2013

Convergent evolution of [D-Leucine1] microcystin-LR in taxonomically disparate cyanobacteria

Convergent evolution of [*D*-Leucine[1]] microcystin-LR in taxonomically disparate cyanobacteria

Shishido *et al.*

BMC
Evolutionary Biology

RESEARCH ARTICLE   **Open Access**

Convergent evolution of [*D*-Leucine¹] microcystin-LR in taxonomically disparate cyanobacteria

Tânia Keiko Shishido[1], Ulla Kaasalainen[2], David P Fewer[1], Leo Rouhiainen[1], Jouni Jokela[1], Matti Wahlsten[1], Marli Fátima Fiore[3], João Sarkis Yunes[4], Jouko Rikkinen[2] and Kaarina Sivonen[1*]

Abstract

Background: Many important toxins and antibiotics are produced by non-ribosomal biosynthetic pathways. Microcystins are a chemically diverse family of potent peptide toxins and the end-products of a hybrid NRPS and PKS secondary metabolic pathway. They are produced by a variety of cyanobacteria and are responsible for the poisoning of humans as well as the deaths of wild and domestic animals around the world. The chemical diversity of the microcystin family is attributed to a number of genetic events that have resulted in the diversification of the pathway for microcystin assembly.

Results: Here, we show that independent evolutionary events affecting the substrate specificity of the microcystin biosynthetic pathway have resulted in convergence on a rare [*D*-Leu¹] microcystin-LR chemical variant. We detected this rare microcystin variant from strains of the distantly related genera *Microcystis*, *Nostoc*, and *Phormidium*. Phylogenetic analysis performed using sequences of the catalytic domains within the *mcy* gene cluster demonstrated a clear recombination pattern in the adenylation domain phylogenetic tree. We found evidence for conversion of the gene encoding the McyA$_2$ adenylation domain in strains of the genera *Nostoc* and *Phormidium*. However, point mutations affecting the substrate-binding sequence motifs of the McyA$_2$ adenylation domain were associated with the change in substrate specificity in two strains of *Microcystis*. In addition to the main [*D*-Leu¹] microcystin-LR variant, these two strains produced a new microcystin that was identified as [Met¹] microcystin-LR.

Conclusions: Phylogenetic analysis demonstrated that both point mutations and gene conversion result in functional *mcy* gene clusters that produce the same rare [*D*-Leu¹] variant of microcystin in strains of the genera *Microcystis*, *Nostoc*, and *Phormidium*. Engineering pathways to produce recombinant non-ribosomal peptides could provide new natural products or increase the activity of known compounds. Our results suggest that the replacement of entire adenylation domains could be a more successful strategy to obtain higher specificity in the modification of the non-ribosomal peptides than point mutations.

Keywords: Adenylation domain, Phylogeny, Substrate specificity, Gene conversion, Point mutations, Recombination

Background

Non-ribosomal peptides are an important class of secondary metabolites produced by a range of bacteria and fungi [1-3]. These peptides have many biotechnological and pharmaceutical applications such as the antibiotics penicillin [4] and daptomycin [5], and the anticancer bleomycin [6]. They are synthesized on large modular non-ribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) enzyme complexes. Each NRPS module is responsible for the recognition and incorporation of an amino acid during elongation of the peptide intermediate. The basic NRPS module has a condensation (C), adenylation (A), and peptidyl carrier protein (PCP) domains. The adenylation domain is responsible for the selection and activation of amino acids in the form of aminoacyl adenylates [7]. It is followed by the peptidyl carrier protein and condensation domains, the former holding the activated amino acid and the latter making a peptide bond between two adjacent amino acids [2]. In addition, auxiliary enzymes may be present and have activities such as the epimerization, cyclisation, *N*-methylation,

* Correspondence: kaarina.sivonen@helsinki.fi
[1]Department of Food and Environmental Sciences, Division of Microbiology and Biotechnology, University of Helsinki, Viikki Biocenter (Viikinkaari 9) P.O. Box 56, Helsinki FIN-00014, Finland
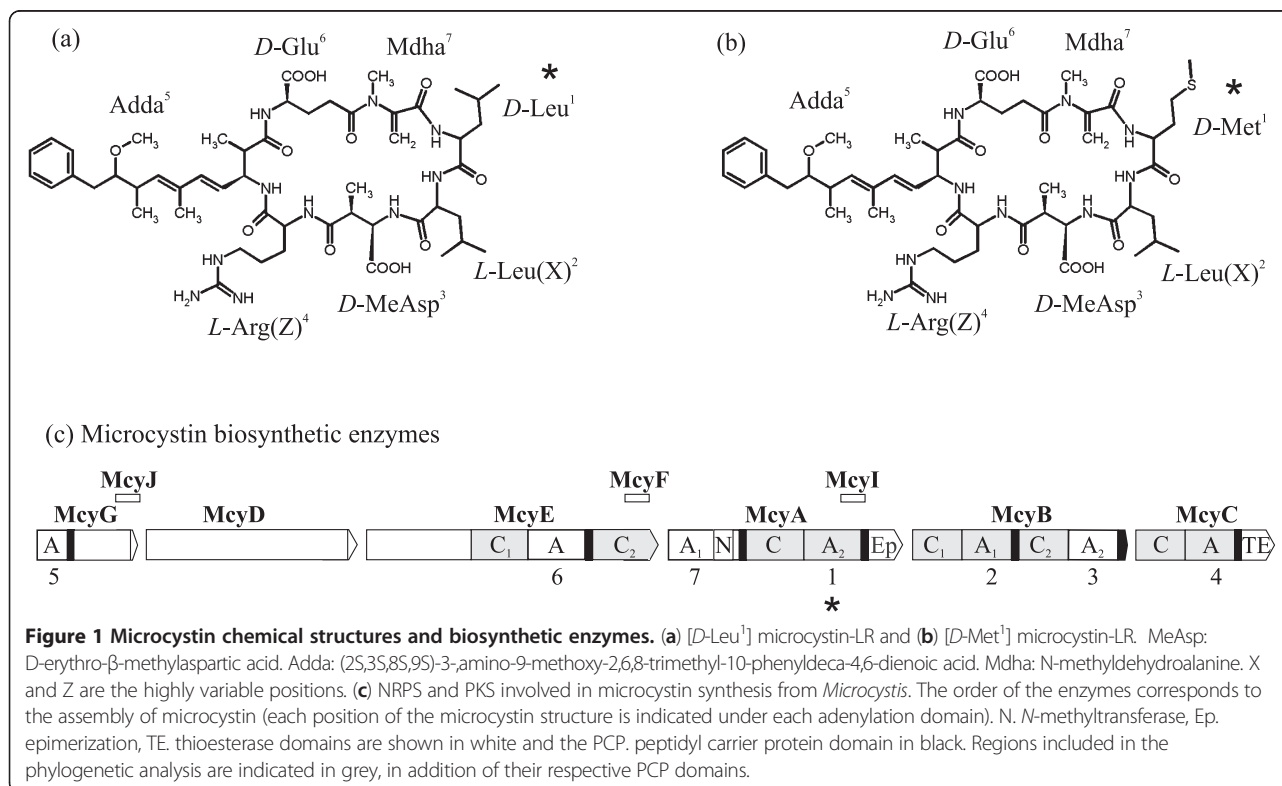Full list of author information is available at the end of the article

formylation, and reduction of amino acids [1,3]. The large variety of non-proteinogenic amino acids and hydroxyl acids that can be incorporated and further modified by tailoring enzymes allows the production of highly complex peptides.

Microcystins are the most frequently reported cyanobacterial toxins in aquatic blooms. They are small cyclic heptapeptides with extensive variation in amino acid residue composition and are commonly produced by planktonic strains in fresh and brackish water [8]. Microcystins are potent inhibitors of serine/threonine protein phosphatases (PP1 and PP2A) and can cause human intoxication, tumor promotion, and death [9]. The general structure of microcystin is cyclo(−D-Ala[1]-X[2]-D-MeAsp[3]-Z[4]-Adda[5]-D-Glu[6]-Mdha[7]-) (Figure 1a and b). A hybrid NRPS-PKS enzyme complex directs the synthesis of microcystin and is encoded in the 55–55.6 kb (*mcy*) gene cluster (Figure 1c) [10-13].

Phylogenetic studies indicate that the *mcy* gene cluster has an ancient origin among cyanobacteria [14-17]. Other studies suggest that horizontal transfer, gene loss, and recombination events in the microcystin gene cluster explain the distribution and variation of the genes among the closely related *Microcystis* spp. [18-20]. Recombination events affecting the adenylation domain have been described in genes encoding McyA$_1$, McyB$_1$ and McyC [16,18,19,21-25]. Positive selection acting on the adenylation domains of McyB$_1$ and McyC was reported as

the possible cause of the large number of microcystin variants produced by cyanobacteria [24]. Deletion of the entire *N*-methyltransferase domain of *mcyA* in *Anabaena* or point mutations in this gene in *Microcystis* were associated with the absence of *N*-methylation in the microcystins produced by these strains [23,26]. Furthermore, recombination in the same region (*mcyA$_1$*) was related to the synthesis of microcystins containing 2-amino-2-butenoic acid (Dhb) in some strains of the genus *Planktothrix* [25]. Such genetic rearrangements, positive selection, and recombination events act to increase the chemical variability of microcystins found in nature. However, the production of the same rare microcystin variant in disparate taxonomic lineages of cyanobacteria raises questions about the genetic mechanism underlying this phenomenon [27-31]. Evolutionary diversification of NRPS and PKS pathways is achieved through genetic mechanisms such as recombination, duplication, fusion or fission of genes, deletion or substitutions of domains, circular permutations, gene loss and horizontal gene transfer [32,33]. Here, we show that the production of a rare [D-Leu[1]] microcystin (MC) variant in three distantly related genera of cyanobacteria is the result of three independent evolutionary events leading to convergence on the same chemical structure. There is an interest in engineering non-ribosomal peptide biosynthetic pathways in order to increase the production levels of known compounds or create new bioactive compounds [3]. Our results suggest that the replacement

Figure 1 Microcystin chemical structures and biosynthetic enzymes. (a) [D-Leu[1]] microcystin-LR and **(b)** [D-Met[1]] microcystin-LR. MeAsp: D-erythro-β-methylaspartic acid. Adda: (2S,3S,8S,9S)-3-,amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid. Mdha: N-methyldehydroalanine. X and Z are the highly variable positions. **(c)** NRPS and PKS involved in microcystin synthesis from *Microcystis*. The order of the enzymes corresponds to the assembly of microcystin (each position of the microcystin structure is indicated under each adenylation domain). N. *N*-methyltransferase, Ep. epimerization, TE. thioesterase domains are shown in white and the PCP. peptidyl carrier protein domain in black. Regions included in the phylogenetic analysis are indicated in grey, in addition of their respective PCP domains.

of entire adenylation domains might be a more successful strategy for changing substrate specificity in the engineering of NRPS than point mutations.

