Published in final edited form as:

Environ Microbiol. 2011 June ; 13(6): 1601–1610. doi:10.1111/j.1462-2920.2011.02472.x.

Underestimated biodiversity as a major explanation for the perceived rich secondary metabolite capacity of the cyanobacterial genus *Lyngbya*

Niclas Engene¹, Hyukjae Choi¹, Eduardo Esquenazi^{1,2}, Erin C. Rottacker³, Mark H. Ellisman³, Pieter C. Dorrestein^{4,5}, and William H. Gerwick^{1,5,*}

¹Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA.

²Division of Biological Sciences, University of California San Diego, La Jolla, CA, USA.

³National Center for Microscopy and Imaging Research, University of California San Diego, La Jolla, CA, USA.

⁴Departments of Pharmacology, Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA, USA.

⁵Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, USA.

Summary

Marine cyanobacteria are prolific producers of bioactive secondary metabolites responsible for harmful algal blooms as well as rich sources of promising biomedical lead compounds. The current study focused on obtaining a clearer understanding of the remarkable chemical richness of the cyanobacterial genus *Lyngbya*. Specimens of *Lyngbya* from various environmental habitats around Curaçao were analyzed for their capacity to produce secondary metabolites by genetic screening of their biosynthetic pathways. The presence of biosynthetic pathways was compared with the production of corresponding metabolites by LC-ESI-MS² and MALDI-TOF-MS. The comparison of biosynthetic capacity and actual metabolite production revealed no evidence of genetic silencing in response to environmental conditions. On a cellular level, the metabolic origin of the detected metabolites was pinpointed to the cyanobacteria, rather than the sheath-associated heterotrophic bacteria, by MALDI-TOF-MS and multiple displacement amplification of single-cells. Finally, the traditional morphology-based taxonomic identifications of these *Lyngbya* populations were combined with their phylogenetic relationships. As a result, polyphyly of morphologically similar cyanobacteria was identified as the major explanation for the perceived chemical richness of the genus *Lyngbya*, a result which further underscores the need to revise the taxonomy of this group of biomedically important cyanobacteria.

Keywords

bioactive secondary metabolites; biosynthetic pathways; *Lyngbya*; multiple displacement amplification

*For correspondence. wgerwick@ucsd.edu; Tel. (+1) 858 534 0578; Fax (+1) 858-534-0529.

Introduction

Natural products discovery programs focused on the cyanobacterial genus *Lyngbya* continue to yield an extraordinary diversity of biologically active secondary metabolites (Gerwick *et al.*, 2008; Tidgewell *et al.*, 2009). Marine, tropical forms of *Lyngbya* in particular have been very affluent in this regard and have yielded over 260 different secondary metabolites (>40% of all reported marine cyanobacterial molecules) (MarinLit, 2010). Even more remarkable is the fact that a total of 196 different secondary metabolites have been reported from a single species, *Lyngbya majuscula*.

The natural biological roles of these secondary metabolites are often ascribed to defensive functions with many having potent toxicity (Thacker *et al.*, 1997; Capper *et al.*, 2006). Exploitation of these bioactivities has yielded a number of important natural products with therapeutic potential (Tan, 2007; Gerwick *et al.*, 2008). Hence, a clear understanding of the extent and origin of this chemical diversity in *Lyngbya* is important to enhancing the discovery of new natural products (NP) as well as for predicting some types of harmful algae blooms (HAB).

Despite nearly 30 years of investigation of cyanobacteria for their unique secondary metabolites, the full extent of the biosynthetic capacities of these microorganisms is still largely unknown. Transcriptional expression of cyanobacterial secondary metabolite pathways have been shown to be influenced by environmental factors (Kaebernick *et al.*, 2000; Shalev-Malul *et al.*, 2008; Sorrels *et al.*, 2009). Differential expression of secondary metabolites depending on environmental conditions has also been highlighted as an alternative explanation for the reported chemical diversity from genetically related *Lyngbya* populations at different collection sites (Thacker & Paul, 2004). Unfortunately, at the present time there is little genomic information available for these largely marine, filamentous cyanobacteria, so a true assessment of their capacity for natural products biosynthesis remains hidden (NCBI Microbial Genomes).

The identification of the vast majority of chemically interrogated strains of cyanobacteria has mainly been founded on traditional morphology-based taxonomic systems. These “morpho-species” of the genus *Lyngbya* are traditionally defined as filamentous non-heterocystous cyanobacteria with discoid cells enclosed within distinct sheaths (Castenholz, 2001; Komárek & Anagnostidis, 2005). The recent inclusion of phylogenetics into taxonomic classification, however, greatly enhances understanding of the relatedness of cyanobacteria and has led to major revisions of traditional genera (Hoffman *et al.*, 2005). Along these lines the genus *Lyngbya* has been recognized as a polyphyletic group (Sharp *et al.*, 2009; Engene *et al.*, 2010). Moreover, secondary metabolites described from *Lyngbya* morpho-types have been reisolated from distantly related phylogenetic lineages (Sharp *et al.*, 2009). However, the true taxonomic identities of these different *Lyngbya* lineages have not yet been clarified in reference to type-strains.

