

Article

N-Terminal Protease Gene Phylogeny Reveals the Potential for Novel Cyanobactin Diversity in Cyanobacteria

Joana Martins ^{1,2}, Pedro N. Leão ², Vitor Ramos ^{1,2} and Vitor Vasconcelos ^{1,2,*}

¹ Faculty of Sciences, University of Porto, Rua do Campo Alegre, Porto 4169-007, Portugal; E-Mails: joana.martins@ciimar.up.pt (J.M.); vtr.rms@gmail.com (V.R.)

² CIIMAR/CIMAR—Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Rua dos Bragas 289, Porto 4050-123, Portugal; E-Mail: pleao@ciimar.up.pt

* Author to whom correspondence should be addressed; E-Mail: vmvascon@fc.up.pt; Tel.: +351-22-340-18-37; Fax: +351-22-339-06-08.

Received: 13 October 2013; in revised form: 21 November 2013 / Accepted: 21 November 2013 /

Published: 9 December 2013

Abstract: Cyanobactins are a recently recognized group of ribosomal cyclic peptides produced by cyanobacteria, which have been studied because of their interesting biological activities. Here, we have used a PCR-based approach to detect the N-terminal protease (A) gene from cyanobactin synthetase gene clusters, in a set of diverse cyanobacteria from our culture collection (Laboratory of Ecotoxicology, Genomics and Evolution (LEGE) CC). Homologues of this gene were found in *Microcystis* and *Rivularia* strains, and for the first time in *Cuspidothrix*, *Phormidium* and *Sphaerospermopsis* strains. Phylogenetic relationships inferred from available A-gene sequences, including those obtained in this work, revealed two new groups of phylotypes, harboring *Phormidium*, *Sphaerospermopsis* and *Rivularia* LEGE isolates. Thus, this study shows that, using underexplored cyanobacterial strains, it is still possible to expand the known genetic diversity of genes involved in cyanobactin biosynthesis.

Keywords: cyanobacteria; cyanobactins; bioactive compounds; phylogeny

1. Introduction

Cyanobacteria are one of the preferred microbial sources in the search of novel bioactive compounds due to the great structural diversity of their secondary metabolites [1,2]. Molecules derived

from polyketide synthase (PKS), non-ribosomal peptide synthetase (NRPS) and hybrid PKS/NRPS biosynthetic pathways constitute the majority of known cyanobacterial natural products [2].

Recently, several studies have focused on an emerging group of low-molecular weight cyclic peptides of ribosomal origin. These are structurally and biosynthetically related, thus justifying the establishment of a new metabolite family—the cyanobactins [3–5]. While many cyanobactins are simple peptide macrocycles (*N*-*C* terminal cyclized), most are additionally decorated with diverse posttranslational modifications, such as heterocyclization, isoprenylation or *N*-methylation, making this group of compounds one of the largest classes of diverse peptides [4,6]. Several pharmacological relevant activities, including cytotoxicity [7], antimalarial [8,9] and multidrug reversing activities [10] have been reported for cyanobactins. As an example, trunkamide, produced by *Prochloron* sp., was shown to be highly selective for the human renal UO-31 cell line and to have a favorable COMPARE profile at the National Cancer Institute, being listed as a pre-clinical candidate [11]. Recent studies, using a large diversity spectrum, have estimated that these peptides are found in 10%–30% of all cyanobacteria [1,6], being widespread among symbiotic as well as free-living cyanobacteria isolated from terrestrial, freshwater and marine environments [1,4,12].

Cyanobactin biosynthetic gene clusters, and their respective metabolites, have been described in cyanobacteria belonging to the unicellular genera *Prochloron* (patellamide, lissoclinamides, ulithiacyclamides, patellin and trunkamide) [13], *Microcystis* (microcyclamide, piricyclamide) [13,14] and *Cyanothece* (cyanothecamides) [6,15]; the filamentous non-heterocystous genera *Trichodesmium* (trichamide) [13], *Planktothrix* (prenylagaramide) [6,16], *Lyngbya* (aesturamide) [3,17], *Oscillatoria* (a full gene cluster was discovered, however the corresponding product has not been found) [6] and *Arthrospira* (arthrospiramide) [6]; and the filamentous heterocystous *Anabaena* (anacyclamide) [13] and *Nostoc* (tenuecyclamide) [13]. The patellamide gene cluster (*pat*), from the ascidian symbionts *Prochloron* spp., was the first cyanobactin biosynthetic gene cluster to be described [18]. Most of the homologous clusters reported since then resemble *pat*, although the gene arrangement is not precisely conserved [6,13]. Nevertheless, all clusters feature genes coding for two proteases (*N*-terminal protease—A, and *C*-terminal protease—G), which cleave precursor peptides (E) and macrocyclize the resulting fragments. Ribosomally translated precursor peptides, act as substrates for the posttranslational modifications, which imprint structural diversity to these metabolites. Furthermore, cyanobactin precursor peptides are hypervariable, meaning that identical enzymes can often synthesize numerous products depending on discrete changes in small cassettes [19]. The minimal set of genes present in cyanobactin gene clusters seems to include, apart from genes encoding A, G and E, two genes coding for short conserved hypothetical proteins (B and C) of unknown function [6]. Additionally, the cyclodehydratase (D) gene, responsible for heterocyclization and the prenyltransferase (F) gene, responsible for prenylation, may be present in the gene cluster, as well as thiazoline/oxazoline dehydrogenases, methyltransferases and other non-characterized proteins [20]. More recently, cyanobactins with more than one precursor peptide have been described [6].

Evidence from genomic studies [6,13] points towards the existence of a large number of cyanobactins yet to be discovered. With this in mind, we have employed a PCR-based strategy to identify potential cyanobactin producing strains in our culture collection. To this end, we designed degenerate oligonucleotide primers that were used to detect gene A from cyanobactin gene clusters among a set of diverse cyanobacteria. Phylogenetic relationships were inferred using all the available

A-gene sequences (including sequences from this study), which ultimately led to the identification of promising cyanobacterial strains/phylotypes from which novel cyanobactin chemotypes may be discovered.

2. Results

2.1. PCR Analyses

The selected cyanobacterial strains (see Experimental Section) were screened by PCR for the presence of cyanobactin *N*-terminal protease (A) gene. Apart from the strain used as positive control (*M. aeruginosa* LEGE 91351), this gene was detected in 11 other strains (Table 1). We were unable to detect gene A in cyanobacteria belonging to the order Pleurocapsales. A-gene sequences were obtained for strains belonging to the unicellular genus *Microcystis* (six isolates, including positive control, order Chroococcales), the filamentous non-heterocystous genus *Phormidium* (three isolates, order Oscillatoriales), and one isolate of each of the nostocalean genera *Cuspidothrix*, *Sphaerospermopsis* and *Rivularia*. The majority of these isolates were obtained from freshwater environments, with the exception of the marine strain *Rivularia* sp. LEGE 07159 (Table 1).