Results

Production of the [D-Leu¹]MC variant by taxonomically disparate cyanobacteria

We documented the production of the rare [D-Leu¹] MC variant in morphologically disparate cyanobacteria (Figure 2). These cyanobacteria were assigned to the genera *Nostoc*, *Phormidium* and *Microcystis*. *Nostoc* sp. strain UK89IIa was identified based on morphology and 16S rRNA gene sequence similarity (Figures 2 and 3). The *Phormidium* sp. CENA270 strain has a thick sheath, tangled filaments and small (≤6.3 μm) cells that indicate the presence of similar characters with *Phormidium*, even though the cells in the trichomes are distinctly wider (5–6.3 μm) than their length (1–3 μm). The 16S rRNA gene sequence of strain CENA270 was 98% of *Phormidium* sp. DVL1003c, which also produces microcystin. These two strains form a clade with *Phormidium*, *Lyngbya*, *Oscillatoria* and Oscillatoriales strains in the 16S rRNA trees (Figure 3; Additional file 1: Figure S1). The two *Microcystis* strains had previously been identified [27,34].

Nostoc sp. UK89IIa produced four variants of microcystins, while *Phormidium* sp. CENA270 produced five variants (Table 1). All of the detected microcystin variants

produced by these two strains contained D-Leu, and none of the strains produced detectable levels of microcystins that contained D-Ala. The two *Microcystis* strains produced at least twelve variants altogether (Table 1, Additional file 1: Table S2 and Additional file 2: Table S1). However, 97% of the microcystin variants in *Microcystis aeruginosa* NPLJ-4 and 80% of the microcystin produced by *Microcystis* sp. RST 9501 contained D-Leu at position 1 (Table 1, Additional file 2: Table S1). We carried out further chemical analysis to characterize the new microcystin variants produced by these strains.
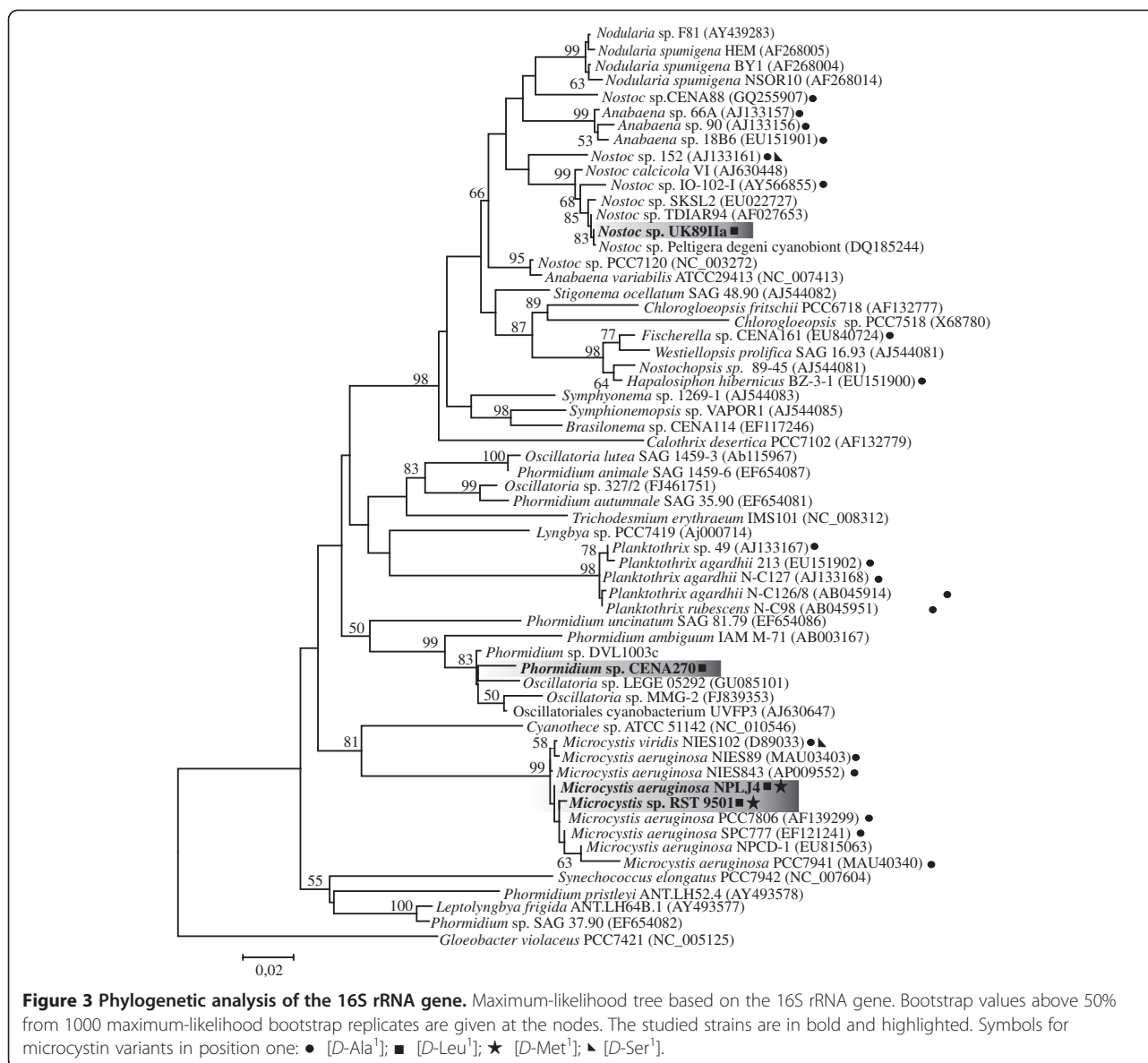
In all known microcystin variants, the amino acid configuration is *D* in position 1 and *L* in positions 2 and 4 (Figure 1a and b). Therefore, the chirality of the amino acids in positions 1, 2 and 4 of the major microcystin variants produced by the studied strain was investigated by using deuterated acid hydrolysis reagents (DCl, D_2O). This confirmed that the hydrolyzed microcystins contained D-Leu (Additional file 1: Table S2).

Production of [Met¹]MC-LR by *Microcystis* spp.

Mass spectrometry strongly suggested that *Microcystis* strains NPLJ-4 and RST 9501 produced new microcystin variants that contained Met instead of D-Leu (Table 1, Figure 1b). In order to confirm these results, *Microcystis* sp. RST 9501 cells were grown with ³²S and ³⁴S as the sole source of sulfur. LC-MS revealed an increase in [Met¹]
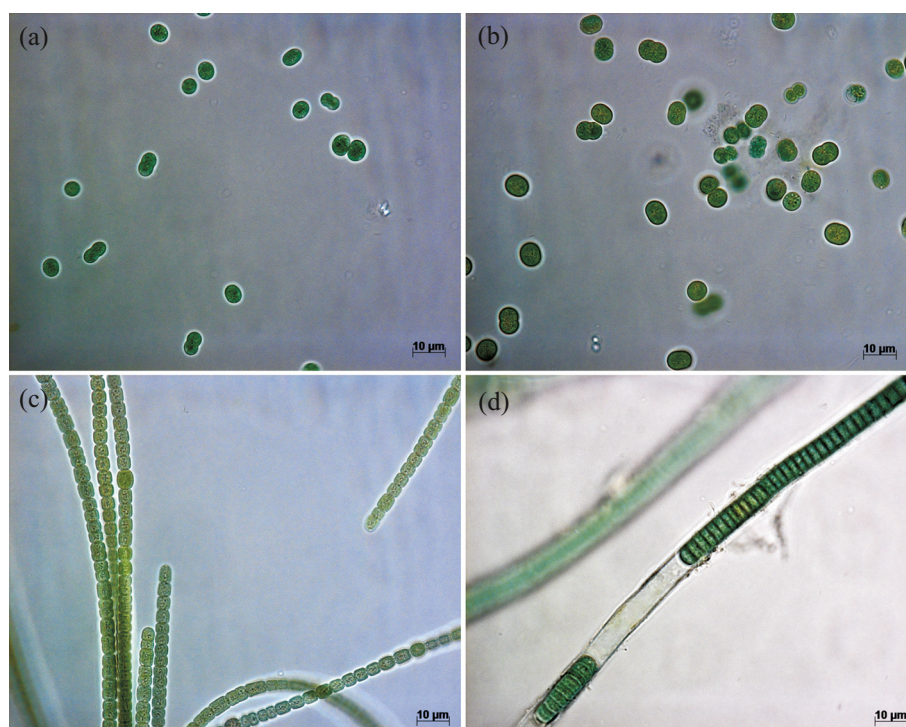
Figure 2 Photomicrographs of the studied strains. (**a**) *Microcystis aeruginosa* NPLJ-4. (**b**) *Microcystis* sp. RST 9501. (**c**) *Nostoc* sp. UK89IIa. (**d**) *Phormidium* sp. CENA270.

Figure 3 Phylogenetic analysis of the 16S rRNA gene. Maximum-likelihood tree based on the 16S rRNA gene. Bootstrap values above 50% from 1000 maximum-likelihood bootstrap replicates are given at the nodes. The studied strains are in bold and highlighted. Symbols for microcystin variants in position one: ● [*D*-Ala[1]]; ■ [*D*-Leu[1]]; ★ [*D*-Met[1]]; ◣ [*D*-Ser[1]].

MC-LR protonated ion mass from m/z 1,055.5 (control containing [32]S) to 1,057.5 ([34]S-labelled samples), indicating the presence of a sulfur atom in the microcystin, and the same differences could be visualized in the fragmentation analysis (Additional file 1: Figure S2). The product ion spectra of [Met[1]]MC-LR could be compared with the spectra of [*D*-Leu[1]]MC-LR and -RR (Additional file 1: Figures S2a and S3a and b). The results confirmed the presence of a new variant of [Met[1]]MC in the two *Microcystis* strains NPLJ-4 and RST 9501. However, *Microcystis* sp. RST 9501 produced 14 times more [Met[1]]MC-LR than strain NPLJ-4. On the other hand, *Microcystis aeruginosa* NPLJ-4 produced a larger diversity of microcystin variants, but most of them in trace amounts, such as [Val[1]]MC-LR and [Phe[1]]MC-LR.