An alternative explanation for the distribution of secondary metabolites among evolutionary distinct *Lyngbya* populations could be that microorganisms associated with the cyanobacterial filaments are responsible for the production of the secondary metabolites. *Lyngbya* populations often form extensive mat-like colonies which are frequently colonized with complex microbial communities of diverse heterotrophic bacteria or other epiphytic cyanobacteria (Engene *et al.*, 2010). Dissection of cyanobacterial assemblages has shown that in some cases cyanobacteria which are epiphytic on *Lyngbya* can also produce bioactive secondary metabolites (Simmons *et al.*, 2008). Moreover, the thick polysaccharide sheaths enveloping *Lyngbya* filaments provide a haven for heterotrophic bacteria (Simmons *et al.*, 2007), and these may also contribute to the metabolites reported from field collections of

this genus. Indeed, some of the natural products isolated from *Lyngbya* specimens structurally resemble those of heterotrophic bacteria (Graber & Gerwick, 1998; Burja *et al.*, 2001). Thus, due to the microbial complexity of *Lyngbya* colonies, the true metabolic origin of secondary metabolites isolated from *Lyngbya* populations remains uncertain.

An excellent example of the extraordinary secondary metabolite diversity found in marine *Lyngbya* derives from the Caribbean island of Curaçao. Despite its rather small coast line (total land area: 444 km²), *Lyngbya* populations from various sites around the leeward half of the island (<40 coast line miles) have yielded at least sixteen novel molecules (Orjala *et al.*, 1995; Orjala & Gerwick, 1997; Wu *et al.*, 1997; Graber & Gerwick, 1998; Wu *et al.*, 2000). Considering that the shallow-water margin along most of Curaçao's perimeter is narrow, the microenvironment found there may well represent the greatest *Lyngbya* chemo-diversity examined to date.

In the current study, cyanobacteria corresponding with the *Lyngbya* morpho-type (herein referred to as “*Lyngbya*”) were sampled from diverse environments around Curaçao. Phylogenetic inferences of the conserved SSU (16S) rRNA genes were included to give a more complete understanding of the relatedness among these populations and insight into how secondary metabolite production was distributed between different lineages. Genetic and chemical approaches were combined to thoroughly explore the biosynthetic capacity of different samples, and comparison between the biosynthetic capacity versus the actual biosynthetic production of discrete indicator natural products allowed for hypotheses concerning variations arising from differential regulation of gene expression. Additionally, specimens from diverse habitats were compared to evaluate if environmental influences might impact the observed distribution of these secondary metabolites. Finally, we also examined at the cellular level whether the secondary metabolites ascribed to *Lyngbya* are truly produced by this cyanobacterium, or if they are produced by microorganisms associated with its surfaces and sheath material. The data from these various analyses, including gene regulation, metabolic origin and taxonomic revision, were synthesized and integrated into a model that provides a better understanding of the underlying mechanisms behind secondary metabolite diversity in marine cyanobacteria.

Results

Sampling and taxonomic identification

A total of 12 “*Lyngbya*” populations were obtained from a variety of environments ranging from a harbor/dock habitat, a sandy beach, an exposed reef and a salt water mangrove (Fig. 1). All 12 specimens were initially identified as species of “*Lyngbya*” based on the following traditional morphological criteria: (i) filamentous (width >6 µm), (ii) lack of specialized cells (e.g. no heterocysts or akinetes), (iii) isopolar trichomes, (iv) discoid cells and (v) presence of sheaths (Table S2). Six of the specimens were identified as “*L. majuscula*” and the other six as “*L. sordida*” based on the cell width to cell length dimensions and the amount of cross-wall constriction between the cells (Table S2).

Distribution of expressed secondary metabolites

A combination of MALDI-TOF-MS of live filaments and LC ESI-MS of crude extracts was used to screen each specimen for known secondary metabolites (Table 1). Identities of proposed metabolites were confirmed by a combination of ESI-MS², LC retention times (RT) and compared against these features for secondary metabolites of reference strains PNG05-4 (tumonoic acid producing) and 3L (curacin/barbamide/carmabin producing) (Supplementary Table S3).

A total of six of the twelve “*Lyngbya*” strains (NAC8-47, NAC8-48, NAC8-49, NAC8-51, NAC8-52, and NAC8-53) contained a suite of bioactive secondary metabolites composed of the crustacean toxins curacin A, curazole, and curacin D, the molluscicidal agents barbamide and dechlorobarbamide, and the lipopeptides carmabins A-B. Additionally, four of the remaining six “*Lyngbya*” strains (NAC8-45, NAC8-46, NAC8-54 and NAC8-55) produced tumonoic acids A-C and F as well as the methyl and ethyl esters, methyl tumonoates A, B and ethyl tumonoate A. No secondary metabolites cataloged in the MarinLit (2010) database were detected from either of the remaining two strains (NAC8-18 and NAC8-50).

Genetic capacity for secondary metabolite production

All “*Lyngbya*” specimens were screened by PCR for genetic markers of specific secondary metabolite pathways in order to verify their biosynthetic capacity to produce these molecules (Table 1). The following biosynthetic gene markers were analyzed from genomic-DNA: (i) *phyH* encoding the α -ketoglutarate dependent halogenase of the curacin A pathway, (ii) *barB1* and *barB2* encoding the BarB1 and BarB2 halogenases of the barbamide pathway, (iii) the tandem methyl transferases (*carMT1*–*MT2*) methylating the *N,O*-dimethyl tyrosine of carmabin A, and (iv) the thioesterase (*carTE*) of the carmabin A pathway. In one case, the DNA extracted from a single filament of “*L. majuscula*” NAC8-47 was amplified by multiple displacement amplification (MDA) prior to PCR-screening, due to the limited amount of biomass obtained from the field collection.