Table 1. Cyanobacterial strains positive for cyanobactin *N*-terminal protease (A) gene.

Strain	Habitat ^a	Source ^b	Accession Number	
			16S rRNA	(A) gene
<i>Cuspidothrix</i> sp. LEGE 03284	F	Montargil, Portugal	KC989703	KF008260
<i>Microcystis aeruginosa</i> LEGE 91347	F	Reservoir, Bemposta dam, Portugal	KC989705	KF008262
<i>Microcystis aeruginosa</i> LEGE 91351	F	Pond, Lagoa das Braças, Portugal	KC311966	KF008265
<i>Microcystis aeruginosa</i> LMECYA 1	F	Reservoir, Montargil dam, Portugal	KC989706	KF008263
<i>Microcystis</i> sp. IZANCYA 45	F	Pond, Lagoa da Vela, Portugal	KC311968	KF008266
<i>Microcystis</i> sp. LEGE 08328	F	Lake Zumpango, Mexico	KC989704	KF008261
<i>Microcystis</i> sp. LEGE 08331	F	Man-made channel, Cuemanco, Mexico	KC989707	KF008264
<i>Phormidium</i> sp. LEGE 06204	F	WWTP, Febros river, Portugal	KC989699	KF008256
<i>Phormidium</i> sp. LEGE 06363	F	WWTP, Febros river, Portugal	KC989700	KF008257
<i>Phormidium</i> sp. LEGE 07215	F	WWTP, Febros river, Portugal	KC989698	KF008255
<i>Rivularia</i> sp. LEGE 07159	M	Beach, Burgau, Portugal	KC989702	KF008259
<i>Sphaerospermopsis</i> sp. LEGE 00249	F	Reservoir, Maranhão dam, Portugal	KC989701	KF008258

^a M—marine, F—freshwater; ^b WWTP—Wastewater Treatment Plant.

2.2. Phylogenetic Analyses

The phylogeny for gene A (Figure 1), constructed from 58 cyanobacterial strains sequences (12 from this study) revealed six distinct clades, three of them (I, II and VI) encompassing a large number of sequences. Clade I (ML bootstrap support value of 70%) is formed by the unicellular *Cyanothece* sp. PCC 7822, one subclade of filamentous non-heterocystous *Oscillatoria* spp. (one of these—strain PCC 6506—known to contain the cyanobactin *osc* gene cluster) and a larger group branched in two major subclades, which in turn form two smaller groups each. One of the branches is composed only by filamentous heterocystous cyanobacteria. It contains a subclade with *Cuspidothrix* sp. LEGE 03284 and isolates assigned to *Anabaena* and *Aphanizomenon* spp., which include the

two strains known to produce anacyclamides (*Anabaena* sp. 90 and *Anabaena planctonica* 1tu33s10). In addition, it includes a subclade of three *Nodularia spumigena* strains. The other branch encompasses a subclade with two unicellular *Snowella litoralis* isolates, and also a subclade formed by the filamentous non-heterocystous genus *Planktothrix*, including the prenylagaramide-producing strain NIES-596. Clade II is formed by a broad subclade of unicellular *Microcystis* spp. sequences and by two phylogenetically distant sequences from distinct oscillatorial cyanobacteria (including the trichamide-producing strain *Trichodesmium erythraeum* IMS101). The subclade of cyanobactin A-gene sequences from *Microcystis* spp. is one of the two clusters comprising members of this genus featured in the phylogenetic tree. This particular cluster (in Clade II) includes the six *Microcystis* strains from this study and also two *M. aeruginosa* strains known to contain the piricyclamide (*pir*) gene cluster. Clade III comprises the sequences of the three filamentous non-heterocystous *Phormidium* spp. LEGE CC strains and that from the filamentous heterocystous *Sphaerospermopsis* sp. LEGE 00249. The unicellular, cyanothecamide-producing *Cyanothece* sp. PCC 7425 grouped with the filamentous heterocystous *Calothrix* sp. PCC 7103 (although with a large genetic distance), forming Clade IV. Clade V is also constituted by two sequences. However, in this case the genetic distance is less pronounced and the strains, *Rivularia* sp. LEGE 07159 and PCC 7116 belong to the same genus. Clade VI constitutes the most heterogeneous cluster in terms of cyanobacterial and cyanobactin diversities. As observed in Clade IV, this clade does not comprise any of the strains screened in this study. It includes the ungrouped sequence from *Leptolyngbya* sp. PCC 7376, a subclade with two strains from the likewise oscillatorial genus *Arthrospira*, including the arthrospiramide producer *A. platensis* NIES-39, and a larger subclade. This latter subclade harbors the ungrouped sequence from the nostocalean strain *Tolypothrix* sp. TOL328, a subclade with two *Prochloron* spp., which are unicellular marine cyanobacteria living symbiotically with ascidians and notably produce several different cyanobactins, and a still broad subclade divided in two main branches. The first forms a subclade with the sequences from other two strains of *Oscillatoria* spp. and from a filamentous heterocystous *Nostoc* species. This nostocalean strain is known to produce the tenuencyclamides. The second branch includes the remaining sequences from *Microcystis* strains. In this case, the isolates belonging to this genus are placed separately in two small clusters, each of them with a strain known to produce microcyclamide cyanobactins. These two groups of sequences are interleaved with those of two *Lyngbya* spp. (Oscillatoriales), one of which (strain PCC 8106) produces aesturamide. Finally, *Oscillatoria* sp. PCC 10802 is placed as a “loner sequence” (*i.e.*, having no close relatives) in the cyanobactin gene A phylogeny.

On the other hand, the phylogenetic tree of 16S rRNA gene sequences (Figure 2) revealed a clearly different topology. It covers almost all the strains included in the A-gene phylogeny (Figure 1), with the exception of four cyanobacteria for which the sequences were not available. In this case, all the 17 *Microcystis* strains show to be phylogenetically very close, clustering in a tight and well-supported clade (bootstrap value of 100%). Similarly, all the filamentous heterocystous cyanobacteria (including *Sphaerospermopsis* sp. LEGE 00249 and *Cuspidothrix* sp. LEGE 03284) cluster in a same clade. The three *Phormidium* LEGE strains have formed a tight subclade within a clade with other oscillatorial species.

