

Research article

Open Access

## Transcriptional analysis of the jamaicamide gene cluster from the marine cyanobacterium *Lyngbya majuscula* and identification of possible regulatory proteins

Adam C Jones<sup>1</sup>, Lena Gerwick<sup>1</sup>, David Gonzalez<sup>2,3</sup>, Pieter C Dorrestein<sup>2,3,4,5</sup> and William H Gerwick<sup>\*1,5</sup>

Address: <sup>1</sup>Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093 USA, <sup>2</sup>Department of Chemistry, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093 USA, <sup>3</sup>Department of Biochemistry, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093 USA, <sup>4</sup>Department of Pharmacology, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093 USA and <sup>5</sup>Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093 USA

Email: Adam C Jones - [acjones@ucsd.edu](mailto:acjones@ucsd.edu); Lena Gerwick - [lgerwick@ucsd.edu](mailto:lgerwick@ucsd.edu); David Gonzalez - [djgonzalez@ucsd.edu](mailto:djgonzalez@ucsd.edu); Pieter C Dorrestein - [pdorrestein@ucsd.edu](mailto:pdorrestein@ucsd.edu); William H Gerwick\* - [wgerwick@ucsd.edu](mailto:wgerwick@ucsd.edu)

\* Corresponding author

Published: 1 December 2009

Received: 8 May 2009

BMC Microbiology 2009, 9:247 doi:10.1186/1471-2180-9-247

Accepted: 1 December 2009

This article is available from: <http://www.biomedcentral.com/1471-2180/9/247>

© 2009 Jones et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** The marine cyanobacterium *Lyngbya majuscula* is a prolific producer of bioactive secondary metabolites. Although biosynthetic gene clusters encoding several of these compounds have been identified, little is known about how these clusters of genes are transcribed or regulated, and techniques targeting genetic manipulation in *Lyngbya* strains have not yet been developed. We conducted transcriptional analyses of the jamaicamide gene cluster from a Jamaican strain of *Lyngbya majuscula*, and isolated proteins that could be involved in jamaicamide regulation.

**Results:** An unusually long untranslated leader region of approximately 840 bp is located between the jamaicamide transcription start site (TSS) and gene cluster start codon. All of the intergenic regions between the pathway ORFs were transcribed into RNA in RT-PCR experiments; however, a promoter prediction program indicated the possible presence of promoters in multiple intergenic regions. Because the functionality of these promoters could not be verified *in vivo*, we used a reporter gene assay in *E. coli* to show that several of these intergenic regions, as well as the primary promoter preceding the TSS, are capable of driving  $\beta$ -galactosidase production. A protein pulldown assay was also used to isolate proteins that may regulate the jamaicamide pathway. Pulldown experiments using the intergenic region upstream of *jamA* as a DNA probe isolated two proteins that were identified by LC-MS/MS. By BLAST analysis, one of these had close sequence identity to a regulatory protein in another cyanobacterial species. Protein comparisons suggest a possible correlation between secondary metabolism regulation and light dependent complementary chromatic adaptation. Electromobility shift assays were used to evaluate binding of the recombinant proteins to the jamaicamide promoter region.

**Conclusion:** Insights into natural product regulation in cyanobacteria are of significant value to drug discovery and biotechnology. To our knowledge, this is the first attempt to characterize the transcription and regulation of secondary metabolism in a marine cyanobacterium. If jamaicamide is light regulated, this mechanism would be similar to other cyanobacterial natural product gene clusters such as microcystin LR. These findings could aid in understanding and potentially assisting the management of toxin production by *Lyngbya* in the environment.

## Background

Over the past 30 years, the search for bioactive secondary metabolites (natural products) from marine organisms has yielded a wealth of new molecules (estimated at ~17,000) with many fundamentally new chemotypes and extraordinary potential for biomedical research and applications [1], and previous references therein]. Marine cyanobacteria continue to be among the most fruitful sources of marine natural products, with nearly 700 compounds described [2,3]. The filamentous marine cyanobacterium *Lyngbya majuscula* (Gomont) is of particular importance, as approximately 35% of all cyanobacterial bioactive compounds have been reported from the genus *Lyngbya*, with 76% of these coming from *L. majuscula* [3]. More recently, compound isolation and structure elucidation from *L. majuscula* has been complemented with the characterization of biosynthetic gene clusters that encode a number of these compounds. The gene clusters for several potent anticancer and neurotoxic agents such as curacin A, barbamide, and the jamaicamides have provided new insight into the biosynthetic strategies and logic used by this organism for compound production, as well as unique enzymes involved in unprecedented molecular tailoring reactions [4-7].