This PCR-screening and sequencing effort revealed that all six of the curacins/barbamide/carmabins-producing specimens contained the three pathways listed above (*phyH* [curacin A], *barB1* and *barB2* [barbamide], and *carMT1*–*MT2* and *carTE* [carmabin A]), and that these were identical or nearly identical to sequences present in the reference strain 3L (p-distance = <99%). The MDA-DNA of “*L. majuscula*” NAC8-47 contained genes from the barbamide pathway and the carmabin A pathway, but not the curacin pathway.¹ None of the remaining specimens contained any of these biosynthetic pathways.

¹ MDA typically amplifies only *ca.* 50–70% of a genome (Lasken, 2007).

Phylogenetic inference of “*Lyngbya*” strains

Phylogenetic inference of the 16S rRNA genes (1372 bp; ~95% of the gene coverage) revealed that the 12 specimens formed three distinct and distantly related lineages (Fig. 2). All five curacins/barbamide/carmabins producing specimens: NAC8-47, NAC8-48, NAC8-49, NAC8-51, NAC8-52, and NAC8-53 shared >99% sequence identity with other tropical marine “*Lyngbya*” species, including the previously described curacin/barbamide/carmabin producing strain 3L from Curaçao (Rossi *et al.*, 1997).

Despite nesting with other tropical marine “*Lyngbya*” specimens this lineage lacked a type-strain associated with it. The closest related genera are *Symploca* (type-strain = PCC 8002^T; mean p-distance = 6.1%) and *Coleofasciculus* (type-strain = PCC 7420^T; mean p-distance = 6.7%). More importantly, this tropical marine “*Lyngbya*” lineage is evolutionarily highly distant from other *Lyngbya* specimens, including the genus type-strain PCC 7419^T (mean p-distance = 9.4%).

The remaining six specimens (NAC8-18, NAC8-45, NAC8-46, NAC8-50, NAC8-54 and NAC8-55) claded with the morphologically similar genera *Oscillatoria* and *Trichodesmium*, including the type-strains *Oscillatoria sancta* PCC 7515^T and *Trichodesmium erythraeum* IMS 101^T (Fig. 2). The tumonoic acid-producing strains (NAC8-45, NAC8-46, NAC8-50, NAC8-54 and NAC8-55) formed a clade with the Pacific tumonoic acid-producing strain PNG05-4. The strain NAC8-18 was most closely related to the viridamide-producing 3L-O s c from Curaçao and the venturamide-producing PAB-21 *Oscillatoria* spp from the

Caribbean coast of Panama. Interestingly, the strain NAC8-50 grouped with marine species of the planktic genus *Trichodesmium*.

Cellular origin of secondary metabolites and their biosynthetic genes

Scanning electron microscopy (SEM) of cultured filaments from the reference strain “*Lyngbya*” 3L revealed a rich diversity of heterotrophic bacteria on the exterior surface of the polysaccharide sheath (Fig. 3A). Transmission electron microscopy (TEM) of filament sections showed that these bacteria were indeed restricted to these exterior surfaces and that the interior space and cells were free from associated bacteria (Fig. 3B). The enveloping sheath was removed from a filament of strain 3L, two single-cells were individually isolated by micro-manipulation and their genomes amplified by multiple displacement amplification (MDA). Multiple copies (>10) of the 16S rRNA gene were PCR-amplified using general bacterial primers and sequenced from each single-cell MDA-DNA to verify that the MDA reaction had only amplified cyanobacterial DNA and not that from any associated heterotrophic bacteria. All ten 16S rRNA gene sequences from each single-cell MDA-DNA had 100% gene sequence identity with either of the two 16S rRNA gene sequences from “*Lyngbya*” 3L.2 Furthermore, the DNA from one single-cell revealed the presence of the biosynthetic genes *phyH*, *barB1*, *barB2*, *carMT1*–*MT2* and *TE* while the other single-cell MDA DNA contained only *barB1* and *barB2*.

²Tropical marine *Lyngbya* genomes have been found to often contain two different ribosomal operons with variable 16S rRNA gene sequences possessing up to 1.1% divergence (Engene *et al.*, 2010). Thus, the two 16S rRNA gene copies found in 3L is assumed to belong to paralogous ribosomal operons (*i.e.* *rrn* A and *rrn* B).

Intact cell MALDI-TOF-MS (ICM) was also performed on additional isolates of single cells from “*Lyngbya*” 3L, and revealed metabolites with molecular weights of m/z 374.2 $[M+H]^+$, m/z 460.9 $[M+H]^+$ and m/z 726.4 $[M+H]^+$, corresponding to curacin A, barbamide and carmabin A, respectively.

Discussion

Cyanobacteria represent one of the most ancient and biologically diverse groups of organisms on earth (Komárek, 2005; Rasmussen *et al.*, 2008). Thus, the extraordinary diversity of secondary metabolites is likely a reflection of this biodiversity. Yet, despite this diversity, over 90% of marine cyanobacterial-derived secondary metabolites have been isolated from a total of ten different genera (MarinLit, 2010). This imbalance in secondary metabolite distribution is best exemplified with the genus *Lyngbya*, which according to our current understanding, is responsible for the production of over 40% of all marine cyanobacterial secondary metabolites (Tidgewell *et al.*, 2010). The extreme secondary metabolite diversity of marine *Lyngbya* was exemplified by the populations surveyed in this study from the island of Curaçao.