Figure 1. Unrooted maximum-likelihood phylogenetic tree of partial cyanobactin N-terminal protease (A) gene sequences (819 bp). The six clusters are identified by I–VI. Numbers along branches indicate the percentage of bootstrap support considering 1000 pseudo-replicates: only those equal to or higher than 70% are indicated. Strains from this study are indicated in light blue, whereas strains producing known cyanobactins are shaded with different colors. Strain *Oscillatoria* sp. PCC 6506 contains the cyanobactin *osc* gene cluster but its product remains unknown (see text for details). Cyanobactin gene cluster from strain *Microcystis aeruginosa* NIES-843 is inactive [14]. Strains marked with an asterisk correspond to isolates with A-genes identified through antiSmash searches, not previously described in the literature.

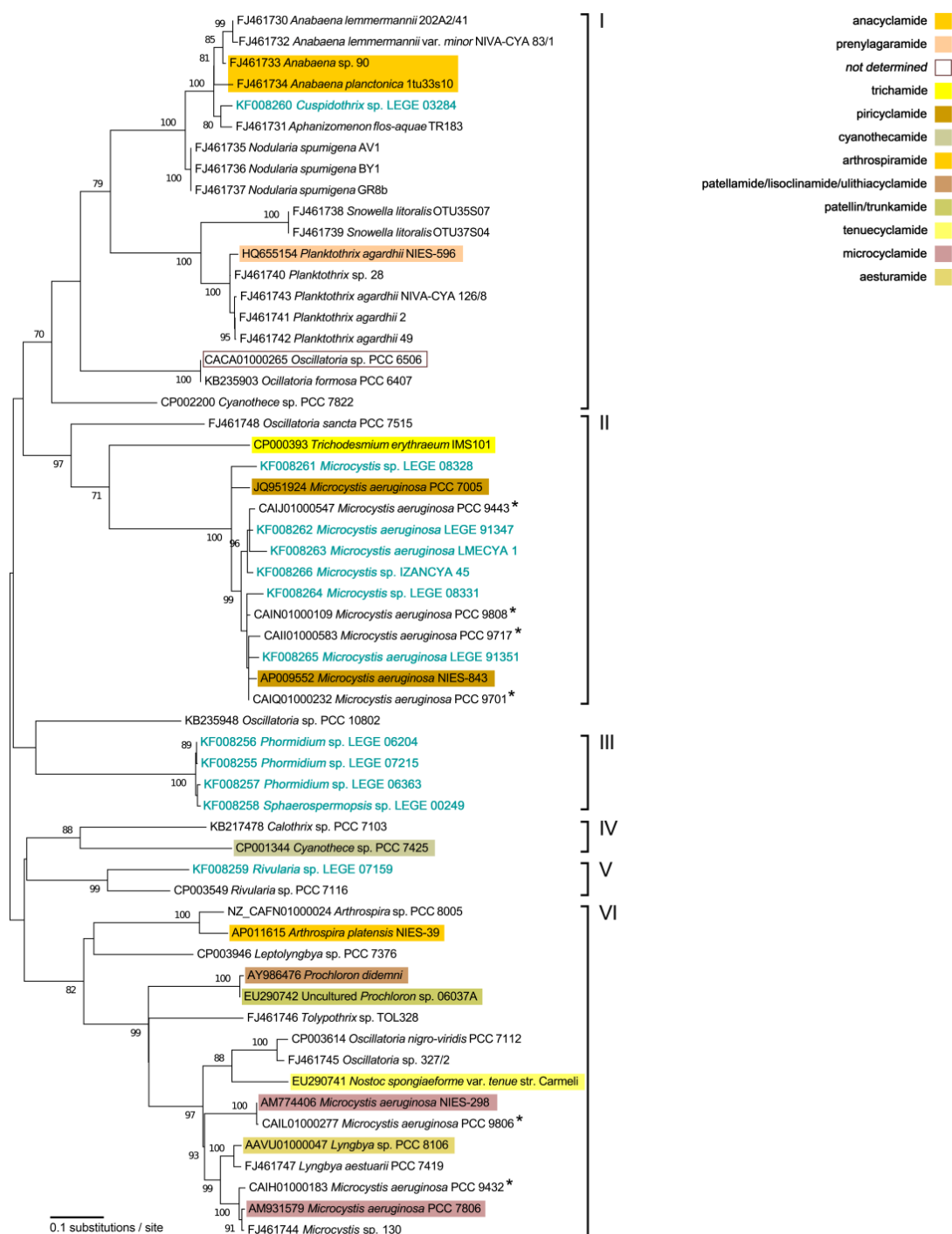
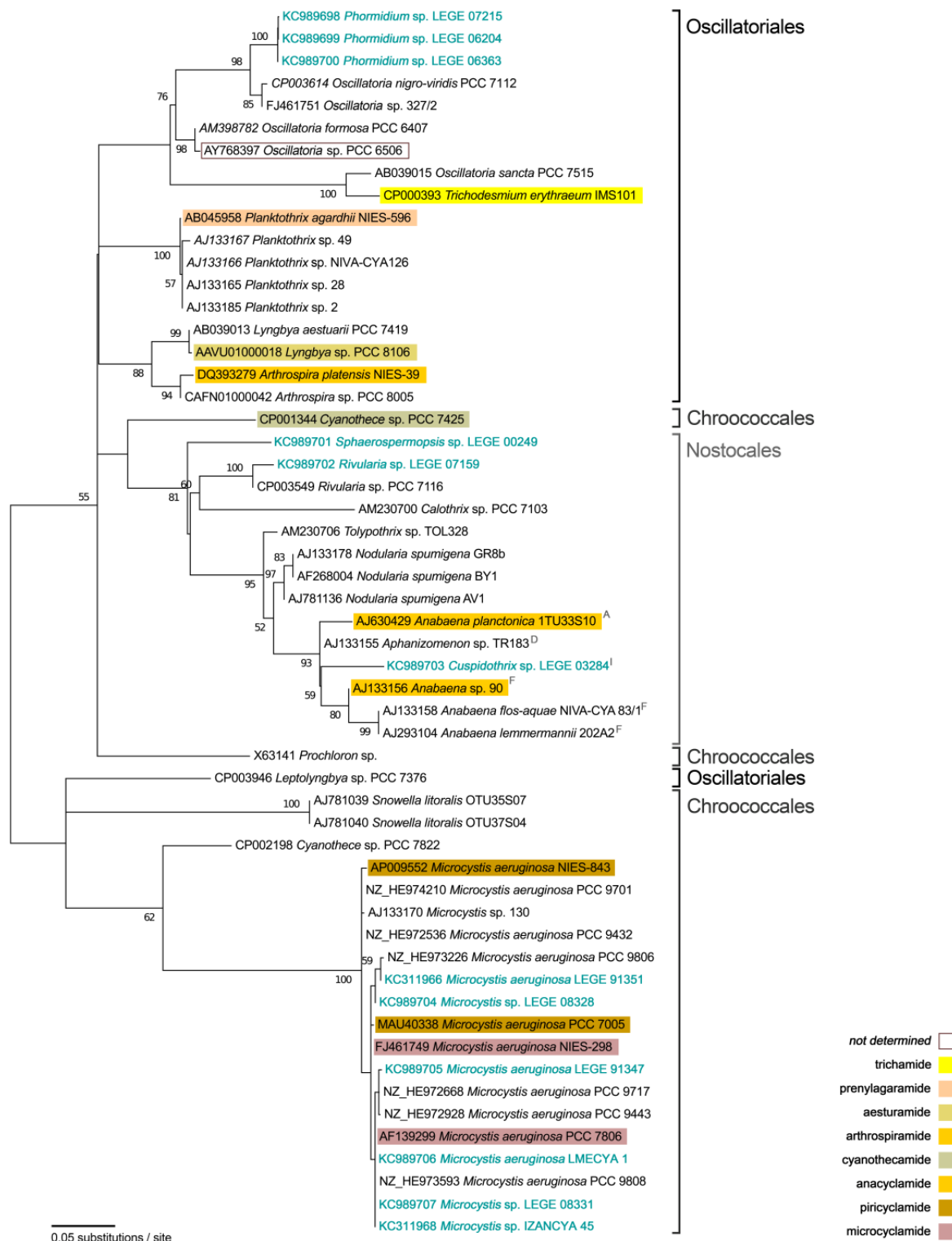