Despite considerable interest in pursuing cyanobacterial lead compounds as potential drug candidates, an adequate supply of these compounds for clinical research is often impossible to obtain without impractically large scale field collections or sophisticated and expensive synthetic methods [8,9]. With some notable examples [10-13] it has been difficult to induce microbial gene clusters to produce their natural products in heterologous hosts, and thus this technology is not currently predictable [14]. Equally problematic, filamentous marine cyanobacteria such as *Lyngbya* grow slowly in laboratory culture, with doubling times in some cases of about 18 days [15].

One avenue for increasing compound production from marine cyanobacteria could be to take advantage of regulatory elements associated with a biosynthetic gene cluster of interest. Although genetic controls of several primary metabolic functions in cyanobacteria including circadian rhythms [16], heterocyst development [17], and nutrient uptake [18] have been described, information regarding transcriptional regulation of cyanobacterial secondary metabolites is currently limited to freshwater toxins such as the microcystins. The microcystins are potent hepatotoxins synthesized by several freshwater cyanobacteria of worldwide occurrence [19] and are generated via a mixed polyketide synthase/non-ribosomal peptide synthetase (PKS/NRPS) gene cluster [20]. Expression of the microcystin gene cluster is positively correlated with increased light intensity and red light in particular [21]. Moreover, the gene cluster has different transcription start sites depend-

ing on light levels [22]. Other environmental factors have been evaluated for their effects on microcystin production, and increasing evidence suggests that iron may be important. Transcription of genes from the microcystin gene cluster increases with iron starvation [23], and in the presence of iron, a ferric uptake regulator (Fur) protein appears to bind to the microcystin bidirectional promoter and may decrease microcystin production [24]. Because it complexes with iron and other metals [25] microcystin may therefore function as a siderophore. Alternatively, microcystin has been proposed to serve in intraspecies communication, where release of the compound is interpreted as cell death by other *Microcystis* sp. and causes increased microcystin production to enhance localized toxicity [26].

As with microcystin, many of the toxins found in *L. majuscula* are also produced by gene clusters comprised of PKS/NRPS architecture. PKS/NRPS gene clusters in other bacteria have been found to include imbedded regulatory proteins, such as the *Streptomyces* Antibiotic Regulatory Proteins (SARPs) found within the confines of several antibiotic pathways in *Streptomyces* [27]. However, cyanobacterial natural product gene clusters identified to date do not contain any apparent associated regulatory proteins.

Insight into the mechanisms used by *L. majuscula* in the transcription of secondary metabolite gene clusters could be of significant value in enhancing the overproduction of potential drug leads in laboratory culture. Increased compound yield would reduce the need and environmental impact of repeated large scale field collections or the time and expense of chemical synthesis. Additionally, because the secondary metabolite biosynthetic gene clusters identified thus far from *L. majuscula* have been from different strains of the same species, transcription of each pathway could be under similar mechanisms of regulation.

This paper provides an analysis of transcriptional regulatory elements associated with the jamaicamide gene cluster from *Lyngbya majuscula*, and to our knowledge is the first such effort for a secondary metabolite gene cluster from a marine cyanobacterium. The jamaicamides are mixed PKS/NRPS neurotoxins that exhibit sodium channel blocking activity and fish toxicity. The molecules contain unusual structural features including a vinyl chloride and alkynyl bromide [6]. The gene cluster encoding jamaicamide biosynthesis is 57 kbp in length, and is composed of 17 ORFs that encode for proteins ranging in length from 80 to 3936 amino acids. Intergenic regions between 5 and 442 bp are located between all but two of the ORFs, and a region of approximately 1700 bp exists between the first jamaicamide ORF (*jamA*, a hexanoyl ACP synthetase) and the closest upstream (5') ORF outside of the cluster (a