Although twelve different secondary metabolites were detected in various “*Lyngbya*” populations surveyed in this study, the majority of the molecules previously isolated from marine “*Lyngbya*” from Curaçao were not identified in any of the current specimens. Previously isolated molecules that were not observed in any of the current specimens included malyangamides H-L (Wu *et al.*, 1997), antillatoxin A (Orjala *et al.*, 1995), quinones A and B (Orjala & Gerwick, 1997), kalkipyron (Graber & Gerwick, 1998) and kalkitoxin (Wu *et al.*, 2000). Thus, the true metabolic origin of the molecules that were not detected in this study can only be speculated upon. For example, a close structural resemblance has been noted between kalkipyron and secondary metabolites of heterotrophic bacteria (Graber & Gerwick, 1998), and thus, their production from ephemeral populations of

associated heterotrophic bacteria may explain their variable isolation from environmental samples of cyanobacteria.

However, for those natural products detected in the current study, we were able to demonstrate both the presence of the biosynthetic pathways and the detection of secondary metabolites from single cells of “*Lyngbya*”, thereby firmly establishing that these are products of cyanobacterial genes and biochemical processes. Multiple displacement amplification (MDA) and sequencing analysis revealed that the curacin A, barbamide and carmabin A biosynthetic pathways were all present in the cyanobacterial genomes. This was matched by analysis of intact single-cells (ICA) by MALDI-TOF-MS which demonstrated that these secondary metabolites were physically present in the cyanobacterial cells. To our knowledge, this is the first direct combined genetic and chemical proof of secondary metabolite biosynthesis on a microbial single-cell level.

Moreover, this correlation between biosynthetic capacity and biosynthetic expression was a prominent feature, despite sampling from diverse and variable environments. In addition to the conserved expression levels found in different environments, all secondary metabolites were produced in subsequent culture conditions (Table 1). Thus, in this model, the surrounding environments appear to have little impact on the genetic expression of these secondary metabolites and consequently, in the distribution of secondary metabolites among different “*Lyngbya*” populations.

The initial classification of the various “*Lyngbya*” species in this study was based solely on traditional morphology-based criteria, as this has been the predominant foundation for identification of secondary metabolite producing cyanobacteria to date. However, the 16S rRNA phylogenetic analysis revealed that these specimens formed distinct and evolutionarily distant lineages. A corresponding polyphyletic grouping has previously been described for the genus “*Lyngbya*” where these different “*Lyngbya*” lineages were assumed to represent species within the same genus (Sharp *et al.*, 2009). Because the 16S rRNA gene is relatively slowly evolving, an uncorrected genetic divergence of approximate ten percent corresponds to a relatively long period of evolution (Wilmotte & Herdman, 2001). In further support of the evolutionarily distance between these different “*Lyngbya*” lineages is the fact that heterocystous-forming cyanobacteria of the order Nostocales nests between the different lineages. This branching point represents an evolutionary event that has been estimated to have occurred between 2,450 and 2,100 mega-annums ago (Tomitani *et al.*, 2006). Thus, these lineages are clearly evolutionarily distinct and need to be considered as different generic entities rather than species of the same genus. The fact that these unrelated lineages possess highly comparable morphological features likely results from convergent evolution, perhaps as a result of their occupying similar ecological niches.

An alternative hypothesis to homoplasy may simply be limitations in the number and degree of distinguishing characters that form the basis for our traditional morphology-based taxonomic systems. Cyanobacterial taxonomy has been predominantly founded upon temperate soil or freshwater specimens. Therefore, strains from recently explored environments, such as tropical marine, have typically been identified based on their morphological similarities with previously described taxa.

In this study, we show that none of the “*Lyngbya*” specimens were related to the genus type-strain. Instead, some of the specimens were related to the morphologically similar genera *Oscillatoria* and *Trichodesmium*, and should be grouped with these taxonomic entities. These findings underscore the need to reevaluate the morphological characters differentiating these other genera from *Lyngbya*. By contrast, the curacin/barbamide/carmabins-producing “*Lyngbya*” specimens formed an evolutionarily distant lineage without

any related type-strain. This lineage clearly represents a novel cyanobacterial group that has been positioned with the genus *Lyngbya* solely because of morphological similarities. However, this lineage is phylogenetically distinguished from the genus *Lyngbya* and needs to be described as a new generic entity.

The different evolutionarily paths of these lineages have resulted in different metabolic capacities as is shown in their produced secondary metabolites. By grouping morphologically similar but evolutionarily distant specimens together, these groups have become extensively overrepresented in their perceived chemical richness, and this explains the imbalance in secondary metabolite distribution ascribed above. In this study, we show that polyphyly is the major reason for the misconception that *Lyngbya* is such a secondary metabolite-rich group. Ongoing analyses of secondary metabolite-producing cyanobacteria corresponding with the “*Lyngbya*” morpho-type from other geographic regions support this hypothesis. Moreover, to our knowledge, no secondary metabolites have been isolated from specimens related to the genus type-strain.

In this study, we showed that “*Lyngbya*” is a polyphyletic group and that bioactive secondary metabolites attributed to “*Lyngbya*” are actually produced by morphologically similar but phylogenetically distant lineages. A corollary to this conclusion is that it is the morphological resemblance of different cyanobacterial groups that has contributed to the perception that *Lyngbya* is so remarkably rich in secondary metabolites. Thus, taxonomic clarification and revision of polyphyletic cyanobacterial lineages is essential for developing an accurate understanding of the distribution of bioactive secondary metabolites, and can be used to direct targeted and more efficient natural products discovery programs in the future.