Phylogenetic analysis of microcystin catalytic domains

In order to understand the order and timing of genetic events leading to the production of the rare [*D*-Leu[1]]MC variant, we conducted phylogenetic analysis of the *mcy* gene cluster. Phylogenetic trees based on concatenated *mcyD* and *mcyE* gene sequences and the housekeeping 16S rRNA gene were robust and found to have a similar pattern (Additional file 1: Figure S4). Strains producing the [*D*-Leu[1]]MC variant do not group together but instead group with strains that produce the [*D*-Ala[1]]MC variant in both trees (Additional file 1: Figure S4). We constructed alignments based on the NRPS catalytic domains encoded in each of the microcystin biosynthetic genes (Figures 1c, 4 and 5). The condensation domains and peptidyl carrier proteins of each module encoded in

Table 1 Chemical variants of microcystin detected in LC-MS/MS analyses

Microcystin variant	[M+H]$^+$ (m/z)	*Microcystis aeruginosa* NPLJ-4	*Microcystis* sp. RST9501	*Nostoc* sp. UK89IIa	*Phormidium* sp. CENA270
[Leu[1]]MC-LR	1037	76	76	96	31
[Leu[1], Asp[3]]MC-LR	1023	19	<1	2	18
[Met[1]]MC-LR	1055	1	18	-	-
[Met[1], Asp[3]]MC-LR	1041	2	-	-	-
[Leu[1], Dha[7]]MC-LR	1023	1	-	1	-
[Leu[1]]MC-HilR	1051	-	3	-	-
[Leu[1]]MC-HphR	1085	-	-	2	-
[Leu[1]]MC-LHar	1051	-	-	-	7
[Leu[1]]MC-RR	1080	-	-	-	31
[Leu[1], Asp[3]]MC-RR	1066	-	-	-	13

Values given as the percentage of relative area of the peaks in the protonated molecular ion chromatograms for the main microcystins variants produced by the studied strain. Hil, homoisoleucine. Hph, homophenylalanine. Har, homoarginine.

the *mcy* gene cluster grouped together according to their encoding gene and were placed in separate clades in the phylogenetic tree (Figure 4a and b). A similar pattern was also observed for the adenylation domains of McyG, McyE, and McyB$_2$ (Figure 5), and the epimerization domains of McyA (Figure 4c). However, the adenylation domain sequences of McyA$_1$, McyA$_2$, McyB$_1$, and McyC are mixed and do not form separate clades (Figure 5). A phylogenetic tree was constructed using adenylation domain sequences of McyA$_2$ and other NRPSs obtained from BLASTp searches of the *nr* database at NCBI (Figure 6). The McyA$_2$ adenylation domains of the *Microcystis* strains were grouped in a single well-supported clade, irrespective of whether they produced [*D*-Leu[1]]MCs or [*D*-Ala[1]]MCs variants. By contrast, the McyA$_2$ adenylation domain of *Phormidium* sp. CENA270 and *Nostoc* sp. UK89IIa did not group together with other McyA$_2$ adenylation domains. Instead, the *Phormidium* sp. CENA270 McyA$_2$ adenylation domain was placed in the same clade of McyB$_1$ adenylation domain of *Microcystis* strains. The McyA$_2$ adenylation domain of *Nostoc* sp. UK89IIa grouped with adenylation domain sequences from *Nostoc punctiforme* PCC73102 and *Nostoc* sp. GSV224.

### Substrate specificity of the McyA$_2$ adenylation domain

Conservation of the tertiary structure of adenylation domains makes it possible to predict amino acid binding pockets and consequently the substrate specificity. All strains that produce the [*D*-Ala[1]]MC variants analyzed in this study have identical predicted McyA$_2$ adenylation domain binding pocket sequences, with the exception of *Planktothrix agardhii* NIVA-CYA 126/8 (Table 2). There are only two conserved residues (D235 and K517) in the predicted binding pocket in strains producing [*D*-Ala[1]] or [*D*-Leu[1]]MC variants. However, *Nostoc* sp. UK89IIa and *Phormidium* sp. CENA270 have identical predicted binding pockets, differing substantially from those of strains producing

[*D*-Ala[1]]MC variants (Table 2). *Microcystis aeruginosa* NPLJ-4 and *Microcystis* sp. RST 9501 have identical predicted binding pockets, which differ by three amino acids at positions 301, 330, and 331 in comparison to strains that activate *L*-Ala (Table 2). Strains producing [*D*-Ala[1]]MC variants have polar amino acids (Thr or Ser) at position 330 of the binding pocket, while strains producing [*D*-Leu[1]]MC variants have hydrophobic non-polar amino acids (Ile and Val) at this position (Table 2).

### Genetic variations in the McyA$_2$ adenylation domain

Microcystis spp. NPLJ-4 and RST 9501, *Nostoc* sp. UK89IIa and *Phormidium* sp. CENA270 produce the same rare [*D*-Leu[1]]MC variant [27,31]. Recombination events affecting the substrate specificity of the *mcyA$_2$* gene were detected in *Nostoc* sp. UK89IIa and *Phormidium* sp. CENA270 by four different methods (Table 3). Breakpoints with statistical support within the *mcyA* gene were identified in UK89IIa and CENA270 (Figure 7a). The identified predicted breakpoints were visualized in a recombination breakpoint distribution plot (Figure 7b). A lengthy region replaced in a recombination event in the *mcyA* gene in *Nostoc* sp. UK89IIa (1029 bp) and *Phormidium* sp. CENA270 (167 and 707 bp) was observed in the present study (Figure 7a). The predicted binding pockets responsible for amino acid selection and activation were found within this region in both cases.