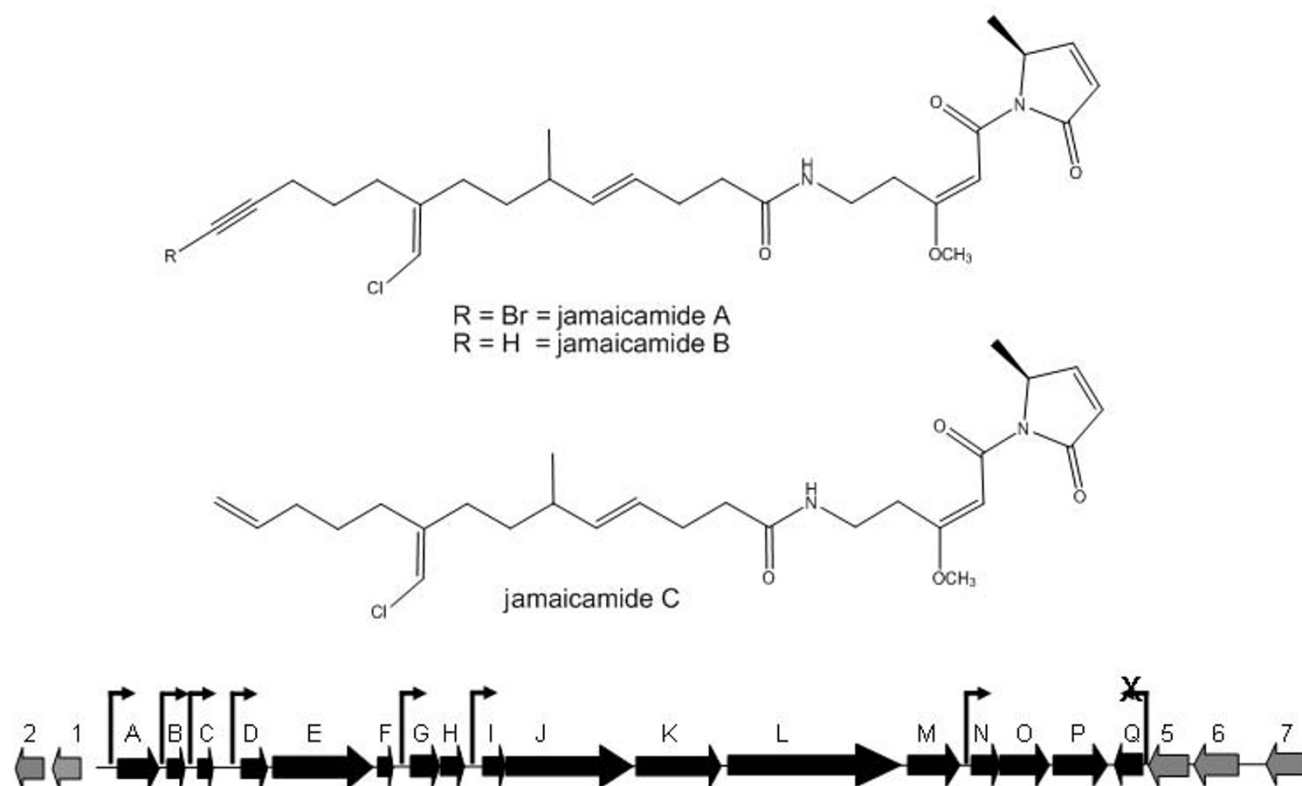
putative transposase). In this study, we used RT-PCR to locate the transcriptional start site (TSS) of the jamaicamide gene cluster. Because it is not yet possible to perform genetics in filamentous marine cyanobacteria such as *Lyngbya*, we used a reporter gene assay to identify several possible internal pathway promoters. We also isolated at least one possible regulatory protein using pulldown experiments that is able to bind to the region upstream of the transcription start site in gel shift assays. Bioinformatic analyses conducted with the protein sequence suggest a correlation between secondary metabolite production and complementary chromatic adaptation (CCA) in cyanobacteria.

## Results

### RT-PCR using *L. majuscula* RNA to search for the transcriptional start site (TSS) and promoter regions in the jamaicamide pathway

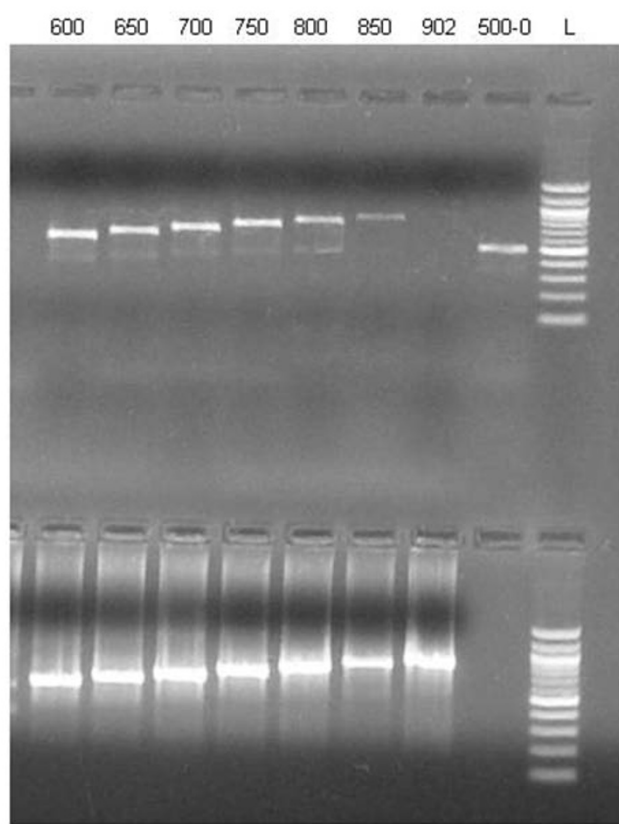
The initial characterization of the jamaicamide gene cluster [6] described that the first 16 ORFs of the gene cluster (*jamA-jamP*) are all transcribed in the forward direction,