Figure 2. Unrooted maximum-likelihood phylogenetic tree of partial 16S rRNA gene sequences (760 bp). Numbers along branches indicate the percentage of bootstrap support considering 1000 pseudo-replicates: only those equal to or higher than 50% are indicated. Strains from this study are indicated in light blue. A, D, F and I refer to clades identified by Rajaniemni *et al.* [21]. Note: 16S rRNA gene sequences of *Oscillatoria* sp. PCC 10802, *Nostoc spongiaeforme* var. *tenue* str. Carmeli, *Prochloron didemni* and *Prochloron* sp. 06037A are not available and for that reason were not included in the tree.



3. Discussion

As a consequence of the strain selection criteria, nearly half of the isolates tested here correspond to cyanobacterial genera for which the presence of cyanobactin genes had not been studied or discovered. We were able to detect cyanobactin genes in approximately 13% of the tested strains (85), a smaller proportion than that previously reported by Leikoski *et al.* [12], which have found cyanobactin A-genes in 36% of a varied set of 132 cyanobacterial strains belonging to the major taxonomic divisions. Nevertheless, Shih and co-workers [1] found similar results to the ones obtained in this study, having detected cyanobactin genes in 19% of the available cyanobacterial genomes (24 out of 126).

The present phylogenetic analysis of the A-gene employed a much larger number of sequences than previous studies and benefited from both our screening effort and from the genome sequencing studies that have been completed recently (e.g., [1]). The A-gene phylogeny allows the clear identification of six clades, which should correspond to different cyanobactin chemotypes (Figure 1). Leikoski *et al.* [12] have studied the phylogenetic relationships among A-gene sequences from 25 cyanobacteria. The resulting tree revealed clustering into the clades represented here as Clades I, II and VI. Later, and with a larger number of cyanobactin gene sequences available, Donia and Schmidt [6] identified four A-protease phylotypes and linked them to the corresponding chemotypes. These formerly described phylotypes and chemotypes are congruent with our analysis and correspond to four out of the six clades reported here, namely Clades I, II, IV and VI (Genotypes II, III, IV and I, respectively, in the afore-mentioned study). The remaining Clades, III and V, constitute in this manner two new putative cyanobactin groups resulting from this study.

Considering the previously defined groups, Clade I comprises cyanobacterial strains producing anacyclamides and prenylagaramides. The A-gene from the cyanobactin *osc* gene cluster (from *Oscillatoria* sp. PCC 6506) is also part of this clade, however its cyanobactin product has not been reported to date. As pointed out by Donia and Schmidt [6], cyanobactin clusters bearing the A-gene from this clade lead to the production of cyclic peptides without heterocyclization or oxidation, which is consistent with the absence of the cyclodehydratase (D) gene and oxidase domain (in gene G) (as an illustration see Supplementary Figure S1). The exception stands for the *osc* gene cluster, since while it features a similar gene arrangement than other cyanobactin synthetase clusters from this clade, it also contains all the genes required for the production of oxidated heterocycles (a characteristic of the chemotypes found in Clade VI). Hence, the *osc* gene cluster appears to be a hybrid pathway between those with their A-genes present in Clades I and VI. In fact, Donia and Schmidt [6] predict that the cyanobactin(s) encoded in this pathway will present structural features from both phylotypes. Moreover, this gene cluster may be an ideal candidate to study cyanobactin evolution.

Trichamide- and piricyclamide-producing strains are encompassed in Clade II. The *tri* gene cluster, from *Trichodesmium erythraeum* IMS101, has previously been considered an individual and separate phylotype because it presents distinctive characteristics. It contains most of the genes that are present in the gene clusters from Clade VI, but with a very different organization. Furthermore, the oxidase domain is present in a separate open reading frame [6] (as an illustration see Supplementary Figure S1). Our results show the inclusion of A-genes from the *pir* cluster in this clade. This can be explained because the proteases (A and G) from the *pir*-bearing strains *Microcystis aeruginosa* PCC 7005 and

NIES-843 are almost identical to those of the *tri* pathway [14]. The A-gene from the *thc* gene cluster (*Cyanothece* sp. PCC 7425), placed in Clade IV, constitutes a separate genotype since the organization of this pathway is currently unique [6] (as an illustration see Supplementary Figure S1). Cyanobacterial strains that produce arthrospiramide (*Arthrospira platensis* NIES-39), patellamide, lissoclinamide, ulithiacyclamide, patellin and trunkamide (*Prochloron*), tenuocyclamide (*Nostoc spongiaeforme* var. *tenue* str. Carmeli), microcycyclamide (*Microcystis aeruginosa*) and aesturamides, (*Lyngbya* sp. PCC 8106) fit in Clade VI of the A-gene phylogeny. Cyanobactin gene clusters present in this clade lead to the production of oxidated heterocyclic compounds, with the exception of the *tru* gene cluster [6] (as an illustration see Supplementary Figure S1).