No evidence for recombination was detected in the *mcyA$_2$* gene from *Microcystis aeruginosa* NPLJ-4 or *Microcystis* sp. RST 9501. However, the McyA$_2$ adenylation domain of *Microcystis* spp. NPLJ-4 and RST 9501 differed by 17 amino acids residues compared to other McyA$_2$ adenylation domains of *Microcystis* strains that produce microcystins containing *D*-Ala (Additional file 1: Figure S5). Five of the 17 amino acids residues were predicted to be located within 8 Å of the substrate, and three of them were considered to be within the binding pocket. Mutants based on *Microcystis* sp. RST 9501 McyA$_2$ adenylation
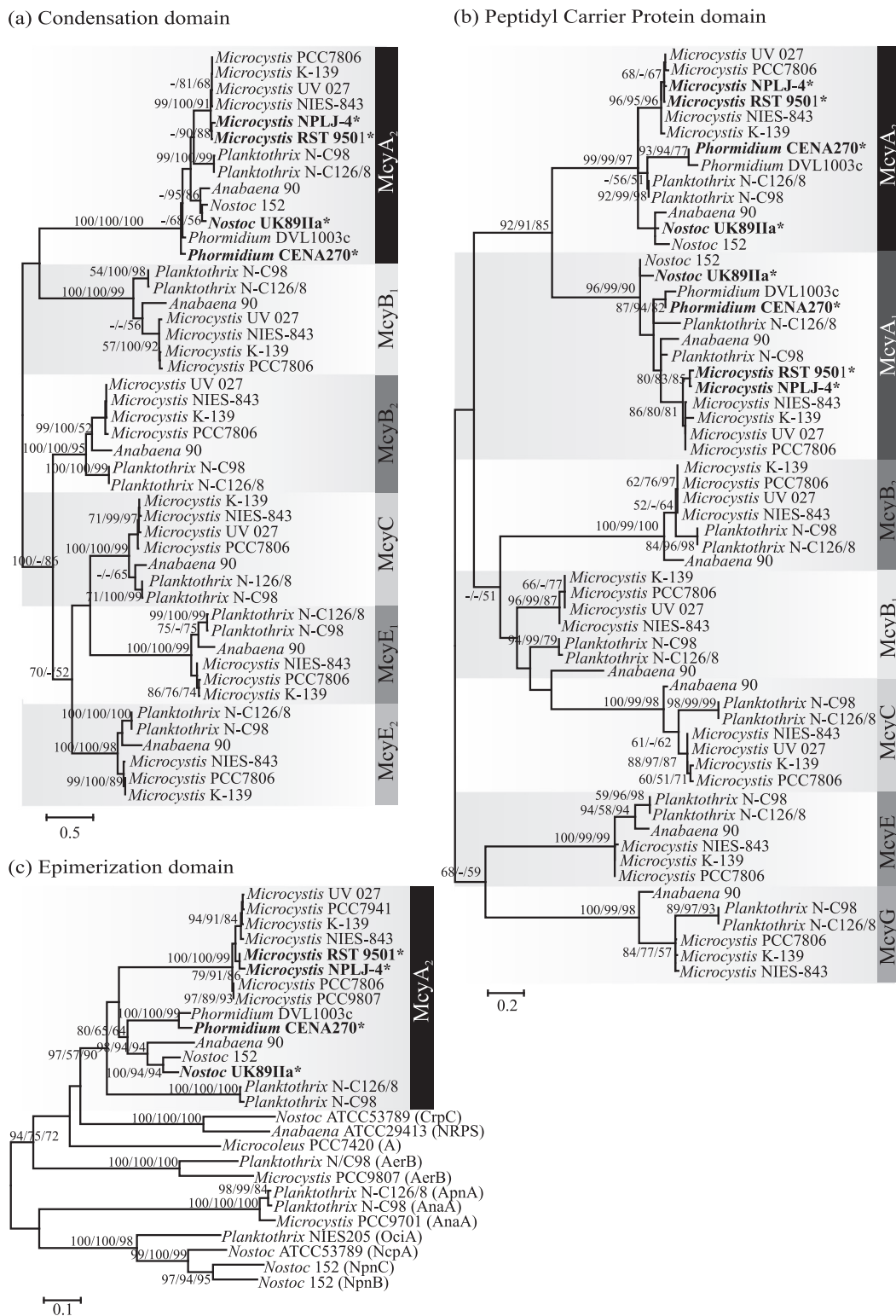
Figure 4 Conservative evolutionary history of domains surrounding the McyA₂ adenylation domain in the microcystin biosynthetic gene cluster. Maximum-likelihood tree based on amino acid sequences of the (**a**) condensation domain, (**b**) peptidyl carrier protein domain and (**c**) epimerization domain within the *mcy* gene cluster. Phylogenetic tree inferred using MEGA 5. Bootstrap values above 50 per cent from 1000 respectively neighbor-joining, maximum parsimony and maximum-likelihood bootstrap replicates are given at the nodes. The studied strains are in bold and indicated with *.
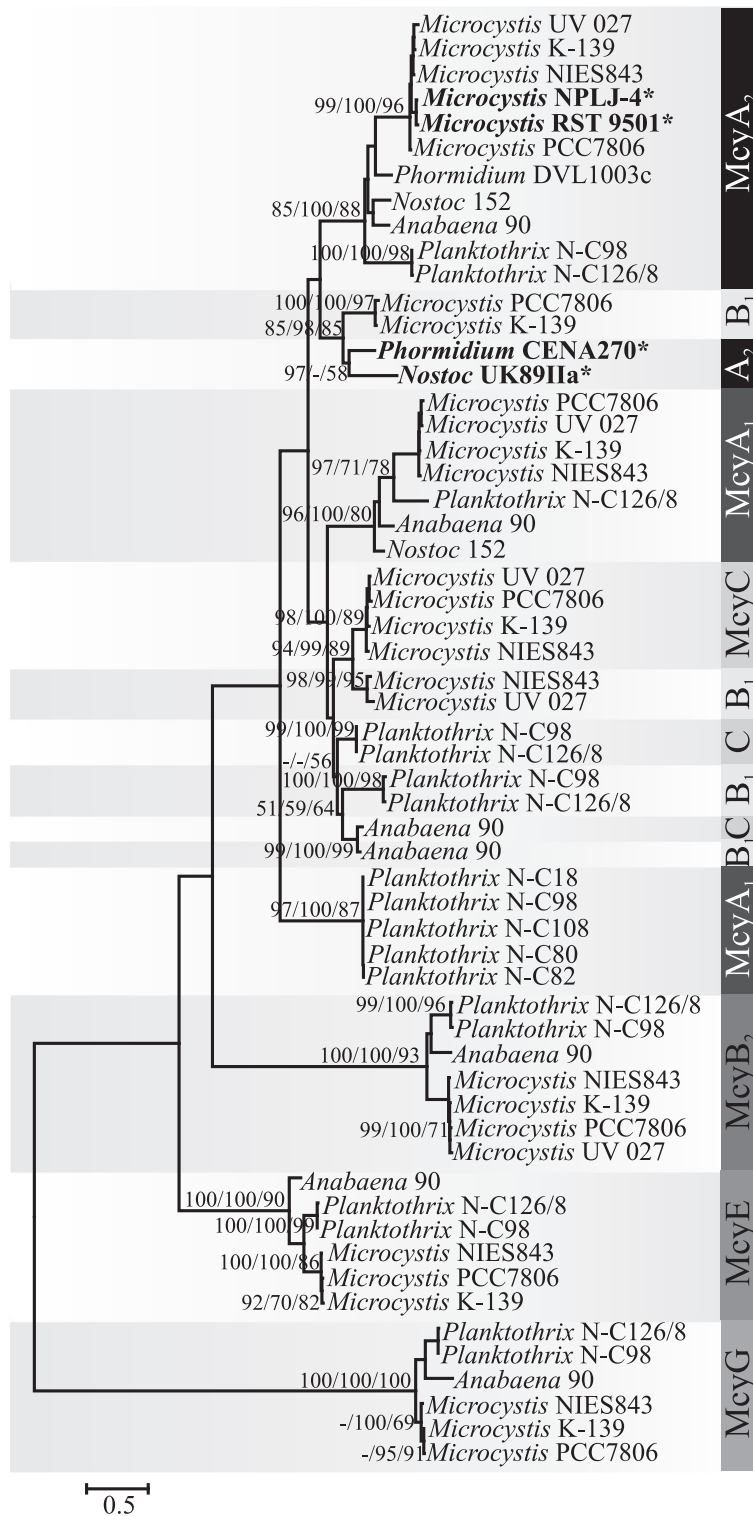
Figure 5 Evolutionary history of adenylation domains of microcystin biosynthetic gene cluster. Maximum-likelihood tree based on amino acids sequences of adenylation domain within *mcy* gene cluster. Phylogenetic tree inferred using MEGA 5. Bootstrap values above 50 per cent from 1000 respectively neighbor-joining, maximum parsimony and maximum-likelihood bootstrap replicates are given at the nodes. The studied strains are in bold and indicated with *.
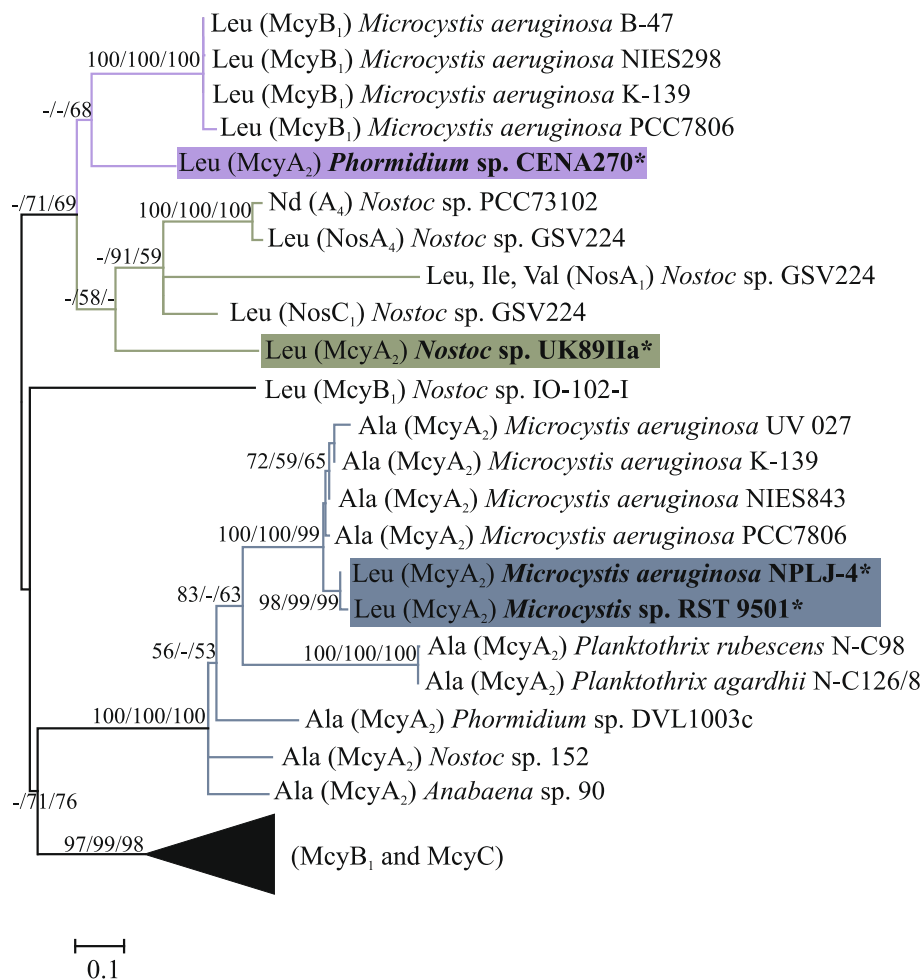
Figure 6 Independent evolutionary history of McyA₂ adenylation domain. Maximum-likelihood tree based on amino acids sequences inferred using MEGA 5. Bootstrap values above 50 per cent from 1000 respectively neighbor-joining, maximum parsimony and maximum-likelihood bootstrap replicates are given at the nodes. The studied strains are in bold and indicated with *.

domain were constructed to analyze the effect of these three amino acids residues on binding of the predicted substrate amino acids. Surprisingly, the results obtained in the ATP-pyrophosphate exchange assay showed the highest specificity for *L*-Val in all the mutants and the wild type (Additional file 1: Figure S6). Furthermore, the results indicated a higher activation of *L*-Ile instead of the expected *L*-Leu and *L*-Ala. The high levels of miscognate activation indicate that the catalytic efficiency of the enzymes in recognizing nonpolar amino acid was generally higher, although some specificity was observed for *L*-Tyr (Additional file 1: Figure S6).

Discussion

Identification of the rare [*D*-Leu¹]MC and other microcystin variants

Microcystins form a large family of cyclic toxins characterized by a highly conserved chemical structure with an extensive amino acid composition at the two variable

positions, two and four (Additional file 1: Figure S7 and Table S3). In this study, we detected rare [*D*-Leu¹]MC variants from strains of the distantly related genera *Microcystis*, *Nostoc*, and *Phormidium*. Almost all microcystins reported in the literature contain *D*-Ala in position 1 [8]. Previously, microcystins containing *D*-Leu [27-31], Gly [35], and *D*-Ser [36,37] have also been reported. The [*D*-Leu¹]MC variant has previously been found from *Microcystis aeruginosa* NPLJ-4, *Microcystis* sp. RST 9501, and water blooms from Brazil and Canada dominated by *Microcystis* strains [27-31,34]. The production of microcystins in lichen thalli by *Nostoc* symbionts has previously been reported [37,38]. *Nostoc* strains isolated from lichen symbiosis produce a large variety of microcystins including the [*D*-Leu¹] MC variant [31].