while the last ORF (*jamQ*, a putative condensation domain thought to be involved in the cyclization of the pyrrolinone ring of the molecule) is transcribed in the reverse direction (Figure 1). In order to determine the location of the transcriptional start site (TSS) of the gene cluster, RNA was isolated from the jamaicamide producing strain of *Lyngbya majuscula* (JHB). First strand cDNA was synthesized using reverse transcriptase and a reverse primer designed as a complement to the 5' end of the *jamA* gene (Additional file 1: Table S1). Initial experiments creating second strand cDNA using the first strand cDNA as template found that an unusually long untranslated leader region of at least 500 bp preceded *jamA*. A primer extension experiment was conducted in which second strand cDNA was amplified in 50 bp increments beyond this 500 bp location. The experiment indicated that transcription of RNA began between 850 bp and 902 bp upstream of the *jamA* ORF start site (Figure 2). Using comparisons to consensus promoter and transcription start regions in *E. coli* [28-30], a putative promoter was identified which, relative to a probable TSS (844 bp



**Figure 1**

**Structures of the jamaicamides and the jamaicamide biosynthetic gene cluster [6].** Genes associated with the pathway are represented by black arrows, and genes flanking the pathway are represented in gray. Elevated arrows above the upstream regions of selected open reading frames indicate where promoter activity was detected using the  $\beta$ -galactosidase reporter assay. The region upstream of *jamQ* did not have any detectable promoter activity in the assay.



**Figure 2**  
**Transcription start site (TSS) primer extension**  
**experiment using first strand cDNA upstream of**  
**jamA (top) or jamA fosmid (bottom) as PCR tem-**  
**plates.** The upstream region sizes (e.g., 600-0, 650-0) are  
 indicated above each lane.

upstream of *jamA*), included conserved hexamer RNA polymerase (RNAP) binding sites at -35 and -10 bp, a conserved extended -10 TGN region upstream of the -10 box, and an optimal DNA length between the hexamers (17 bp) (Figure 3).

We also evaluated whether the jamaicamide gene cluster contained non-transcribed intergenic regions between ORFs that could indicate the presence of breaks in the transcripts. Primers designed for those intergenic regions in the pathway 20 bp in size or larger (all but 2 intergenic regions) were used for synthesizing and amplifying cDNA to determine whether they were transcribed. All of the intergenic regions in the jamaicamide pathway tested were amplified into second strand cDNA, including the intergenic region between *jamP* and *jamQ*. Intergenic regions between the two ORFs downstream of *jamQ* (putative transposases) were also transcribed. These results indicated that the majority of the jamaicamide gene cluster is composed of the operon *jamABCDEFGHI-*

*JKLMNOP*. Because no apparent breaks in transcription occurred between *jamQ* and at least the two neighboring downstream transposases (ORF5 and ORF6) and a hypothetical protein (ORF7), one contiguous transcript may encode the translation of all of these proteins. Transcription of the intergenic region between *jamP* and *jamQ* indicated that a transcript including *jamP* must extend at least into the complementary strand of *jamQ* before termination, although transcription in the opposite direction would be necessary to generate *jamQ* mRNA.