Experimental Procedures

Sampling and characterization

Cyanobacterial specimens were collected by SCUBA from 12 sites along the leeward coast of Curaçao, Netherlands Antilles (Fig. 1). Morphological identification of the specimens was performed in accordance with modern taxonomic systems (Castenholz *et al.*, 2003; Komárek & Anagnostidis, 2005). Specimens were directly cleaned from meio/macro-fauna under a dissecting scope in seawater filtered through 0.2 µm Acrodisc® Syringe filters for culturing and morphological analysis. Additional biomass (~200 mg) was preserved for genetic analysis in 10 mL RNAlater® (Ambion Inc., Austin, TX, USA) and for chemical analysis in seawater/EtOH (1:1) at -20 °C (Supplementary Table S1). Light microscopy was performed using an Olympus IX51 epifluorescent microscope equipped with an Olympus U-CMAD3 camera. Samples for scanning electron microscopy (SEM) were placed on indium-tin-oxide glass slides that had been coated with 0.1% polyethylenimine to facilitate adhesion. The samples were then fixed in 2.5% glutaraldehyde buffered in 1X PBS for 30 min followed by a secondary fix of 2% osmium tetroxide (OsO₄) for 15 min. Dehydration was done in a graded EtOH series. Samples were dried in a Balzers critical-point dryer with liquid carbon dioxide as the transition fluid and then sputter-coated with gold palladium using a Polaron E5100 SEM Coating. A Hitachi SU6600 Field Emission SEM was used to view the samples. Samples for transmission electron microscopy (TEM) were fixed overnight in a 4% formaldehyde and 1% glutaraldehyde solution (4F:1G) buffered in 1X PBS followed by a secondary fix of 2% OsO₄ for 1–2 h. Dehydration was done in a graded EtOH series. Samples were then embedded in Spurr's resin and left to polymerize for 48 h. Thin sections (70 nm) were obtained using the Reichart Ultracut E ultramicrotome and picked up on 75 mesh copper grids. The grids were subsequently stained with uranyl acetate and Sato lead (Hanaichi *et al.*, 1986). A JEOL 1200EX TEM was used to view and record images of these samples.

Genomic DNA extraction and multiple displacement amplification (MDA)

Cyanobacterial filaments were cleaned and pretreated using TE (10 mM Tris; 0.1M EDTA; 0.5% SDS; 20 $\mu\text{g} \cdot \text{mL}^{-1}$ RNase)/lysozyme (1 $\text{mg} \cdot \text{mL}^{-1}$) at 37 °C for 30 min followed by incubation with proteinase K (0.5 $\text{mg} \cdot \text{mL}^{-1}$) at 50 °C for 1 h. Genomic DNA was extracted using the Wizard® Genomic DNA Purification Kit (Promega Inc, Madison, WI, USA) following the manufacturer's specifications. DNA concentration and purity was measured on a DU® 800 spectrophotometer (Beckman Coulter). Single-cells were isolated using a MM3-All micromanipulator (World Precision Instruments Inc, Fullerton, CA, USA) as previously described (Engene *et al.*, 2010). DNA was amplified from single-cell genomes using the REPLI-g® Mini Kit (Qiagen Inc, Valencia, CA, USA) following the manufacturer's specifications. All MDA reactions were performed in 50 μL reaction volume for 16 h at 30 °C.

Screening of biosynthetic genes

PCR-primers were designed using Primer3Plus (Untergasser *et al.*, 2007) (Table S4). The PCR reaction volumes were 25 μL containing 0.5 μL (~50 ng) of DNA, 2.5 μL of 10 $\times$ PfuUltra IV reaction buffer, 0.5 μL (25 mM) of dNTP mix, 0.5 μL of each primer (10 μM), 0.5 μL of PfuUltra IV fusion HS DNA polymerase and 20.5 μL dH₂O. The PCR reactions were performed in an Eppendorf® Mastercycler® gradient as follows: initial denaturation for 2 min at 95 °C, 25 cycles of amplification: 20 sec at 95 °C, 20 sec at 50 °C and 1.5 min at 72 °C, and final elongation for 3 min at 72 °C. PCR products were purified using a MinElute® PCR Purification Kit (Qiagen) before being subcloned using the Zero Blunt® TOPO® PCR Cloning Kit (Invitrogen) following the manufacturer's specifications. Plasmid DNA was isolated using the QIAprep® Spin Miniprep Kit (Qiagen) and sequenced with M13 primers. Putative biosynthetic genes were compared with the original barbamide (AF516145) and curacin (AY652953) pathways as well as a putative, unpublished biosynthetic pathway for carmabin obtained from a partly sequenced genome of "*L. sordida*" strain 3L (only sequences with >99% sequence identity were accepted). The gene sequences are available in the DDBJ/EMBL/GenBank databases under acc. Nr. GU724195-GU724208.