The novel clades arising from this study (III and V), although corresponding only to gene A, are expected to be part of cyanobactin biosynthetic pathways different from those currently known. Thus, the study of the *Phormidium* sp., *Sphaerospermopsis* sp. and *Rivularia* sp. LEGE strains bearing these novel A-genes, both at the genetic and chemical levels, is of great interest. This potential becomes more evident if we take into account that even cyanobacterial strains with A-genes belonging to the same clade (e.g., Clade I and VI) are able to produce different types of cyanobactins.

Other A-gene sequences obtained from this study grouped into previously described clades. *Cuspidothrix* sp. LEGE 03284 is present in a sub-clade (Clade I) that contains anacyclamide-producing strains. In a similar manner, A-genes from all the *Microcystis* strains from this study are found in the same sub-clade (Clade II), which also contains the *pir*-bearing *Microcystis* strains. It is thus expectable that the aforementioned strains from this study will produce anacyclamides and piricyclamides, respectively. The study of such eventually produced cyanobactins should be of interest, in particular as piricyclamides and anacyclamides are known to display a considerable chemical diversity [14,22].

The existence of cyanobactin genes has been previously reported in *Anabaena* and *Aphanizomenon* [12] strains. However, recent taxonomic revisions have established several novel genera arising from the two aforementioned taxa. For instance, the case of the new proposed genus *Dolichospermum* [23], which comprehends the strain 1tu33s10, previously assigned to *Anabaena planctonica* (Figure 2). In contrast, the presence of cyanobactin synthetase genes in the genera *Cuspidothrix* and *Sphaerospermopsis*, which have also been recently established from *Anabaena* and *Aphanizomenon* taxonomic revisions [21,24,25], is unprecedented to our knowledge (Figure 2). This is also the first report of cyanobactin genes in the genus *Phormidium*.

The phylogeny of gene A also reveals several cyanobacterial strains, apart from those selected herein, from which new cyanobactins may be discovered. As an example, the *Nodularia spumigena* and *Snowella litoralis* with A-genes present in Clade I, are not known to produce cyanobactins, nor are other members of these genera. One other good candidate is *Cyanothece* sp. PCC 7822, which is present in the phylogenetic tree in a separate, distant branch from the cyanothecamide producer *Cyanothece* sp. PCC 7425. *Oscillatoria* strains present throughout the phylogeny, and in particular the strain containing the A-gene “loner sequence”, *Oscillatoria* sp. PCC 10802, represent potential candidates in the discovery of novel cyanobactins, which have not been reported for any member of this genus. Even though we could not retrieve the 16S rRNA gene sequence for this strain, the work of Shih *et al.* [1] allows inferring that this strain would be placed, in our phylogeny (Figure 2), in the clade harboring *Oscillatoria* spp. and *Phormidium* sp. sequences.

The lack of congruence between the A-gene and 16S rRNA gene phylogenies (Figures 1 and 2), highlights that the ability to produce cyanobactins has, most probably, been laterally transferred (via horizontal transfer of gene clusters) from one taxon to another, as earlier suggested by Leikoski *et al.* [12]. This becomes more evident from the comparison of the cyanobacterial strains that are clustered in Clade VI (Figure 1) to their diffuse distribution among the 16S rRNA gene phylogeny (Figure 2). One other good example concerns the *Microcystis* spp. strains that are grouped separately (Clades II and VI, Figure 1), on the A-gene phylogeny and clustered together in the 16S rRNA gene phylogeny (Figure 2). Still, and as expected from previous studies, the phylogenetic tree (Figure 2) features the well-known polyphyly of most of the cyanobacterial orders [26–28], and even genera (as examples, see [28–30]), recognized in traditional systematics. However, the also documented monophyly of the heterocystous cyanobacteria [31] and that of the “traditional” genus *Microcystis* [32] was verified. The closely related *Phormidium* LEGE strains included in this study are placed within a clade with *Oscillatoria* spp., which, in turn, and unlike what can be seen in the A-gene phylogeny (Figure 1), have clustered together. These observations suggest that this particular group (*i.e.*, 16S rRNA gene phylotype) of filamentous non-heterocystous cyanobacteria can/may possess the ability to produce a variable set of currently unknown cyanobactins.

4. Experimental Section

4.1. Cyanobacterial Strains

A selection of cyanobacterial strains (Table 2 and Supplementary Figure S1) from the LEGE Culture Collection (Laboratory of Ecotoxicology, Genomics and Evolution; CIIMAR, Porto, Portugal) was performed in order to cover the diversity in the collection and the different ecosystems from which the strains were isolated. We tried to include cyanobacterial genera that had been tested in former studies [12], as well as genera that had not been previously examined. These were screened for the presence of the N-terminal protease (A) gene, and belonged to the orders Chroococcales (33 strains), Pleurocapsales (5 strains), Oscillatoriales (29 strains) and Nostocales (18 strains), according to the traditional systems of classification [33–36]. The isolates were originally obtained from freshwater (37 strains), estuarine (two strains) and marine (46 strains) environments. Additionally, three ESSACC (Estela Sousa e Silva Algal Culture Collection) strains [37] from freshwater environments were also included (hereafter named as LMECYA strains).

4.2. Degenerate Primers

PCR primers were developed with the aim of being “universal” both in terms of taxa and cyanobactin diversity. For that purpose, publicly available (*i.e.*, in GenBank) cyanobactin A-gene sequences from different cyanobacterial taxa were aligned using the Clustal W algorithm [38] in the MEGA 5.05 software package [39]. Conserved regions among the homologous genes were thus identified and evaluated for candidate priming sites. This allowed us to design the primer set CBT_AF (5'-TTVGGYTAYGAYTTYGG-3') and CBT_AR (5'-AGACCARGAACGRACTTC-3'), which amplify a 804 bp region of gene A. This primer pair was tested using a DNA template from *Microcystis aeruginosa* LEGE 91351 (formerly known as IZANCYA 41), a strain that was previously

shown to be positive for the presence of cyanobactin genes [14]. Still, the corresponding A-gene sequence was obtained, as described below, and had not been reported yet. The PCR protocol was optimized, by performing reactions under an annealing temperature gradient (48 °C–60 °C, optimal 52 °C).

Table 2. Cyanobacterial genera tested for the presence of cyanobactin *N*-terminal protease (A) gene.