#### **Use of promoter prediction and -galactosidase reporter gene assays to search for promoter activity**

The large size (approximately 55 kbp) of the main jamaicamide operon (*jamA-P*) suggested that multiple promoters would likely be needed for efficient jamaicamide transcription. Because transcripts were found for each of the intergenic regions between the ORFs, these promoters may function intermittently and could be subject to promoter occlusion [22]. A software prediction program (BPROM, <http://www.softberry.com>) was used to predict whether the intergenic regions from the jamaicamide pathway contained conserved promoter binding regions. Several of these regions were predicted to contain at least one potential pair of -35 and -10 binding sites (Table 1). Because transformation methods into *L. majuscula* have not yet been developed, we used a reporter gene assay in *E. coli* to determine whether any of these upstream (up-) regions could function as promoters. Each region predicted to contain a promoter (*upjamA*, *upjamB*, *upjamC*, *upjamD*, *upjamG*, *upjamI*, *upjamN*, and *upjamQ*), as well as the promoter upstream of the jamaicamide TSS, was amplified with specific primers from fosmids containing different portions of the jamaicamide biosynthetic pathway ([6]; Additional file 1: Table S1). Each of these regions were individually ligated into the pBLUE TOPO vector (Invitrogen) and transformed into TOP-10 *E. coli*. The resulting constructs were evaluated for relative promoter activity using the  $\beta$ -galactosidase reporter gene assay (Invitrogen), standardized against total soluble protein content measured by BCA assay (Pierce). For *upjamA*, two regions were evaluated, including the region predicted to contain the initial promoter, as well as immediately upstream of the *jamA* gene (a region with high activity in preliminary assays). The arabinose promoter from *E. coli* was amplified from the pBAD vector (Invitrogen) and ligated into the pBLUE vector as a positive control, while a 49 bp segment of a jamaicamide pathway gene (*jamG*) ligated into pBLUE vector was used as a negative control.

Several of the tested intergenic regions exhibited significantly stronger promoter activity than the positive control, including the promoter identified from the primer extension experiment (*upjamA*-902 - -832 bp), as well as

**Figure 3**  
**Location of identified promoter regions and transcription start site (TSS) upstream of *jamA*.** The consensus -35 and -10 boxes of each region are underlined. The conserved extended -10 TGn box of the primary pathway promoter is double underlined. The putative TSS is noted at +1, and was chosen based on similarities to the consensus *E. coli* TSS nucleotide region [29]. The first four codons of the *jamA* gene are noted at the end of the sequence.

To more precisely localize the promoter regions upstream of two of these genes, a series of additional assays were conducted using truncated regions of *upjamA* (immediately upstream of the *jamA* gene) and *upjamI*. For *upjamA*, promoter activity was comparable relative to the positive control when testing longer upstream fragments (-500 - 0 and -200 - 0 bp; Figure 5). However, when small fragments closer to the *jamA* ORF start site were used, the promoter activity increased significantly, with maximal

To quantitatively determine the promoter activities of the DNA fragments, a series of  $\beta$ -galactosidase assays incorporating a serial dilution of *E. coli* soluble protein lysate was



**Table 1: Predicted -35 and -10 promoter regions (bold) and transcription start sites (TSS; nucleotides in bold italics at the end of each sequence) for intergenic regions in the jamaicamide gene cluster (accession #AY522504).**

Upstream region of gene	Predicted TSS location (bp)	ORF start (bp)	
<b>upjamA</b>	6626	6630	CTGACT <b>TTTCC</b> ACGACATGGGACTGATGGGAAATGT <b><i>TAT</i></b> TTTAT TTGA
<b>upjamB</b>	8464	8591	GTGGG <b>TTG</b> ATTGATCAAGTTTGATGAT <b><i>TATA</i></b> TTTGATTTA
<b>upjamB</b>	8501	8591	TTTAAT <b>TTTAC</b> AGGGATACCGCCAATTCGGT <b><i>TA</i></b> CTGGAAAA
<b>upjamC</b>	9614	9718	AAAAC <b>TTGT</b> CAACCTGAACAAGATCCTGAA <b><i>CA</i></b> AAATATTGTTG
<b>upjamD</b>	10433	10463	ACAGT <b>TTG</b> ATGGTGCCGCTATTTTGAGTTGG <b><i>AA</i></b> AAATTTTTTA
<b>upjamG</b>	18145	18222	ATTTG <b>TTG</b> TTTGGGAATCGGGAATTGGT <b><i>TAT</i></b> AGTAGTGGAA
<b>upjamI</b>	20776	20982	CGGAAT <b>TTCA</b> AAATTCAAATTCAAAATGCTTATGG <b><i>ATT</i></b> ATGGAA GTAAA
<b>upjamI</b>	20989	20982	CCAGG <b>TTG</b> ACAACCATGATAAAGCT <b><i>TAT</i></b> AGTATGTATTA
<b>upjamN</b>	51787	51811	TGGAG <b>TATA</b> AAAAAACAGAGCCTGGT <b><i>GAT</i></b> AGTTAATTAA
<b>upjamQ</b>	63710 <sup>a</sup>	63646 <sup>a</sup>	GAACT <b>TTGA</b> ATCCTCTATTTTGATT <b><i>TA</i></b> AATTGGAGA

<sup>a</sup>: Numbers correspond to bp in complementary 3' - 5' direction.