Microcystins are best known from aquatic habitats, where they are frequently reported from blooms. Although microcystins are more commonly detected in

Table 2 Predicted binding pockets of the adenylation domain of McyA$_2$

Strain	Binding pocket										Score %(a)	Predicted (a)	Activated
	235	236	239	278	299	301	322	330	331	517			
Anabaena sp. 90	D	L	F	N	N	A	L	T	Y	K	100	Ala	Ala
Microcystis aeruginosa NIES-843	-	-	-	-	-	-	-	-	-	-	100	Ala	Ala
Microcystis aeruginosa PCC7806	-	-	-	-	-	-	-	-	-	-	100	Ala	Ala
Microcystis aeruginosa UV027	-	-	-	-	-	-	-	-	-	-	100	Ala	Ala
Microcystis aeruginosa K-139	-	-	-	-	-	-	-	-	-	-	100	Ala	Ala
Nostoc sp. 152	-	-	-	-	-	-	-	-	-	-	100	Ala	Ala*
Phormidium sp. DVL1003c	-	-	-	-	-	-	-	-	-	-	100	Ala	Ala
Planktothrix rubescens NIVA-CYA 98	-	-	-	-	-	-	-	-	-	-	100	Ala	Ala
Planktothrix agardhii NIVA-CYA 126/8	-	-	-	-	-	-	-	S	-	-	90	Ala	Ala
Microcystis sp. RST 9501	-	-	-	-	-	G	-	I	C	-	70	Cys	Leu, Met
Microcystis aeruginosa NPLJ-4	-	-	-	-	-	G	-	I	C	-	70	Cys	Leu, Met ±
Nostoc sp. UK89lla	-	A	W	F	L	G	N	V	V	-	100	Leu	Leu
Phormidium sp. CENA270	-	A	W	F	L	G	N	V	V	-	100	Leu	Leu

The adenylation domain of McyA$_2$ is responsible for amino acid activation in position one of microcystins.
(a) Prediction obtained using the NRPSpredictor2 program [72,73]. Trace amounts of *Serine and ± Valine or Phenylalanine detected in LC-MS/MS.

planktonic strains, terrestrial and benthic strains have also been reported to be producers of these compounds [8,39]. Benthic environmental samples containing microcystins have been reported from Switzerland [40], Spain [41], Australia [42], and Antarctica [43]. Microcystin production in isolated cyanobacterial strains from benthic environments has been reported from Egypt [44], New Zealand [45], and the USA [46]. It is not always clear which cyanobacterium produces the toxin in benthic mats of cyanobacteria. A strain of the genus *Phormidium* was isolated from the walls of a reservoir in the USA and shown to produce a range of microcystin variants, all of which contained *D*-Ala (Additional file 1: Table S3) [46]. Our results demonstrated that *Phormidium* sp. CENA 270, isolated from a pond in the northeast of Brazil, also produces microcystins but with *D*-Leu in place of *D*-Ala. In the phylogenetic analysis of the 16S rRNA gene, the two *Phormidium* strains cluster together with *Lyngbya*, *Oscillatoria*, *Phormidium* and Oscillatoriales (Additional file 1: Figure S1). The biomass of benthic strains can go unnoticed to the casual observer but be massive enough to cause animal poisonings [40,45]. Moreover, when the strains lyse, microcystins are released into the water, which

suggests that the analysis of toxic benthic cyanobacteria is also important in water-quality management.

The chemical structure of microcystins is highly conserved, with variation at X and Z positions (Figure 1a and b) resulting in over 86 reported variants [39]. In this study, we demonstrated that *Microcystis* strains NPLJ-4 and RST 9501 produce new microcystin variants containing methionine in addition to the rare [*D*-Leu^1]MC variants. Met is also present in oscillamide B [47] and in microcystin-M (O)R and -YM [48]. *Nodularia spumigena* strains also produce nodulapeptins, which commonly contain Met [49,50]. According to analysis using the NRPS Norine database, Met contains a methylthiol group and is rare in non-ribosomal peptides [51]. The presence of the highly active sulfhydryl group in the thiol group could explain the scarcity of secondary metabolites containing Met or Cys. If amino acid recognition by the McyA$_2$ adenylation domain is not strict, the incorporation of Met instead of Leu is logical because of the similar size and hydrophobicity of the side chains. Here, we demonstrated that methionine is incorporated in the microcystins produced by *Microcystis* strains from brackish water. Microcystin variants are constantly being discovered, making the

Table 3 Recombination events affecting the substrate specificity of the adenylation domain of the *mcyA*$_2$ gene

Strain	Parents*		Average P-value						
	Major	Minor	RDP	GENECON	BootScan	MaxChi	Chimaera	SiScan	3Seq
Nostoc sp. UK89lla	N152	M9501	1.423x10^{-9}	5.882x10^{-9}	2.306x10^{-3}	5.938x10^{-12}	7.331x10^{-3}	7.789x10^{-16}	2.192x10^{-24}
Phormidium sp. CENA270	±P1003c	±Ana90	1.242x10^{-6}	1.793x10^{-18}	3.361x10^{-6}	6.848x10^{-14}	1.200x10^{-3}	7.894x10^{-27}	2.410x10^{-105}
	¦P1003c	¦M027	4.209x10^{-3}	-	-	2.274x10^{-5}	9.115x10^{-4}	6.746x10^{-12}	-

Analysis was performed in the RDP3 program and breakpoint positions are indicated in Figure 7a. The P-value cutoff was chosen as 0.05 and the best signals for recombination are associated with the lowest P-values.
* M7806 = *Microcystis aeruginosa* PCC7806; M139 = *Microcystis aeruginosa* K-139; N152 = *Nostoc* sp. 152; M9501 = *Microcystis* sp. RST 9501; P1003e = *Phormidium* sp. DVL1003c; Ana90 = *Anabaena* sp. 90; M027 = *Microcystis aeruginosa* UV027. Breakpoint positions: ± 432 (undetermined) to 1139 and ¦ 216 to 383 (undetermined).
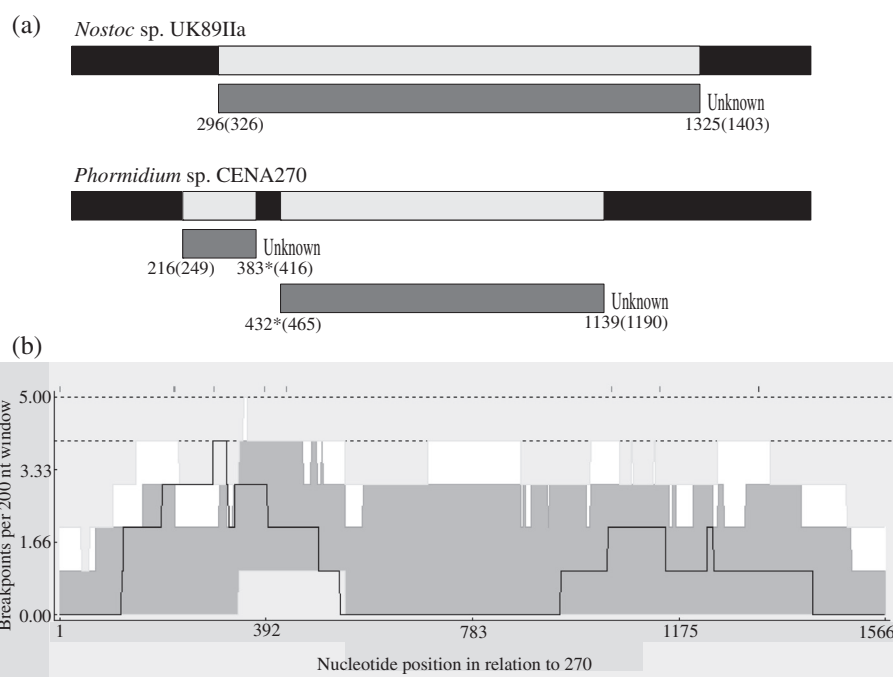
Figure 7 Recombination breakpoints in the *mcyA*$_2$ gene detected by the RDP3 program. (a) Breakpoint positions detected in three cyanobacterial strains. The position in the alignment is in parentheses. *The breakpoint position is undetermined. (**b**) Potential recombination hotspots within the alignment of *mcyA*$_2$ genes. The black line corresponds to the breakpoint count within a moving 200-nt-long window. White and grey areas respectively indicate local 99% and 95% confidence intervals of the hot/cold spot test. Upper and lower broken lines respectively indicate 99% and 95% confidence thresholds for the global hotspot test.

microcystin family extremely diverse, and posing a challenge for the detection of microcystins from water samples.

Convergence on [*D*-Leu^1]MC variant chemical structure

Phylogenetic analysis of the McyA$_2$ adenylation domains provided evidence for independent evolutionary events affecting the substrate specificity of the enzyme in three disparate genera of cyanobacteria (Figures 5 and 6). Breakpoint analysis suggests the replacement of almost the entire substrate specificity-conferring portion of the adenylation domain in *Phormidium* sp. CENA270 and *Nostoc* sp. UK89IIa. These gene conversions dramatically altered the predicted substrate specificity of the McyA$_2$ adenylation domain in these strains and are linked to the synthesis of the [*D*-Leu^1]MC variant. However, point mutations affecting the substrate specificity of the McyA$_2$ adenylation domain in *Microcystis* strains NPLJ-4 and RST 9501 led to the synthesis of the [*D*-Leu^1]MC and [Met^1]MC variants.

Phormidium sp. CENA270 and *Nostoc* sp. UK89IIa are not grouped together with other McyA$_2$ adenylation domains (Figure 6). The McyA$_2$ adenylation domain of *Phormidium* sp. CENA270 grouped with the McyB$_1$ adenylation domain from *Microcystis* strains, which produce microcystin variants containing *D*-Ala [10,52]. The adenylation domains of NosA$_1$ and NosC$_1$ from the

nostopeptolide gene cluster are placed in the same clade with the McyA$_2$ adenylation domain from *Nostoc* sp. UK89IIa (Figure 6). They are involved in the incorporation of Ile/Leu/Val and Leu, respectively, in *Nostoc* sp. GSV224 [53].

Genetic variation in the microcystin synthetases can be visualized in the phylogenetic trees showing two different patterns. While the amino acids of condensation, peptidyl carrier protein, and epimerization domain regions can be grouped according to the enzyme sequence (McyA, McyB, McyC, McyE, and McyG grouped together), the adenylation domain phylogeny clearly indicates recombination (Figures 4 and 5). Recombination in adenylation domains has previously been described for the adenylation and condensation domains of McyB$_1$ and McyC [16]. The recombination and positive selection in the McyB$_1$ adenylation domain are involved in the high variability of amino acids incorporated at position 2 of the microcystin [16,18,22,24,25]. These genetic events have been related to the increase in the chemical diversity of microcystin. Interestingly, our results show that these different evolutionary events are involved in the convergence of the [*D*-Leu^1]MC-LR.

Nevertheless, the selective forces behind this convergent evolution remain unclear. Competition in brackish water

and different seasonal periods have possibly acted as selective forces. The chemical diversity of microcystins could be related to protein phosphatase inhibition as a form of chemical defense, for example against predators. Previous studies have indicated that microcystins can affect some predators, acting as metal chelators, in gene regulation, or in the inter- and intra-specific signaling [54-59]. However, microcystins join a large number of secondary metabolites produced by different organisms that have no assigned biological function. According to the most accepted view, these compounds are produced due their ecological or physiological function and benefits for the producer organisms [60]. However, more information is still needed concerning the advantages in the production of these secondary metabolites. The biological role of a mixture of different bioactive compounds produced by the same strain would be interesting to study.