Order	Genus	Habitat(s) ^a	Number of tested strains	Number of positive strains
Chroococcales	<i>Cyanobium</i>	M	9	0
	<i>Gloeocapsa</i>	M	1	0
	<i>Microcystis</i>	F	15	5
	<i>Synechococcus</i>	M	5	0
	<i>Synechocystis</i>	F; M	3	0
Pleurocapsales	<i>Chroococcopsis</i>	M	2	0
	<i>Chroococciopsis</i>	M	2	0
	<i>Hyella</i>	M	1	0
Oscillatoriales	<i>Leptolyngbya</i>	F; M	9	0
	<i>Nodosilinea</i>	M	2	0
	<i>Oscillatoria</i>	F	3	0
	<i>Phormidium</i>	F; E	4	3
	<i>Planktothrix</i>	F	3	0
	<i>Plectonema</i>	M	1	0
	<i>Pseudanabaena</i>	M	4	0
	<i>Romeria</i>	M	1	0
	<i>Schizothrix</i>	M	1	0
	Unidentified	M	1	0
	Pseudanabaenaceae			
Nostocales	<i>Anabaena</i>	F	3	0
	<i>Aphanizomenon</i>	F	4	0
	<i>Calothrix</i>	M	3	0
	<i>Cuspidothrix</i>	F	1	1
	<i>Cylindrospermopsis</i>	F	2	0
	<i>Nodularia</i>	E	1	0
	<i>Nostoc</i>	M	1	0
	<i>Rivularia</i>	M	1	1
	<i>Scytonema</i>	M	1	0
<i>Sphaerospermopsis</i>	F	1	1	

^a M—marine; F—freshwater; E—estuarine.

4.3. DNA Extraction, PCR Analyses and Sequencing

Exponentially growing cultures of each strain were harvested (1.5–2 mL) and the biomass centrifuged for 2 min at 7000× *g*. Total genomic DNA was extracted from cyanobacterial pellets using the Purelink Genomic DNA Mini Kit (Invitrogen, Carlsbad, CA, USA) following the manufacturer's procedure for Gram-negative bacteria.

Cyanobacteria-specific primers 27F and CYA781R [40,41] and the primers CBT_AF and CBT_AR were used to amplify 16S rRNA genes (small subunit ribosomal gene) and cyanobactin A-genes, respectively, from the strains. PCR reactions were prepared in a volume of 20 μ L containing 2.5 mM MgCl₂, 125 μ M of each deoxynucleotide triphosphate, 10 μ M of each of the primers, 0.5 U of GoTaq[®] DNA polymerase (Promega, Fitchburg, WI, USA) and 5–10 ng of DNA. Thermal cycling was carried out using a MyCycler (Bio-Rad, Hercules, CA, USA). In order to amplify the 16S rRNA gene, PCR analysis was performed with an initial denaturation step at 94 °C for 4 min followed by 30 cycles of 94 °C for 45 s, annealing temperature at 50 °C for 45 s, 72 °C for 1.20 min and a final extension step at 72 °C for 8 min. For the CBT_AF/CBT_AR primer-pair, thermal cycling conditions were as follows: initial denaturation at 94 °C for 3 min followed by 35 cycles of 94 °C for 45 s, annealing temperature at 52 °C for 45 s, 72 °C for 50 s and a final extension step at 72 °C for 7 min.

PCR products were separated in a 1.5% agarose gel in 1 \times TAE Buffer, stained with ethidium bromide (Bio-Rad, Hercules, CA, USA) and photographed under UV transillumination. DNA fragments from PCR products were gel extracted and purified with Cut&Spin DNA Gel Extraction Columns (Grisp, Porto, Portugal), following the manufacturer's procedure. The purified DNA was then cloned into the pGEM[®]-T Easy Vector (Promega, Fitchburg, WI, USA). The vector was used to transform chemically competent *Escherichia coli* ONE Shot[®] TOP10 cells (Invitrogen, Carlsbad, CA, USA). Plasmid DNA was isolated from the transformed cells using the GenElute[™] Plasmid Miniprep kit (Sigma-Aldrich, Saint Louis, MO, USA) and sequenced at Macrogen (Seoul, Korea) using M13/pUC sequencing primers.

4.4. Nucleotide Sequence Accession Numbers

16S rRNA gene and cyanobactin A-gene sequences obtained in this study were deposited in GenBank under the accession numbers [KC989698-KC989707] and [KF008255-KF008266], respectively (see also Table 1).

4.5. Phylogenetic Analyses

Phylogenetic studies of A-genes and 16S rRNA genes from the different cyanobacterial strains were performed. The putative A-gene sequences obtained in this study were independently submitted as the queries in a BLASTn search in order to confirm their identity. Since portions of the A- and G-genes are homologous, some of the sequences corresponded to the G-gene and were discarded. Cyanobactin A-gene sequences previously reported for cyanobacteria were also retrieved from the same database, as well as their respective 16S rRNA gene sequences (whenever available). Shih *et al.* [1] predicted cyanobactin synthetase gene clusters in several cyanobacterial genome sequences but in some cases, their genes were not fully annotated. Hence, we sought to identify the position of A-genes in the genome of those cyanobacterial strains by using anti SMASH [42] searches. This allowed retrieving the corresponding nucleotide sequences, but also identifying cyanobactin A-genes, hitherto not described in the literature, in other cyanobacterial genomes. These sequences were also included in the phylogenetic analyses.

MEGA 5.05 was used to perform the phylogenetic analyses. Multiple sequence alignments were obtained using the MUSCLE algorithm [43]. To compute the phylogenetic trees of A-genes and

16S rRNA genes, Maximum-Likelihood (ML) [44] methodology was employed using the model of evolution that best fitted the dataset. For both analyses, the Kimura 2-parameter model with gamma-distributed rate variation, using five substitution rate categories, and a proportion of invariant sites (K2P+ Γ +I) was then selected, and 1000 bootstrap replicates were generated.

5. Conclusions

Phylogenetic analysis of cyanobactin *N*-terminal protease (A) genes seems to be an effective strategy to screen for cyanobacterial strains that may produce novel cyanobactins. Using this approach, two novel putative cyanobactin groups (Clades III and V, Figure 1) emerged, providing an up-to-date picture of cyanobactin A-gene diversity. The findings from this study emphasize the importance of prospecting the still uncovered cyanobacterial “tree of life”. Pursuing this will certainly expand the number of compounds (and chemotypes) from this important group of secondary metabolites.