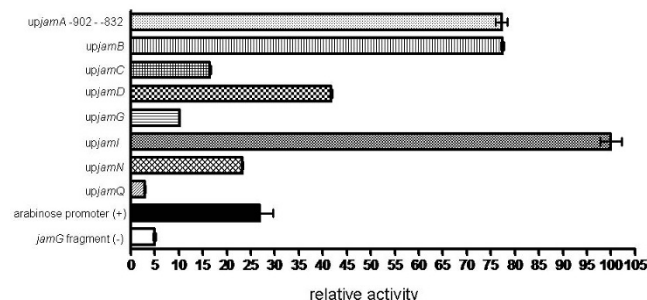
Bold -35 and -10 binding regions were predicted using BPROM (softberry.com) in comparison to the *E. coli*  $\sigma^{70}$  consensus -35 (TTGACA) and -10 (TATAAT) promoter regions. If upstream regions had more than one predicted promoter region, the region receiving the highest predicted score is provided (with the exception of *upjamB*, which had two high scoring regions, and *upjamI*, in which both predicted promoters were tested in the  $\beta$ -galactosidase assay). Each upstream region listed in the first column had activity in the reporter assay, except those not shown in bold text. The italic ATG in the second *upjamI* predicted promoter region indicates the *jaml* start codon. TSS nucleotides were predicted to be A or G based on comparisons to the most common TSS nucleotides in *E. coli*[29].

also used in order to avoid saturation problems in color development (Figure 6). These data were used to calculate  $\beta$ -galactosidase activity in terms of nmol ONPG hydrolyzed min<sup>-1</sup> mg soluble protein<sup>-1</sup> for each of the upstream fragments with any detectable promoter activity. The strongest promoter was the section upstream of the jamaicamide TSS (-902 - -832 upstream of *jamA*), with an average of approximately 950 nmol ONPG hydrolyzed min<sup>-1</sup> mg soluble protein<sup>-1</sup>. The promoter immediately upstream of *jamA* (-76 - 0) and those upstream of *jamB*, *jamD*, and *jaml* yielded lower values, with *upjamA*, *upjamB* and *upjamI* between 500-700 nmol ONPG hydrolyzed min<sup>-1</sup> mg soluble protein<sup>-1</sup>, and *upjamD* at approximately 265 nmol ONPG hydrolyzed min<sup>-1</sup> mg soluble

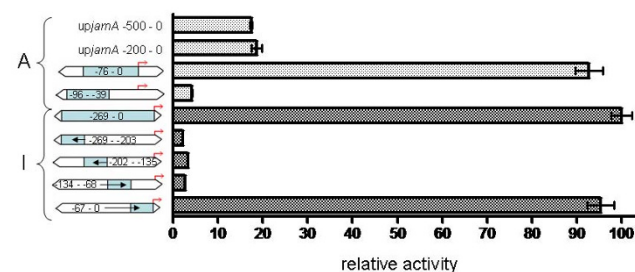
protein<sup>-1</sup>. Reduced activity was found for promoters upstream of *jamC*, *jamG*, and *jamN*, with values ranging from approximately 75 to 150 nmol ONPG hydrolyzed min<sup>-1</sup> mg soluble protein<sup>-1</sup>. The arabinose promoter positive control construct yielded an average value of 170 nmol ONPG hydrolyzed min<sup>-1</sup> mg soluble protein<sup>-1</sup>.

#### Isolation and characterization of possible transcription factors from a pulldown assay

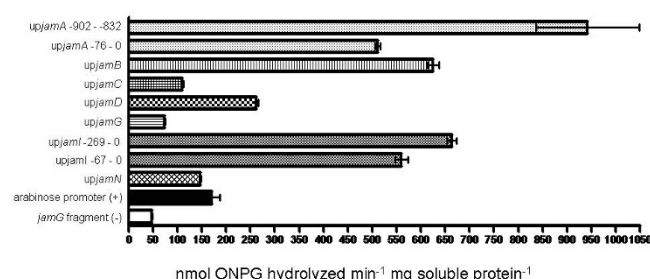
To determine whether jamaicamide regulatory proteins are encoded in the *L. majuscula* JHB genome, we performed DNA - protein "pulldown" experiments to isolate proteins with affinity to the upstream region of *jamA*. A



**Figure 4**  
**Relative activity of the primary promoter upstream of *jamA* and predicted promoters in jamaicamide intergenic regions in the  $\beta$ -galactosidase reporter assay.** Standard error is represented by error bars.