Prediction of McyA$_2$ adenylation domain substrate specificity

The eight to ten amino acid residues forming the adenylation domain binding pocket are the main determinants of substrate specificity [7,61,62]. In our study, *Phormidium* sp. CENA270 and *Nostoc* sp. UK89IIa were shown to produce [*D*-Leu[1]]MC variants and have identical binding pocket sequences (Table 2). Such amino acids signatures had been already described as presenting Leu specificity [7]. The *Microcystis* strains NPLJ-4 and RST 9501 differ in the binding pocket positions 301, 330, and 331 from the strains producing [*D*-Ala[1]]MCs. Residues at positions 301 and 330 are regarded to be less variable than at position 331 [7].

The adenylation domain binding pocket of *Microcystis* strains NPLJ-4 and RST 9501 has three different amino acid residues and a broader diversity of microcystin variants at position 1. Despite the fact that almost the entire binding pocket of *Phormidium* sp. CENA270 and *Nostoc* sp. UK89IIa differs from the other studied strains, only microcystin variants containing Leu at position 1 were detected. Re-engineering of non-ribosomal peptides has been a challenge in order to synthesize new peptides or to increase the activity of known compounds. The engineering of NRPSs to change substrate specificity can in some cases be achieved by point mutations. However, our results suggest that the replacement of entire domains might be a more successful strategy for producing a single product.

Replacement of almost the entire McyA$_2$ adenylation domain in *Nostoc* sp. UK89IIa and *Phormidium* sp. CENA270 resulted in specificity towards Leu. Neither strain produced detectable levels of microcystin variants that contain other amino acids at this position. The recombination detected in the *mcyA$_2$* gene of these strains affects the substrate-conferring portion of the McyA$_2$

adenylation domain, which is important for the selection and activation of amino acids [7,61,62]. Previously, it has been reported that recombination among different adenylation domains from *mcyB1* and *mcyC* genes has led to a change in amino acid activation [24].

We designed an experiment in order to test whether point mutations at positions 301, 330, and 331 could change the substrate specificity of the adenylation domain. However, single amino acids changes did not have the expected results. All the constructs and the wild type were found to activate valine in ATP-pyrophosphate (PPi) exchange assays. A previous study [7] demonstrated that in the case of single or multiple mutations, the specificity of the wild type is not lost, but there is an increase in new substrate specificity. A comparison of adenylation domains from *Microcystis* strains that activate Ala and Leu reveals that several amino acid residues differ between them (Additional file 1: Figure S5). Of these different amino acid residues, five are 8 Å or less distant from the substrate and only three belong to the binding pocket. Although it is predicted that amino acid residues in the binding pocket are involved in selectivity, the catalytic efficiency could also be affected by the tertiary structure and proteinogenic surrounding area of the adenylation domain [63]. Promiscuity of the enzymes, allowing them to activate different substrates, could also be involved in the high variability of microcystin variants. Promiscuous activation of amino acids with a hydrophobic side chain by TycA, involved in the synthesis of the antibiotic tyrocidine A, has been reported [64]. Moreover, adenylation domains activating multiple substrates have been described from the fengycin [65], lychenysin [66], nostopeptolide [53], and cyanopeptolin [67] biosynthetic pathways.

Conclusion

Our study revealed that independent gene conversion events and point mutations led to the production of the same microcystin variant by strains belonging to three different cyanobacterial orders. The large chemical diversity of microcystins is proposed to be mostly the result of genetic rearrangements, positive selection, and recombination acting to increase structural diversity. Furthermore, the replacement of the entire adenylation domain seems to result in a more specific change in non-ribosomal peptides than point mutation. New variants of [Met[1]]MCs were detected in *Microcystis* strains NPLJ-4 and RST 9501. Our study also revealed a new cyanobacterial strain (*Phormidium* sp CENA270) producing a rare variant of the potent hepatotoxic microcystin. This finding expands on the recent increase in the detection of microcystin-producing terrestrial and benthic cyanobacterial strains.

Methods

Studied strains and culture conditions

Phormidium sp. CENA270 was isolated from a pond during the rainy season in Paulista (Paraiba, Brazil) and maintained in BG-11 medium [68], but supplemented with half the described amount of combined nitrogen. *Nostoc* sp. UK89IIa was isolated from the lichen *Peltigera neopolydactyla* sampled in Laukaa (Finland) and maintained in Z8 medium [69] lacking a source of combined nitrogen. *Microcystis* strains NPLJ-4 and RST 9501 were respectively isolated from the brackish water of Jacarepaguá Lagoon, Rio de Janeiro [34], and the Patos Lagoon estuary, Rio Grande do Sul [27], in Brazil and maintained in Z8 medium [69] with and without a source of combined nitrogen. The biomass for chemical and molecular analysis was obtained by growing the strains in 2 x 3 L of culture medium under constant light of 3–8 μmol m^{-2} s^{-1} and a temperature of 24 ± 1°C. *Microcystis* sp. RST 9501 was grown in Z8 medium replaced with stable isotope ^{34}S-labeled MgSO$_4$.7H$_2$O (catalogue no. IS7080; 90 atom % ^{34}S; Icon) for the detection and identification of sulfur-containing microcystins.

Sequencing and phylogenetic analysis

Freeze-dried cells (6–50 mg) were used for DNA extraction with a DNeasy Plant mini kit (Qiagen). Two different sizes of glass beads (180 μm and 425–600 μm, Sigma-Aldrich) were added and the cells were disrupted by shaking at 6.5 m s^{-1} for 60 seconds in 3 batches using a FastPrep (M.P. Biomedicals). An additional 1 h incubation at 100°C was necessary in order to extract *Phormidium* sp. CENA270 genomic DNA due the thick sheath surrounding the trichomes of this strain. PCR reaction conditions are described in Additional file 3. The PCR products were cloned in pCR®2.1-TOPO (TOPO TA Cloning, Invitrogen) with the following modifications from the manufacturer's instructions: the entire cloning reaction was used for transformation in 25 μL of TOP10 competent cells, 100 μL of SOC medium was added to the mixture after heat shock and incubated for 20 min in shaker (160 RPM) at 37°C, and the entire reaction volume was plated in LB plates containing 150 μg mL^{-1} of ampicillin.

Plasmid extraction was performed using the QIAprep Spin Miniprep kit (Qiagen) and cycle sequencing was carried out using an ABI Prism® BigDye® Terminator v3.1 cycle sequencing kit. The oligonucleotide primers used to sequence PCR products are given in Additional file 1: Table S4. Sanger sequencing was performed in an ABI PRISM 310 Genetic Analyzer. Contigs were aligned in the program Phred/Phrap/Consed (Philip Green, University of Washington, Seattle, USA), accepting bases with quality >20. BLASTn was used to search the nr database at NCBI for related strains to be included in

phylogenetic analysis. GenBank accession numbers are indicated in Additional file 1: Table S5. Maximum-likelihood trees were constructed in MEGA 5.0 [70]. The best substitution model for each sequence was chosen based on analysis in MEGA 5.0 (K2 + G for the small tree of 16S rRNA; GTR + G + I for *mcyD* concatenated with *mcyE*; K2 + G + I for long tree of 16S rRNA; JTT + G + I for condensation, adenylation, and epimerization domains sequences of McyA; and JTT + G for peptidyl carrier protein domain sequences of McyA). Neighbor-joining and maximum parsimony trees were constructed using the respective methods: JTT + G and CNI on random trees. The domains present in the McyA amino acid sequences were detected in the program PKS/NRPS Analysis [71]. Adenylation domain substrate specificity prediction was performed using the program NRPSpredictor2 [72,73].

Recombination test

The nucleotide sequences of 13 cyanobacterial strains (Additional file 1: Table S5) corresponding to second adenylation domain of *mcyA* (*mcyA$_2$*) were aligned in ClustalW (MEGA 5.0). The recombination analyses were performed in the program RDP3 [74,75], which implements different methods to detect recombination. The methods used in this study were: original RDP [76], BOOTSCAN [77], GENECONV [78], MAXCHI [79], CHIMAERA [80], SISCAN [81] and 3Seq [82]. Default parameters were used and a P-value cutoff was chosen as 0.05. Nucleotide sequence comparison was performed in BLASTn (NCBI database). A breakpoint distribution was plotted, allowing the visualization of recombination hotspots [74]. The phylogenetic compatibility matrix was calculated by TreeOrder scan in the program SSE [83].

McyA$_2$ adenylation domain mutation and heterologous expression

A fragment containing the *mcyA$_2$* adenylation domain was obtained from *Microcystis* sp. RST 9501 by PCR using the primers RSTPETF and RSTPETR (Additional file 1: Table S6). The PCR reaction conditions are described in Additional file 3. The 1,646 bp DNA fragment was cloned in pET101/D-TOPO (Invitrogen) and transformed in TOP10 chemically competent *Escherichia coli* cells. Mutants were obtained by site-directed mutation with PCR mutagenesis using specific primers (Additional file 1: Table S6). Detailed information on mutagenesis is presented in Additional file 3. Mutant G301A contains an Ala instead of Gly at position 301, mutant I330T a Thr instead of Ile at position 330, mutant C331Y a Tyr instead of Cys at position 331, and mutant G301A, I330T, C331Y has all three mutations. The presence of the desired mutation was verified by sequencing. Mutated and wild type adenylation domains were cloned into pFN18A

(HaloTag® 7) T7 Flexi® vector (Promega, WI, USA) and transformed in *Escherichia coli* KRX competent cells (Promega) following the manufacturer's instructions. The adenylation domain was heterologously expressed and purified using the HaloTag® Protein Purification System (Promega). The ATP-pyrophosphate exchange assay was performed as previously described [49].

Chemical analysis

Identification of microcystin variants was performed using an Agilent 1100 Series LC/MSD Trap XCT Plus high-performance liquid chromatograph mass spectrometer (Agilent Technologies). Freeze-dried biomass was extracted with 1 mL of methanol in a FastPrep homogenizer (M.P. Biomedicals). The supernatant was injected to a Luna C18 (2) column (150 × 2.1 mm, 5 µm, Phenomenex) for the LC-MS analysis. Electrospray ionization in positive mode was used and the product ion spectra of protonated microcystins were analyzed to identify the structure of the variant.