Acknowledgments

The authors acknowledge funding through the Project MARBIOTECH (reference NORTE-07-0124-FEDER-000047) within the SR&TD Integrated Program MARVALOR—Building research and innovation capacity for improved management and valorization of marine resources, supported by the Programa Operacional Regional do Norte (ON.2—O Novo Norte) and by the European Regional Development Fund (ERDF), INTERREG IV project Pharamatlantic. This work was also partially financed by ERDF through the COMPETE program and by national funds through FCT (Fundação para a Ciência e a Tecnologia), under the framework of projects PTDC/MAR/099642/2008, PTDC/MAR/102258/2008, PesT-C/MAR/LA0015/2013 and PTDC/MAR-BIO/2818/2012. Joana Martins and Vitor Ramos acknowledge FCT, Fundo Social Europeu (FSE) and Ministério da Ciência, Tecnologia e Ensino Superior (MCTES) for co-funding, respectively, the PhD scholarships SFRH/BD/80426/2011 and SFRH/BD/80153/2011. Pedro N. Leão was supported by a postdoctoral fellowship from FCT (SFRH/BPD/70233/2010).

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Shih, P.M.; Wu, D.; Latifi, A.; Axen, S.D.; Fewer, D.P.; Talla, E.; Calteau, A.; Cai, F.; Tandeau de Marsac, N.; Rippka, R.; *et al.* Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc. Natl. Acad. Sci. USA* **2012**, *110*, 1053–1058.
2. Tidgewell, K.; Clark, B.R.; Gerwick, W.H. The Natural Products Chemistry of Cyanobacteria. In *Comprehensive Natural Products II Chemistry and Biology*; Mander, L., Lui, H.-W., Eds.; Elsevier: Oxford, UK, 2010; Volume 2, pp. 141–188.
3. Donia, M.S.; Ravel, J.; Schmidt, E.W. A global assembly line for cyanobactins. *Nat. Chem. Biol.* **2008**, *4*, 341–343.

4. Donia, M.S.; Schmidt, E.W. Cyanobactins—Ubiquitous Cyanobacterial Ribosomal Peptide Metabolites. In *Comprehensive Natural Products II. Chemistry and Biology*; Mander, L., Lui, H.-W., Eds.; Elsevier: Oxford, UK, 2010; Volume 2, pp. 539–558.
5. Schmidt, E.W.; Donia, M.S. Chapter 23. Cyanobactin ribosomally synthesized peptides—A case of deep metagenome mining. *Meth. Enzymol.* **2009**, *458*, 575–596.
6. Donia, M.S.; Schmidt, E.W. Linking chemistry and genetics in the growing cyanobactin natural products family. *Chem. Biol.* **2011**, *18*, 508–519.
7. In, Y.; Doi, M.; Inoue, M.; Ishida, T.; Hamada, Y.; Shioiri, T. Patellamide A, a cytotoxic cyclic peptide from the ascidian *Lissoclinum patella*. *Acta Crystallogr. C* **1994**, *50*, 432–434.
8. Linington, R.G.; Gonzalez, J.; Urena, L.D.; Romero, L.I.; Ortega-Barria, E.; Gerwick, W.H. Venturamides A and B: Antimalarial constituents of the panamanian marine cyanobacterium *Oscillatoria* sp. *J. Nat. Prod.* **2007**, *70*, 397–401.
9. Portmann, C.; Blom, J.F.; Kaiser, M.; Brun, R.; Juttner, F.; Gademann, K. Isolation of aerucyclamides C and D and structure revision of microcyclamide 7806A: Heterocyclic ribosomal peptides from *Microcystis aeruginosa* PCC 7806 and their antiparasite evaluation. *J. Nat. Prod.* **2008**, *71*, 1891–1896.
10. Ogino, J.; Moore, R.E.; Patterson, G.M.; Smith, C.D. Dendroamides, new cyclic hexapeptides from a blue-green alga. Multidrug-resistance reversing activity of dendroamide A. *J. Nat. Prod.* **1996**, *59*, 581–586.
11. Caba, J.M.; Rodriguez, I.M.; Manzanares, I.; Giralt, E.; Albericio, F. Solid-phase total synthesis of trunkamide A (1). *J. Org. Chem.* **2001**, *66*, 7568–7574.
12. Leikoski, N.; Fewer, D.P.; Sivonen, K. Widespread occurrence and lateral transfer of the cyanobactin biosynthesis gene cluster in cyanobacteria. *Appl. Environ. Microbiol.* **2009**, *75*, 853–857.
13. Sivonen, K.; Leikoski, N.; Fewer, D.P.; Jokela, J. Cyanobactins—Ribosomal cyclic peptides produced by cyanobacteria. *Appl. Microbiol. Biotechnol.* **2010**, *86*, 1213–1225.
14. Leikoski, N.; Fewer, D.P.; Jokela, J.; Alakoski, P.; Wahlsten, M.; Sivonen, K. Analysis of an inactive cyanobactin biosynthetic gene cluster leads to discovery of new natural products from strains of the genus *Microcystis*. *PLoS One* **2012**, *7*, e43002.
15. Houssen, W.E.; Koehnke, J.; Zollman, D.; Vendome, V.; Raab, A.; Smith, M.C.M.; Naismith, J.H.; Jaspars, M. The discovery of new cyanobactins from *Cyanothece* PCC 7425 defines a new signature for processing of patellamides. *ChemBioChem* **2012**, *13*, 2625–2776.
16. Murakami, M.; Itou, Y.; Ishida, K.; Shin, H.J. Prenylagaramides A and B, new cyclic peptides from two strains of *Oscillatoria agardhii*. *J. Nat. Prod.* **1999**, *62*, 752–755.
17. McIntosh, J.A.; Lin, Z.; Tianero, M.D.; Schmidt, E.W. Aesturamides, a natural library of cyanobactin cyclic peptides resulting from isoprene-derived Claisen rearrangements. *ACS Chem. Biol.* **2013**, doi:10.1021/cb300614c.
18. Schmidt, E.W.; Nelson, J.T.; Rasko, D.A.; Sudek, S.; Eisen, J.A.; Haygood, M.G.; Ravel, J. Patellamide A and C biosynthesis by a microcin-like pathway in *Prochloron didemni*, the cyanobacterial symbiont of *Lissoclinum patella*. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 7315–7320.