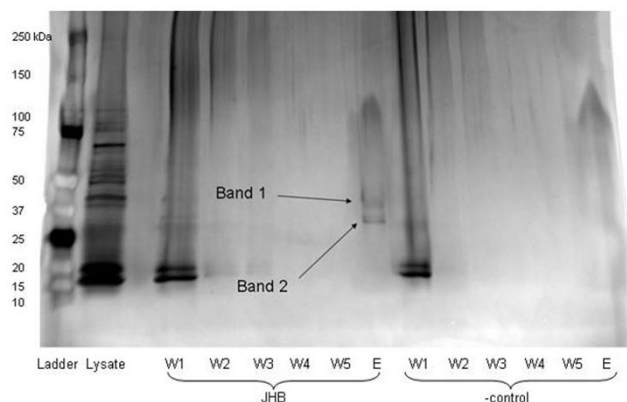


**Figure 5**  
**Activity of truncated *upjamA* and *upjamI* regions in the  $\beta$ -galactosidase assay.** Trimmed regions are represented by blue shaded figures with associated base pair numbers. Red arrows indicate the start codon of the downstream ORF (*jamA* or *jaml*). Relative activity was calculated on same scale as Figure 4. Standard error is represented by error bars.



**Figure 6**  
**Specific activity of the strongest promoters in the  $\beta$ -galactosidase assay.** Base pair number relative to gene ORF start site is provided when necessary. Standard error is represented by error bars.

biotinylated, DNA probe extending from 1000 bp upstream of *jamA* to 20 bp into the *jamA* gene (encompassing both the putative promoter region and entire untranslated leader region; primers in Additional file 1: Table S1) was used to label streptavidin coated magnetic Dynabeads (Invitrogen), which were then incubated with a soluble protein lysate from *L. majuscula* JHB. A series of wash steps were first conducted to remove proteins non-specifically bound, followed by elution of those proteins specifically bound to the probe. This elution was visualized using SDS-PAGE and revealed at least two bands of approximately 30-45 kDa in size (Figure 7). The protein bands from the gel, as well as crude fractions eluted from the magnetic beads in repeated experiments, were analyzed with LC-MS/MS.



**Figure 7**  
**Results from JHB soluble protein pulldown experiment.** From left to right: Ladder, JHB soluble protein lysate, wash fractions (W1 - 5) and elution (E) for incubations with the 1020 bp probe (labeled JHB) or without probe (labeled - control). Note the presence of two bands eluted from the beads containing the probe, indicating successful binding of possible regulatory proteins to the upstream region of *jamA*.

The fragmented peptides generated from the LC-MS/MS analysis of the gel bands were used to query the unfinished *Lyngbya majuscula* 3L genome (a strain from Curaçao that produces several natural products, including barbamide and curacin A) using the MS/MS post-processing program InSpecT [31]. By this approach, two proteins were identified with high confidence from "band 2" (Figure 7), which had a global distribution (N-terminal to C-terminal) among the identified peptides: (i) All4300 protein (39.2% coverage and a molecular weight of 32 kDa), and (ii) hypothetical protein (35.9% coverage and a molecular weight of 33 kDa). Manual annotation of the most abundant peptide identified within the primary sequence of All4300 demonstrated the b and y ion series fell within a mass error of 5-400 ppm. Furthermore, the b and y-ion series for this peptide showed 22/30 possible fragmentations covered with several contingent ion series. The ion series for the hypothetical protein showed similar results to the All4300 protein. Results from the LC-MS/MS of the PAGE gel "band 1" (Figure 7) were inconclusive. Separate analyses of the elution fractions identified with high confidence the same All4300 and hypothetical protein from band 2, as well as a number of putative proteins in the 3L genome such as a peptidase (~45 kDa) and an AP endonuclease (~30 kDa). Several pigment related proteins were also identified that were not visually apparent by SDS-PAGE (smaller than the two main bands indicated on Figure 7), including C-phycoerythrin class 1 subunit alpha (~19 kDa), allophycocyanin alpha subunit (~17 kDa), and photosystem I (PsaD) (~16 kDa).

BLAST analyses of the All4300 and the hypothetical protein (referred to subsequently as protein 5335 and protein 7968, respectively, from the *L. majuscula* 3L genome; annotations in progress) both yielded a number of hypothetical protein matches in other cyanobacteria including *Anabaena variabilis*, *Microcoleus chthonoplastes*, *Nostoc punctiforme*, and *Trichodesmium erythraeum* (JHB protein BLAST hits in Table 2; see below). Interestingly, both proteins also matched (although significantly better for 7968) with the protein RcaD, an activator protein from the cyanobacterium *Calothrix* (= *Fremyella diplosiphon* or *Tolypothrix*) known to regulate complementary chromatic adaptation [32-35]. Complementary chromatic adaptation (CCA) is a phenomenon exhibited by many cyanobacteria in response to changes in light wavelength and intensity. CCA allows cyanobacteria to alter pigment levels so as to optimize their capacity for photosynthesis, and usually involves variation between green and red phenotypes [36]. RcaD is a protein that binds to the promoter for phycocyanin 2 (*cpc2*) and alters the expression of several red light operons in the acclimation phase of CCA [34,35]. Another protein, RcaG, is located downstream of RcaD and has been identified as a putative ATPase. RcaG may facilitate binding of RcaD to DNA, and could require