Phylogenetic inference

Gene sequences were aligned bi-directionally using the L-INS-i algorithm in MAFFT 6.717. A total of 1 378 bp (310 parsimony informative sites) of the 16S rRNA gene were analyzed without data exclusion. The evolutionary distant unicellular cyanobacterium *Gloeobacter violaceus* PCC 7421^T (NC005125) was included as an out-group. Representative type-strains (^T) were selected from *Bergey's Manual* (Castenholz *et al.*, 2001). Pair-wise sequence divergences were calculated in PAUP* 4.0b10. Appropriate nucleotide substitution models were compared and selected using uncorrected/corrected Akaike Information Criterion (AIC/AICc), Bayesian Information Criterion (BIC), and the Decision-theoretic (DT) in jModelTest 0.1.1. The GTR+I+G model was selected by AIC/AICc/BIC/DT criteria assuming a heterogeneous substitution rates and gamma substitution of variable sites (proportion of invariable sites (pINV) = 0.292, shape parameter (α) = 0.316, number of rate categories = 4). The Maximum likelihood (ML) inference was performed using GARLI 1.0 for the GTR+I+G model with 1 000 bootstrap re-sampling. Bayesian analysis was conducted using MrBayes 3.1. Four Metropolis-coupled MCMC chains (one cold and three heated) was run for 10 000 000 generations. MCMC convergence was determined using AWTY and the first 1 000 000 generations (10%) were discarded as burn-in and the following data set were being sampled with a frequency of every 1 000 generations. The maximum parsimony (MP) analysis was performed in PAUP* 4.0b10 using a heuristic search through the branch-swapping tree-bisection-reconnection (TBR) algorithm with the addition of 10 000 random replicates to find the most parsimonious tree. Bootstrap support was obtained from 1 000 replicates.

Secondary metabolite detection

Algal biomass (~1 g) of each specimen was exhaustively extracted with CH₂Cl₂-MeOH (2:1). The extract was dried under vacuum and the dried residues were re-dissolved in MeOH at a concentration of 1 mg · mL⁻¹. Each sample (10 µL) was injected into an LC ESI-MS system (LCQ Advantage Max spectrometer and UV-profiles by Surveyor PDA plus detector, Thermo Finnigan) and separated on an RP HPLC column (HP Lichrosphere 100 RP-18, 4 × 125 mm, 5.0 µm) with step gradient elution of 0.1% formic acid in water (eluent A) and 100% ACN (eluent B). Gradient program: 0–5 min, B, 45%; 5–55 min, B, 45–100%; 55–65 min, B, 100%; flow rate, 700 µL · min⁻¹. The column temperature was kept at 30 °C. The MS and MS² spectra and retention time of each peak was recorded using the positive ion detection mode. For MALDI-TOF-MS, 5–10 µg (wet wt) of each specimen were extracted with 1 µL · µg⁻¹ matrix solution (70 mg · mL⁻¹ alpha-cyano-4-hydroxycinnamic acid and 2,5-dihydroxybenzoic acid (1:1), 750 µL acetonitrile, 248 µL dH₂O, 2 µL trifluoroacetic acid) in a 96-well plastic plate for 20–30 sec. One µL of matrix extract was deposited on a well of a Bruker Microflex MSP 96 Stainless Steel Target Plate and run on a Bruker microflexTM mass spectrometer equipped with flexControl 3.0. Identification of secondary metabolites required support of predicted isotope patterns, corresponding MS² fragmentations and conserved retention times (RT).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We gratefully acknowledge the government of Curaçao and the Carmabi research station for support and permits for this research project. We also thank M Vermeij and JK Nunnery for aid with the collection as well as RV Grindberg for help selecting genetic markers for the biosynthetic pathways. We would also like to express appreciation for the guidance and advice provided by T Deerinck with regard to electron microscopy studies. The research was generously supported by the Halliday award (SIO) and NIH CA CA100851 and Sea Grant SG-100-TECH-N. Facilities used at the National Center for Microscopy and Imaging Research were supported by NIH through grant number P41-RR004050 to MHE.

References

- Burja AM, Banaigs B, Abou-Mansour E, Burgess JG, Wright PC. Marine cyanobacteria-a prolific source of natural products. *Tetrahedron*. 2001; 46:9347–9377.
- Capper A, Cruz-Rivera E, Paul VJ, Tibbetts IR. Chemical deterrence of a marine cyanobacterium against sympatric and non-sympatric consumers. *Hydrobiologia*. 2006; 553:319–326.
- Castenholz, RW.; Phylum, BX. *Bergey's Manual of Systematic Bacteriology*. New York: Springer; 2001. Cyanobacteria oxygenic photosynthetic bacteria; p. 473-553.
- Chang Z, Flatt PM, Gerwick WH, Nguyen V-A, Willis CL, Sherman DH. The barbamide biosynthetic gene cluster: a novel marine cyanobacterial system of mixed polyketide synthase (PKS)-non-ribosomal peptide synthetase (NRPS) origin involving an unusual trichloroleucyl starter unit. *Gene*. 2002; 296:235–247. [PubMed: 12383521]
- Chang Z, Sitachitta N, Rossi JV, Roberts MA, Flatt PM, Jia J, Sherman DH, Gerwick WH. Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*. *J Nat Prod*. 2004; 67:1356–1367. [PubMed: 15332855]
- Engene N, Coates RC, Gerwick WH. 16S rRNA gene heterogeneity in the filamentous marine cyanobacterial genus *Lyngbya*. *J Phycol*. 2010; 46:591–601.
- Erhard M, von Döhren H, Jungblut P. Rapid typing and elucidation of new secondary metabolites of intact cyanobacteria using MALDI-TOF mass spectrometry. *Nat Biotechnol*. 1997; 15:906–909. [PubMed: 9306409]