The remains of the methanol extracts were mixed with water and dichloromethane in equivalent proportions (1:1:1). The hydrophilic upper phase was diluted with water and passed through a preconditioned SPE column (strata 8B-S100-UBJ, Phenomenex). Microcystin-containing fractions were recovered in 1 mL of methanol. Microcystins were isolated and purified by HPLC (HP 1100 Series modular chromatograph, Agilent Technologies). The microcystin fractions were hydrolyzed with 200 µL of 6 M deuterium chloride (catalog no. 543047; 35 wt % solution in D_2O, 99 atom % D, Sigma-Aldrich) and prepared for enantiomeric amino acid analysis using FDAA (Pierce) as a Marfey reagent, as described previously [84]. Chiral analysis of the amino acids alanine, leucine, arginine, and homoarginine was carried out using LC-MS. Amino acids from the common [*D*-Ala[1]] MC-LR and -RR variants produced by *Anabaena* sp. 90 were analyzed as references. Detailed information on the chemical analysis is provided in Additional file 3.

Additional files

Additional file 1: Table S2. Chiral analysis of microcystin amino acids. Table S3: Amino acids incorporated by each adenylation domain of microcystin biosynthetic enzymes for the studied strains. Table S4: Primers used in this study. Table S5: Access number of strains compared in this study. Table S6. Primers designed for PCR mutagenesis. Figure S1: Phylogenetic analysis of 16S rRNA gene focusing in *Phormidium* sp. CENA270. Figure S2: Product ion spectra of [Met[1]] MC-LR of *Microcystis* sp. RST 9501 in the labeling experiment. Figure S3: Product ion spectra of protonated [Leu[1]]microcystins of *Phormidium* sp. CENA270. Figure S4. Phylogenetic congruence between housekeeping and microcystin synthetase genes. Figure S5: Comparison of McyA$_2$ adenylation domain sequences from *Microcystis* strains. Figure S6: ATP-PPi exchange assay. Figure S7: Relative quantity of amino acids in detected microcystins for strains included in the phylogenetic tree.

Additional file 2: Table S1. Assignment of the main ions from microcystin variants produced by the studied strains in LC-MS/MS.

Additional file 3: Supplementary information in Methods. Additional file 3 is a PDF file that contains further information to complement the Methods sub-sections: Sequencing and phylogenetic analysis; McyA$_2$ adenylation domain mutation and heterologous expression; and Chemical analysis.

Abbreviations

NRPS: Non-ribosomal peptide synthetase; PKS: Polyketide synthase; NRPS-PKS: Hybrid non-ribosomal peptide synthetase and polyketide synthase; A: Adenylation domain; C: Condensation domain; PCP: Peptidyl carrier protein domains; E: Epimerization domain; MC: Microcystin.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed to writing of the manuscript and interpretation of the data. TKS, DPF, JJ and MW participated in the design of the experiments. TKS and DPF carried out the molecular experiments and analysis. TKS and LR carried out the biochemical assay. TKS, JJ, UK and MW carried out the chemical experiments and analysis. MFF, JSY and LR revised the manuscript. KS made the study financially feasible and was involved in the design, drafting and revision of the manuscript. All authors read and approved the final manuscript.
Cover page photo: *Peltigera neopolydactyla* (Photo: Jouko Rikkinen); Patos Lagoon estuary, Rio Grande do Sul (Photo: João Sarkis Yunes); and Paulista, Paraiba, Brazil (Photo: João Luiz da Silva).

Author details

[1]Department of Food and Environmental Sciences, Division of Microbiology and Biotechnology, University of Helsinki, Viikki Biocenter (Viikinkaari 9) P.O. Box 56, Helsinki FIN-00014, Finland. [2]Department of Biosciences, University of Helsinki, Viikki Biocenter (Viikinkaari 1), P.O. Box 65, Helsinki FIN-00014, Finland. [3]Center for Nuclear Energy in Agriculture, University of São Paulo, Avenida Centenário 303, Piracicaba, SP 13400-970, Brazil. [4]Unidade de Pesquisas em Cianobactérias, Federal University of Rio Grande, Av. Itália km 8 - Caixa Postal 474, Rio Grande, RS 96.201-900, Brazil.

References

1. Walsh CT, Chen H, Keating TA, Hubbard BK, Losey HC, Luo L, Marshall CG, Miller DA, Patel HM: **Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines.** *Curr Opin Chem Biol* 2001, **5**:525–534.
2. Finking R, Marahiel MA: **Biosynthesis of nonribosomal peptides.** *Annu Rev Microbiol* 2004, **58**:453–488.
3. Sieber SA, Marahiel MA: **Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics.** *Chem Rev* 2005, **105**:715–738.
4. Kallow W, Neuhof T, Arezi B, Jungblut P, Von Döhren H: **Penicillin biosynthesis: intermediates of biosynthesis of delta-alpha-aminoadipyl-cysteinyl-valine formed by ACV synthetase from *Acremonium chrysogenum*.** *FEBS Lett* 1997, **414**:74–78.
5. Steenbergen J, Alder J, Thorne G, Tally F: **Daptomycin: a lipopeptide antibiotic for the treatment of serious gram-positive infections.** *J Antimicrob Chemother* 2005, **55**:283–288.
6. Shen B, Du L, Sanchez C, Edwards D, Chen M, Murrell J: **The biosynthetic gene cluster for the anticancer drug bleomycin from *Streptomyces***

verticillus ATCC 15003 as a model for hybrid peptide-polyketide natural product biosynthesis. *J Ind Microbiol Biotechnol* 2001, **27**:378–385.

7. Stachelhaus T, Mootz HD, Marahiel MA: **The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases.** *Chem Biol* 1999, **6**:493–505.

8. Sivonen K, Jones G: **Cyanobacterial toxins.** In *Toxic cyanobacteria in water.* Edited by Chorus I, Bartram J. London: E & FN Spon; 1999:41–111.

9. Kuiper-Goodman T, Falconer IR, Fitzgerald DJ: **Human health aspects.** In *Toxic Cyanobacteria in Water: a Guide to their Public Health Consequences, Monitoring and Management.* Edited by Chorus I, Bartram J. London: E and FN Spon; 1999:113–153.

10. Tillett D, Dittmann E, Erhard M, Von Döhren H, Börner T, Neilan BA: **Structural organization of microcystin biosynthesis in** *Microcystis aeruginosa* **PCC7806: an integrated peptide polyketide synthetase system.** *Chem Biol* 2000, **7**:753–764.

11. Christiansen G, Fastner J, Erhard M, Börner T, Dittmann E: **Microcystin biosynthesis in** *Planktothrix*: **Genes, Evolution, and Manipulation.** *J Bacteriol* 2003, **185**:564–572.

12. Rouhiainen L, Vakkilainen T, Siemer BL, 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.

13. Dittmann E, Börner T: **Genetic contributions to the risk assessment of microcystin in the environment.** *Toxicol Appl Pharmacol* 2005, **203**:192–200.

14. Rantala A, Fewer DP, Hisbergues M, Rouhiainen L, Vaitomaa J, Borner T, Sivonen K: **Phylogenetic evidence for the early evolution of microcystin synthesis.** *Proc Natl Acad Sci USA* 2004, **101**:568–573.

15. Jungblut AD, Neilan BA: **Molecular identification and evolution of the cyclic peptide hepatotoxins, microcystin and nodularin, synthetase genes in three orders of cyanobacteria.** *Arch Microbiol* 2006, **185**:107–114.

16. Fewer DP, 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.

17. Christiansen G, Molitor C, Philmus B, Kurmayer R: **Nontoxic strains of cyanobacteria are the result of major gene deletion events induced by a transposable element.** *Mol Biol Evol* 2008, **25**:1695–1704.

18. Mikalsen B, Boison G, Skulberg OM, Fastner J, Davies W, Gabrielsen TM, Rudi K, Jacobsen KS: **Natural variation in the microcystin synthetase operon** *mcyABC* **and impact on microcystin production in** *Microcystis* **strains.** *J Bacteriol* 2003, **185**:2774–2785.

19. Tanabe Y, Kaya KT, Watanabe MM: **Evidence for recombination in the microcystin synthetase (***mcy***) genes of toxic cyanobacteria** *Microcystis* **spp.** *J Mol Evol* 2004, **58**:633–641.

20. Tanabe Y, Sano T, Kasai F, Watanabe MM: **Recombination, cryptic clades and neutral molecular divergence of the microcystin synthetase (***mcy***) genes of toxic cyanobacterium** *Microcystis aeruginosa*. *BMC Evol Biol* 2009, **9**:115.

21. 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.

22. 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.

23. Tooming-Klunderud A, Mikalsen B, Kristensen T, Jakobsen KS: **The mosaic structure of the** *mcyABC* **operon in** *Microcystis*. *Microbiology* 2008, **154**:1886–1899.

24. Tooming-Klunderud A, Fewer DP, Rohrlack T, Jokela J, Rouhiainen L, Sivonen K, Kristensen T, Jakobsen KS: **Evidence for positive selection acting on microcystin synthetase adenylation domains in three cyanobacterial genera.** *BMC Evol Biol* 2008, **8**:256.

25. Christiansen G, Yoshida WY, Blom JF, 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.

26. Fewer DP, Tooming-Klunderud A, Jokela J, Wahlsten M, Rouhiainen L, Kristensen T, Rohrlack T, Jakobsen KS, Sivonen K: **Natural occurrence of microcystin synthetase deletion mutants capable of producing microcystins in strains of the genus** *Anabaena* **(Cyanobacteria).** *Microbiology* 2008, **154**:1007–1014.

27. Matthiensen A, Beattie KA, Yunes JS, Kaya K, Codd GA: **[D-Leu¹]Microcystin-LR, from the cyanobacterium** *Microcystis* **RST 9501 and from a** *Microcystis* **bloom in the Patos Lagoon estuary, Brazil.** *Phytochemistry* 2000, **55**:383–387.

28. Park H, Namikoshi M, Brittain SM, Carmichael WW, Murphy T: **[D-Leu(1)] microcystin-LR, a new microcystin isolated from waterbloom in a Canadian prairie lake.** *Toxicon* 2001, **39**:855–862.

29. Schripsema J, Dagnino D: **Spectral assignments and reference data complete assignment of the NMR spectra of [D-Leu¹]-microcystin-LR and analysis of its solution structure.** *Magn Reson Chem* 2002, **40**:614–617.

30. Albuquerque EC Jr, Melo LFC, Franco TT: **Use of solid-phase extraction, high-performance liquid chromatography, and MALDI-TOF identification for [D-Leu¹] MCYST-LR analysis in treated water: validation of the analytical methodology.** *Can J Anal Sci Spectrosc.* 2007, **52**:1–10.

31. Kaasalainen U, Fewer DP, Jokela J, Wahlsten M, Sivonen K, Rikkinen J: **Cyanobacteria produce a high variety of hepatotoxic peptides in symbiosis.** *Proc Natl Acad Sci USA* 2012, **109**:5886–5891.

