluster from *Nostoc* sp. GSV224. *Gene* **2003**, *311*, 171–180.
59. Rounge, T.B.; Rohrlack, T.; Kristensen, T.; Jakobsen, K.S. Recombination and selectional forces in cyanopeptolin NRPS operons from highly similar, but geographically remote *Planktothrix* strains. *BMC Microbiol.* **2008**, *8*, 141.
60. Meyer, S.; Kehr, J.C.; Mainz, A.; Dehm, D.; Petras, D.; Süssmuth, R.D.; Dittmann, E. Biochemical dissection of the natural diversification of microcystin provides lessons for synthetic biology of NRPS. *Cell. Chem. Biol.* **2016**, *23*, 462–471.
61. Wang, H.; Sivonen, K.; Rouhiainen, L.; Fewer, D.P.; Lyra, C.; Rantala-Ylinen, A.; Vestola, J.; Jokela, J.; Rantasärkkä, K.; Li, Z.; et al. Genome-derived insights into the biology of the hepatotoxic bloom-forming cyanobacterium *Anabaena* sp. strain 90. *BMC Genomics* **2012**, *13*, 613.
62. Nurk, S.; Bankevich, A.; Antipov, D.; Gurevich, A.A.; Korobeynikov, A.; Lapidus, A.; Pribelski, A.D.; Pyshkin, A.; Sirotkin, A.; Sirotkin, Y.; Stepanauskas, R.; Clingenpeel, S.R.; Woyke, T.; McLean, J.S.; Lasken, R.; Tesler, G.; Alekseyev, M.A.; Pevzner, P.A. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J. Comput. Biol.* **2013**, *20*, 714–737.
63. Margulies, M.; Egholm, M.; Altman, W.E.; Attiya, S.; Bader, J.S.; Bemben, L.A.; Berka, J.; Braverman, M.S.; Chen, Y.J.; Chen, Z.; Dewell, S.B.; Du, L.; Fierro, J.M.; Gomes, X.V.; Godwin, B.C.; He, W.; Helgesen, S.; Ho, C.H.; Irzyk, G.P.; Jando, S.C.; Alenquer, M.L.; Jarvie, T.P.; Jirage, K.B.; Kim, J.B.; Knight, J.R.; Lanza, J.R.; Leamon, J.H.; Lefkowitz, S.M.; Lei, M.; Li, J.; Lohman, K.L.; Lu, H.; Makhijani, V.B.; McDade, K.E.; McKenna, M.P.; Myers, E.W.; Nickerson, E.; Nobile, J.R.; Plant, R.; Puc, B.P.; Ronan, M.T.; Roth, G.T.; Sarkis, G.J.; Simons, J.F.; Simpson, J.W.; Srinivasan, M.; Tartaro, K.R.; Tomasz, A.; Vogt, K.A.; Volkmer, G.A.; Wang, S.H.; Wang, Y.; Weiner, M.P.; Yu, P.; Begley, R.F.; Rothberg, J.M. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **2005**, *437*, 376–380.
64. Wood, D.E.; Salzberg, S.L. Kraken: Ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* **2014**, *15*, R46.
65. Hall, T.A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* **1999**, *41*, 95–98.
66. Bachmann, B.O.; Ravel, J. *In silico* prediction of microbial secondary metabolic pathways from DNA sequence data. *Methods Enzymol.* **2009**, *458*, 181–217.
67. Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739.