19. Donia, M.S.; Hathaway, B.J.; Sudek, S.; Haygood, M.G.; Rosovitz, M.J.; Ravel, J.; Schmidt, E.W. Natural combinatorial peptide libraries in cyanobacterial symbionts of marine ascidians. *Nat. Chem. Biol.* **2006**, *2*, 729–735.
20. Arnison, P.G.; Bibb, M.J.; Bierbaum, G.; Bowers, A.A.; Bugni, T.S.; Bulaj, G.; Camarero, J.A.; Campopiano, D.J.; Challis, G.L.; Clardy, J.; *et al.* Ribosomally synthesized and post-translationally modified peptide natural products: Overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **2012**, *30*, 108–160.
21. Rajaniemi, P.; Hrouzek, P.; Kaštovská, K.; Willame, R.; Rantala, A.; Hoffmann, L.; Komárek, J.; Sivonen, K. Phylogenetic and morphological evaluation of the genera *Anabaena*, *Aphanizomenon*, *Trichormus* and *Nostoc* (Nostocales, Cyanobacteria). *Int. J. Syst. Ecol. Microbiol.* **2005**, *55*, 11–26.
22. Leikoski, N.; Fewer, D.P.; Jokela, J.; Wahlsten, M.; Rouhiainen, L.; Sivonen, K. Highly diverse cyanobactins in strains of the genus *Anabaena*. *Appl. Environ. Microbiol.* **2010**, *76*, 701–709.
23. Wacklin, P.; Hoffmann, L.; Komárek, J. Nomenclatural validation of the genetically revised cyanobacterial genus *Dolichospermum* (Ralfs ex Bornet et Flahault) comb. nova. *Fottea* **2009**, *9*, 59–64.
24. Rajaniemi, P.; Komárek, J.; Hrouzek, P.; Willame, R.; Kaštovská, K.; Hoffmann, L.; Sivonen, K. Taxonomic consequences from the combined molecular and phenotype evaluation of selected *Anabaena* and *Aphanizomenon* strains. *Arch. Hydrobiol. Algol. Stud.* **2005**, *117* (Suppl.), 371–391.
25. Zapomělová, E.; Jezberová, J.; Hrouzek, P.; Hisem, D.; Rěháková, K.; Komárková, J. Polyphasic characterization of three strains of *Anabaena reniformis* and *Aphanizomenon aphanizomenoides* (cyanobacteria) and their reclassification to *Sphaerospermum* gen. nov. (incl. *Anabaena kisseleviana*). *J. Phycol.* **2009**, *45*, 1363–1373.
26. Gugger, M.F.; Hoffmann, L. Polyphyly of true branching cyanobacteria (Stigonematales). *Int. J. Syst. Evol. Microbiol.* **2004**, *54*, 349–357.
27. Ishida, T.; Watanabe, M.M.; Sugiyama, J.; Yokota, A. Evidence for polyphyletic origin of the members of the orders of Oscillatoriales and Pleurocapsales as determined by 16S rDNA analysis. *FEMS Microbiol. Lett.* **2001**, *201*, 79–82.
28. Litvaitis, M.K. A molecular test of cyanobacterial phylogeny: Inferences from constraint analyses. *Hydrobiologia* **2002**, *468*, 135–145.
29. Honda, D.; Yokota, A.; Sugiyama, J. Detection of seven major evolutionary lineages in cyanobacteria based on the 16S rRNA gene sequence analysis with new sequences of five marine *Synechococcus* strains. *J. Mol. Evol.* **1999**, *48*, 723–739.
30. Marquardt, J.; Palinska, K.A. Genotypic and phenotypic diversity of cyanobacteria assigned to the genus *Phormidium* (Oscillatoriales) from different habitats and geographical sites. *Arch. Microbiol.* **2007**, *187*, 397–413.
31. Henson, B.J.; Hesselbrock, S.M.; Watson, L.E.; Barnum, S.R. Molecular phylogeny of the heterocystous cyanobacteria (subsections IV and V) based on *nifD*. *Int. J. Syst. Evol. Microbiol.* **2004**, *54*, 493–497.
32. Otsuka, S.; Suda, S.; Shibata, S.; Oyaizu, H.; Matsumoto, S.; Watanabe, M.M. A proposal for the unification of five species of the cyanobacterial genus *Microcystis* Kützing ex Lemmermann 1907 under the rules of the Bacteriological Code. *Int. J. Syst. Evol. Microbiol.* **2001**, *51*, 873–879.

33. Waterbury, J.B.; Stanier, R.Y. Patterns of growth and development in pleurocapsalean cyanobacteria. *Microbiol. Mol. Biol. Rev.* **1978**, *42*, 2–44.
34. Komárek, J.; Anagnostidis, K. Modern approach to the classification system of Cyanophytes: 4-Nostocales. *Arch. Hydrobiol. Suppl.* **1989**, *82*, 247–345.
35. Komárek, J.; Anagnostidis, K. Cyanoprokaryota, Part 1: Chroococcales. In *Süßwasserflora von Mitteleuropa*; Ettl, H., Gärtner, G., Heynig, H., Mollenhauer, D., Eds.; Spektrum Akademischer Verlag: Heidelberg, Germany, 1998; Volume 19/1, p. 548.
36. Komárek, J.; Anagnostidis, K. Cyanoprokaryota, Part 2: Oscillatoriales. In *Süßwasserflora von Mitteleuropa*; Büdel, B., Gärtner, G., Krienitz, L., Schagerl, M., Eds.; Elsevier: Munich, Germany, 2005; Volume 19/2, p. 759.
37. Paulino, S.; Sam-Bento, F.; Churro, C.; Alverca, E.; Dias, E.; Valério, E.; Pereira, P. The Estela Sousa e Silva Algal Culture Collection: A resource of biological and toxicological interest. *Hydrobiologia* **2009**, *636*, 489–492.
38. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **1994**, *22*, 4673–4680.
39. 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.
40. Neilan, B.A.; Jacobs, D.; del Dot, T.; Blackall, L.L.; Hawkins, P.R.; Cox, P.T.; Goodman, A.E. rRNA sequences and evolutionary relationships among toxic and non-toxic cyanobacteria of the genus *Microcystis*. *Int. J. Syst. Bacteriol.* **1997**, *47*, 693–697.
41. Nübel, U.; Garcia-Pichel, F.; Muyzer, G. PCR primers to amplify 16S rRNA genes from cyanobacteria. *Appl. Environ. Microbiol.* **1997**, *63*, 3327–3332.
42. Blin, K.; Medema, M.H.; Kazempour, D.; Fischbach, M.A.; Breitling, R.; Takano, E.; Weber, T. antiSMASH 2.0—A versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **2013**, doi:10.1093/nar/gkt449.
43. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797.
44. Felsenstein, J. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* **1981**, *17*, 368–376.