- Gerwick WH, Proteau PJ, Nagle DG, Hamel E, Blokhin A, Slate DL. Structure of curacin A, a novel antimitotic, antiproliferative and brine shrimp toxic natural product from the marine cyanobacterium *Lyngbya majuscula*. *J Org Chem*. 1994; 59:1243–1245.
- Gerwick WH, Coates RC, Engene N, Gerwick LG, Grindberg R, Jones A, Sorrels C. Giant marine cyanobacteria produce exciting potential pharmaceuticals. *Microbe*. 2008; 3:277–284.
- Graber MA, Gerwick WH. Kalkipyron, a toxic γ -pyrone from an assemblage of the marine cyanobacteria *Lyngbya majuscula* and *Tolypothrix* sp. *J Nat Prod*. 1998; 61:677–680. [PubMed: 9599278]
- Hanaichi, T.; Sato, T.; Hoshino, M.; Mizuno, N. Proceedings XIth Int; Kyoto: Congress on Electron Microscopy; 1986.
- Hoffman L, Komárek J, Kaštovský J. System of Cyanoprokaryotes. (Cyanobacteria) - State in 2004. *Arch Hydrobiol*. 2005; 159:95–115.
- Kaebernick M, Neilan BA, Borner T, Dittmann E. Light and the transcriptional response of the microcystin biosynthesis gene cluster. *Appl Environ Microbiol*. 2000; 66:3387–3392. [PubMed: 10919796]
- Komárek, J.; Anagnostidis, K. Süßwasserflora von Mitteleuropa 19/2. Heidelberg: Elsevier/Spektrum; 2005.
- Komárek J. Cyanobacterial taxonomy: current problems and prospects for the integration of traditional and molecular approaches. *Algae*. 2006; 21:349–375.
- Lasken RS. Single-cell genomic sequencing using multiple displacement amplification. *Curr Opin Microbiol*. 2007; 10:510–516. [PubMed: 17923430]
- Marine literature database. (<http://www.chem.canterbury.ac.nz/marinlit/marinlit.shtml>)
- Nübel U, Garcia-Pichel F, Muyzer G. PCR primers to amplify 16S rRNA genes from cyanobacteria. *Appl Environ Microbiol*. 1997; 63:3327–3332. [PubMed: 9251225]
- Orjala J, Gerwick WH. Barbamide, a chlorinated metabolite with molluscicidal activity from the caribbean cyanobacterium *Lyngbya majuscula*. *J Nat Prod*. 1994; 59:427–430. [PubMed: 8699186]
- Orjala J, Gerwick WH. Two quinoline alkaloids from the Caribbean cyanobacterium *Lyngbya majuscula*. *Phytochemistry*. 1997; 45:1087–1090. [PubMed: 9214778]
- Orjala J, Nagle DG, Hsu VL, Gerwick WH. Antillatoxin: an exceptionally ichthyotoxic cyclic lipopeptide from the tropical cyanobacterium *Lyngbya majuscula*. *J Am Chem Soc*. 1995; 117:8281–8282.
- Rasmussen B, Fletcher IR, Brocks JJ, Kilburn MR. Reassessing the first appearance of eukaryotes and cyanobacteria. *Nat Lett*. 2008; 455:1101–1104.
- Rossi JV, Roberts MA, Yoo H-D, Gerwick WH. Pilot scale culture of the marine cyanobacterium *Lyngbya majuscula* for its pharmaceutically-useful natural metabolite curacin A. *J Appl Phycol*. 1997; 9:195–204.
- Sharp K, Arthur KE, Gu L, Ross C, Harrison G, Gunasekera SP, Meickle T, Matthew S, Luesch H, Thacker RW, Sherman DH, Paul VJ. Phylogenetic and chemical diversity of three chemotypes of bloom-forming *Lyngbya* species (Cyanobacteria: Oscillatoriales) from reefs of southeastern Florida. *Appl Environ Microbiol*. 2009; 75:2879–2888. [PubMed: 19270119]
- Simmons TL, Coates RC, Clark BR, Engene N, Gonzalez D, Esquenazi E, Dorrestein PC, Gerwick WH. Biosynthetic origin of natural products isolated from marine microorganism-invertebrate assemblages. *Proc Natl Acad Sci U S A*. 2007; 105:4587–4594. [PubMed: 18250337]
- Simmons TL, Engene N, Urena LD, Romero LI, Ortega-Barria E, Gerwick GL, Gerwick WH. Viridamides A and B, lipodepsipeptides with antiprotozoal activity from the marine cyanobacterium *Oscillatoria nigro-viridis*. *J Nat Prod*. 2008; 71:1544–1550. [PubMed: 18715036]
- Shalev-Malul G, Lieman-Hurwitz J, Viner-Mozzini Y, Sukenik A, Gaathon A, Lebendiker M, Kaplan A. An AbrB-like protein might be involved in the regulation of cylindrospermopsin production by *Aphanizomenon ovalisporum*. *Environ Microbiol*. 2008; 10:988–999. [PubMed: 18093160]
- Sorrels CM, Proteau PJ, Gerwick WH. Organization, evolution, and expression analysis of the biosynthetic gene cluster for scytonemin, a cyanobacterial UV-absorbing pigment. *Appl Environ Microbiol*. 2009; 75:4861–4869. [PubMed: 19482954]