32. Amoutzias GD, Van de Peer Y, Mossialos D: **Evolution and taxonomic distribution of nonribosomal peptide and polyketide synthases.** *Future Microbiol* 2008, **3**:361–370.

33. Dittmann E, Fewer DP, Neilan BA: **Cyanobacterial toxins: biosynthetic routes and evolutionary roots.** *FEMS Microbiol Rev* 2013, **37**:23–43.

34. Soares RM, Magalhaes VF, Azevedo SMFO: **Accumulation and depuration of microcystins (cyanobacteria hepatotoxins) in** *Tilapia rendalli* **(Cichlidae) under laboratory conditions.** *Aquat Toxicol* 2004, **70**:1–10.

35. Wood SA, Mountfort D, Selwood AI, Holland PT, Puddick J, Cary SC: **Widespread distribution and identification of eight novel microcystins in Antarctic cyanobacterial mats.** *Appl Environ Microbiol* 2008, **74**:7243–7251.

36. Sivonen K, Namikoshi M, Evans WR, Fardig M, Carmichael WW, Rinehart KL: **Three new microcystins, cyclic heptapeptide hepatotoxins, from Nostoc sp. strain 152.** *Chem Res Toxicol* 1992, **5**:464–469.

37. Oksanen I, Jokela J, Fewer DP, Wahlsten M, Rikkinen J, Sivonen K: **Discovery of rare and highly toxic microcystins from lichen-associated cyanobacterium Nostoc sp. strain IO-102-I.** *Appl Environ Microbiol* 2004, **70**:5756–5763.

38. Kaasalainen U, Jokela J, Fewer DP, Sivonen K, Rikkinen J: **Microcystin production in the tripartite cyanolichen** *Peltigera leucophlebia*. *Mol Plant Microbe Interact* 2009, **22**:695–702.

39. Sivonen K: **Cyanobacterial Toxins.** In *Encyclopedia of Microbiology.* Edited by Moselio S. Oxford: Elsevier; 2009:290–307.

40. 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.

41. Aboal M, Puig MA: **Intracellular and dissolved microcystin in reservoirs of the river Segura basin, Murcia, SE Spain.** *Toxicon* 2005, **45**:509–518.

42. Dasey M, Ryan N, Wilson J, McGregor G, Fabbro L, Neilan BA, Burns BP, Kankaanpää H, Morrison LF, Codd GA, Rissik D, Bowling L: **Investigations into the taxonomy, toxicity and ecology of benthic cyanobacterial accumulations in Myall Lake, Australia.** *Mar Freshw Res* 2005, **56**:45–55.

43. Hitzfeld BC, Lampert CS, Spaeth N, Mountfort D, Kaspar H, Dietrich DR: **Toxin production in cyanobacterial mats from ponds on the McMurdo Ice Shelf, Antarctica.** *Toxicon* 2000, **38**:1731–1748.

44. Mohamed ZA, El-Sharouny HM, Ali WSM: **Microcystin production in benthic mats of cyanobacteria in the Nile River and irrigation canals. Egypt.** *Toxicon* 2006, **47**:584–590.

45. Wood SA, Heath MW, Holland PT, Munday R, McGregor GB, Ryan KG: **Identification of a benthic microcystin-producing filamentous cyanobacterium (Oscillatoriales) associated with a dog poisoning in New Zealand.** *Toxicon* 2010, **55**:897–903.

46. Izaguirre G, Jungblut AD, Neilan BA: **Benthic cyanobacteria (Oscillatoriaceae) that produce microcystin-LR, isolated from four reservoirs in southern California.** *Water Res* 2007, **41**:492–498.

47. Sano T, Usui K, Ueda H, Osada H, Kaya K: **Isolation of new protein phosphatase inhibitors from two cyanobacteria species.** *Planktothrix spp. J Nat Prod.* 2001, **64**:1052–1055.

48. Namikoshi M, Rinehart KL, Sakai R, Stotts RR, Dahlem AM, Beasley VR, Carmichael WW, Evans WR: **Identification of 12 hepatotoxins from a Homer Lake bloom of the cyanobacteria** *Microcystis aeruginosa, Microcystis viridis*, and *Microcystis wesenbergii*: **nine new microcystins** *J Org Chem* 1992, **57**:866–872.

49. Rouhiainen L, Jokela J, Fewer DP, 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.

50. Fujii K, Sivonen K, Adachi K, Noguchi K, Sano H, Hirayama K, Suzuki M, Harada K-I: Comparative study of toxic and non-toxic cyanobacterial products: novel peptides from toxic *Nodularia spumigena* AV1. *Tetrahedron Lett* 1997, **38**:5525–5528.

51. Caboche S, Leclere V, Pupin M, Kucherov G, Jacques P: Diversity of monomers in nonribosomal peptides: towards the prediction of origin and biological activity. *J Bacteriol* 2010, **192**:5143–5150.

52. Ohtake A, Shirai M, Aida T, Mori N, Harada K-I, Matsuura K, Suzuki M, Nakano M: Toxicity of *Microcystis* species isolated from natural blooms and purification of the toxin. *Appl Environ Microbiol* 1989, **55**:3202–3207.

53. Hoffmann D, Hevel JM, Moore RE, Moore BS: Sequence analysis and biochemical characterization of the nostopeptolide A biosynthetic gene cluster from *Nostoc* sp. GSV224. *Gene* 2003, **311**:171–180.

54. Utkilen H, Gjølme N: Iron-stimulated toxin production in *Microcystis aeruginosa*. *Appl Environ Microbiol* 1995, **61**:797–800.

55. Rohrlack T, Dittmann E, Henning M, Borner T, Kohl JG: Role of microcystins in poisoning and food ingestion inhibition of *Daphnia galeata* caused by the cyanobacterium *Microcystis aeruginosa*. *Appl Environ Microbiol* 1999, **65**:737–739.

56. Rohrlack T, Dittmann E, Börner T, Christoffersen K: Effects of cell-bound microcystins on survival and feeding of *Daphnia* spp. *Appl Environ Microbiol* 2001, **67**:3523–3529.

57. Dittmann E, Erhard M, Kaebernick M, Scheler C, Neilan BA, Von Döhren H, Börner T: Altered expression of two light-dependent genes in a microcystin-lacking mutant of *Microcystis aeruginosa* PCC 7806. *Microbiology* 2001, **147**:3113–3119.

58. Babica P, Bláha L, Maršálek B: Exploring the natural role of microcystins – a review of effects on photoautotrophic organisms. *J Phycol* 2006, **42**:9–20.

59. Schatz D, Keren Y, Vardi A, Sukenik A, Carmeli S, Börner T, Dittmann E, Kaplan A: Towards clarification of the biological role of microcystins, a family of cyanobacterial toxins. *Environ Microbiol* 2007, **4**:965–970.

60. Jenke-Kodama H, Dittmann E: Evolution of metabolic diversity: insights from microbial polyketide synthases. *Phytochemistry* 2009, **70**:1858–1866.

61. Challis GL, Ravel J, Townsend CA: Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem Biol* 2000, **7**:211–224.

62. Conti E, Stachelhaus T, Marahiel MA, Brick P: Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J* 1997, **16**:4174–4183.

63. Eppelmann K, Stachelhaus T, Marahiel MA: Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics. *Biochemistry* 2002, **41**:9718–9726.

64. Villiers BRM, Hollfelder F: Mapping the limits of substrate specificity of the adenylation domain of TycA. *Chembiochem* 2009, **10**:671–682.

65. Lin TP, Chen CL, Chang LK, Tschen JSM, Liu ST: Functional and transcriptional analyses of a fengycin synthetase gene, *fen* C, from *Bacillus subtilis*. *J Bacteriol* 1999, **181**:5060–5067.

66. Konz D, Doekel S, Marahiel MA: Molecular and biochemical characterization of the protein template controlling biosynthesis of the lipopeptide lichenysin. *J Bacteriol* 1999, **181**:133–140.

67. Rounge TB, Rohrlack T, Kristensen T, Jakobsen KS: Recombination and selectional forces in cyanopeptolin NRPS operons from highly similar, but geographically remote *Planktothrix* strains. *BMC Microbiol* 2008, **8**:141.

68. Allen MB: Simple conditions for growth of unicellular blue-green algae on plates. *J Phycol* 1968, **4**:1–4.

69. Kotai J: *Instructions for Preparation of Modified Nutrient Solution Z8 for Algae*. Norwegian Institute for Water Research, publication B-11/69. Oslo: Blindern; 1972.

70. 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.

71. Bachmann BO, Ravel J: In silico prediction of microbial secondary metabolic pathways from DNA sequence data. *Methods Enzymol* 2009, **458**:181–217.

72. Rausch C, Weber T, Kohlbacher O, Wohlleben W, Hudson DH: Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Res* 2005, **33**:5799–5808.

73. Röttig M, Medema MH, Bin K, Weber T, Rausch C, Kohbacher O: NRPSpredictor2—a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res* 2011, **39**:W362–W367.

74. Heath L, van der Walt E, Varsani A, Martin DP: Recombination patterns in aphthoviruses mirror those found in other picornaviruses. *J Virol* 2006, **80**:11827–11832.

75. Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P: RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* 2010, **26**:2462–2463.

76. Martin D, Rybicki E: RDP: detection of recombination amongst aligned sequences. *Bioinformatics* 2000, **16**:562–563.

77. Martin DP, Posada D, Crandall KA, Williamson C: A modified BOOTSCAN algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Res Hum Retroviruses* 2005, **21**:98–102.

78. Padidam M, Sawyer S, Fauquet CM: Possible emergence of new geminiviruses by frequent recombination. *Virology* 1999, **265**:218–225.

79. Maynard Smith J: Analyzing the mosaic structure of genes. *J Mol Evol* 1992, **34**:126–129.

80. Posada D, Crandall KA: Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proc Natl Acad Sci USA* 2001, **98**:13757–13762.

81. Gibbs MJ, Armstrong JS, Gibbs AJ: Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* 2000, **16**:573–582.

82. Boni MF, Posada D, Feldman MW: An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* 2007, **176**:1035–1047.

83. Simmonds P: SSE: A nucleotide and amino acid sequence analysis platform. *BMC Res Notes.* 2012, **5**:50.

84. Jokela J, Herfindal L, Wahlsten M, Permi P, Selheim F, Vasconcelos V, Døskeland SO, Sivonen K: A novel cyanobacterial nostocyclopeptide is a potent antitoxin against microcystins. *Chembiochem* 2010, **11**:1594–1599.