- Tan LT. Bioactive natural products from marine cyanobacteria for drug discovery. *Phytochem.* 2007; 68:954–979.
- Taori K, Liu Y, Paul VJ, Luesch H. Combinatorial strategies by marine cyanobacteria: symplostatin 4, an antimitotic natural dolastatin 10/15 hybrid that synergizes with the coproduced HDAC inhibitor largazole. *ChemBioChem.* 2009; 10:1634–1639. [PubMed: 19514039]
- Tidgewell, K.; Clark, BR.; Gerwick, WH. *Compr Nat Prod Chem.* Oxford: Pergamon Press; 2010. The natural products chemistry of cyanobacteria.
- Thacker RW, Nagle DG, Paul VJ. Effects of repeated exposures to marine cyanobacterial secondary metabolites on feeding by juvenile rabbitfish and parrotfish. *Mar Ecol Prog Ser.* 1997; 147:21–29.
- Thacker RW, Paul VJ. Morphological, chemical, and genetic diversity of tropical marine cyanobacteria *Lyngbya* spp. and *Symploca* spp. (Oscillatoriales). *Appl Environ Microbiol.* 2004; 70:3305–3312. [PubMed: 15184125]
- Tomitani A, Knoll AK, Cavanaugh CM, Ohno T. The evolutionary diversification of cyanobacteria: Molecular–phylogenetic and paleontological perspectives. *Proc Natl Acad Sci U S A.* 2006; 103:5442–5447. [PubMed: 16569695]
- Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM. Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res.* 2007; 35:71–74.
- Willemotte, A.; Herdman, M. *Bergey's Manual of Systematic Bacteriology.* New York: Springer; 2001. p. 487-493.
- Wu M, Milligan KE, Gerwick WH. Three new malyngamides from the marine cyanobacterium *Lyngbya majuscula*. *Tetrahedron.* 1997; 53:15983–15990.
- Wu M, Okino T, Nogle LM, Marquez BL, Williamson RT, Gerwick WH, et al. Structure, synthesis, and biological properties of kalkitoxin, a novel neurotoxin from the marine cyanobacterium *Lyngbya majuscula*. *J Am Chem Soc.* 2000; 122:12041–12042.

Fig. 1.
Geographic map of Curaçao indicating the various collection sites for the current study.
Environmental descriptions of each collection site are available in the Supporting
Information Table S1.

Fig. 2.

Phylogenetic inferences for “*Lyngbya*” specimens from Curaçao based on the SSU (16S) rRNA genes. All specimens are indicated as species, strains and with GenBank accession numbers in brackets. Appropriate type-strains (^T) were obtained from *Bergey’s Manual* for representative genera. Specimens corresponding with *Lyngbya* morphologically, but which are phylogenetically unrelated to the genus type-strain PCC 7419^T are designated as “*Lyngbya*”. Lineages for the morphologically similar genera *Lyngbya*, *Trichodesmium*, and *Oscillatoria* are shown in green, brown, and red boxes, respectively. The “*Lyngbya*” lineage boxed in blue lacks a representative reference strain and needs to be considered a new generic entity. Microphotographs of the different phylotypes are indicated with arrows as well as the secondary metabolites they produce. Support at important nodes are indicated as bootstrap and posterior probability for the Maximum likelihood (GARLI)/Bayesian inference (MrBayes). Well supported nodes (>80% bootstrap and 0.9 posterior probability) are indicated with asterisks (*). The scale bar is equivalent to 0.04 substitutions per nucleotide position.

Fig. 3. (A) Scanning electron micrograph (SEM) of the surface of a cultured “*L. sordida*” 3L filament colonized by heterotrophic bacteria. (B) Transmission electron micrograph (TEM) of a cross section of cultured “*L. sordida*” 3L. Note that heterotrophic bacteria are only visibly present on the exterior surface of the filaments’ polysaccharide sheaths.

Table 1

Detection of biosynthetic pathways and secondary metabolites from “*Lyngbya*” specimens.

	phyH (% Id ^a)	curacin A	curazole	curacin D	barB1 (% Id ^a)	barB2 (% Id ^a)	barbamide	dechloro barbamide	carMT-MT (% Id ^a)	carTE (% Id ^a)	carnabin A	carnabin B
“ <i>L. sordida</i> ” 3L	100	++	+	+	100	100	++	+	100	100	+	–
“ <i>L. sordida</i> ” NAC8-48	100	++	+	+	100	100	++	+	100	100	+	–
“ <i>L. sordida</i> ” NAC8-49	100	++	+	+	99.5	99.8	++	–	99.9	99.9	+	+
“ <i>L. sordida</i> ” NAC8-51	100	++	+	+	99.8	100	++	–	100	100	+	+
“ <i>L. sordida</i> ” NAC8-52	100	+	+	+	100	100	++	–	99.8	99.9	+	+
“ <i>L. sordida</i> ” NAC8-53	99.7	+	+	+	99.8	100	+	–	100	100	+	–
“ <i>L. majuscula</i> ” NAC8-47	ND ^{**}	+	+	NA [*]	99.7 ^{**}	99.8 ^{**}	+	– [*]	100 ^{**}	99.7 ^{**}	+	+

	tumonoic acid			methyl tumonoate		ethyl tumonoate	
	A	B	C	F	A	B	A
“ <i>L. majuscula</i> ” NAC8-45	++	++	++	+	+	++	++
“ <i>L. majuscula</i> ” NAC8-46	++	++	+	+	+	++	++
“ <i>L. majuscula</i> ” NAC8-54	++	++	+	+	+	++	++
“ <i>L. majuscula</i> ” NAC8-55	++	++	++	+	+	++	++
“ <i>L. majuscula</i> ” NAC8-18	–	–	–	–	–	–	–
“ <i>L. majuscula</i> ” NAC8-50	–	–	–	–	–	–	–

Abbreviations: ND - not detected. NA - no data available (*i.e.* due to overlapping matrix ions).

^aPercent sequence identity with the original biosynthetic pathways.

(++) main or major constituent, (+) minor constituent or trace compound estimated based on ESI abundance, (–) not detected.

(*) Secondary metabolites detection based only on MALDI-TOF-MS.

(**) Genes screened using MDA genomes.