**Table 2: BLAST results with *Lyngbya majuscula* JHB proteins 5335 and 7968.**

5335 (279 aa)		Size (aa)	identity	similarity	e value	accession #
<b>Best BLAST hit</b>	<b>BLAST organism</b>					
hypothetical protein	<i>Nostoc punctiforme</i> PCC 73102	217	56	70	8.00E-62	YP_001867255
hypothetical protein	<i>Microcoleus chthonoplastes</i> PCC 7420	245	56	71	2.00E-59	ZP_05025825
hypothetical protein all4300	<i>Nostoc</i> sp. PCC 7120	227	49	68	4.00E-54	NP_488340
hypothetical protein	<i>Anabaena variabilis</i> ATCC 29413	221	49	65	1.00E-51	YP_321771
hypothetical protein	<i>Lyngbya</i> sp. PCC 8106	224	47	64	9.00E-47	ZP_01623947
hypothetical protein	<i>Lyngbya</i> sp. PCC 8106	156	33	56	2.00E-11	ZP_01621638
hypothetical protein	<i>Nodularia spumigena</i> CCY9414	100	41	61	3.00E-11	ZP_01628571
hypothetical protein	<i>Arthrospira maxima</i> CS-328	131	32	60	1.00E-08	ZP_03271683
<b>RcaD protein</b>	<b><i>Tolypothrix</i> sp. PCC 7601</b>	<b>285</b>	<b>22</b>	<b>48</b>	<b>0.2</b>	<b>CAC39267</b>
7968 (304 aa)		Size (aa)	identity	similarity	e value	accession #
<b>Best BLAST hit</b>	<b>BLAST organism</b>					
hypothetical protein	<i>Cyanothece</i> sp. PCC 7424	274	49	69	2.00E-68	YP_002380360
<b>RcaD protein</b>	<b><i>Tolypothrix</i> sp. PCC 7601</b>	<b>285</b>	<b>43</b>	<b>63</b>	<b>3.00E-54</b>	<b>CAC39267</b>
hypothetical protein	<i>Trichodesmium erythraeum</i> IMS101	272	40	59	3.00E-52	YP_720119
hypothetical protein	<i>Nodularia spumigena</i> CCY9414	280	44	62	1.00E-50	ZP_01631082
hypothetical protein	<i>Microcoleus chthonoplastes</i> PCC 7420	287	41	62	7.00E-50	ZP_05025219
hypothetical protein	<i>Synechococcus</i> sp. PCC 7335	199	33	57	3.00E-22	ZP_05035072

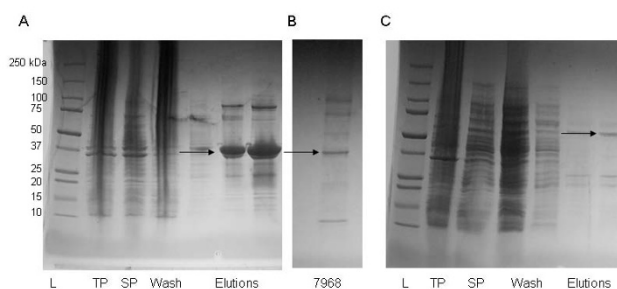
phosphorylation to complete this task [34]. Bioinformatic analysis of the *L. majuscula* 3L genome revealed that the proteins immediately downstream of 5335 and 7968 both resulted in BLAST hits with RcaG, although as with RcaD, the protein neighboring 7968 (7969) had much stronger sequence identity than the neighboring protein to 5335 (5336).

Primers were designed from each of the gene sequences for the two proteins identified above using the *L. majuscula* 3L unfinished genome, and were successful in amplifying homologous gene sequences from *L. majuscula* JHB genomic DNA. The JHB homolog to 5335 encodes for a protein that differs from the 3L protein by only one amino acid (99.6% identical), while the 7968 homolog in JHB encodes for a protein 89.5% identical to the 7968 protein in 3L. Alignments of each JHB protein with their nearest respective BLAST hits (alignment of protein 7968 shown in Additional file 2: Figure S1) indicated several conserved sequence regions, with the highest level of conservation found toward the C terminal end of the proteins (a region in the RcaD protein thought to be involved in DNA binding) [34].

#### Recombinant expression of identified proteins and Electromobility Shift Assays (EMSAs)

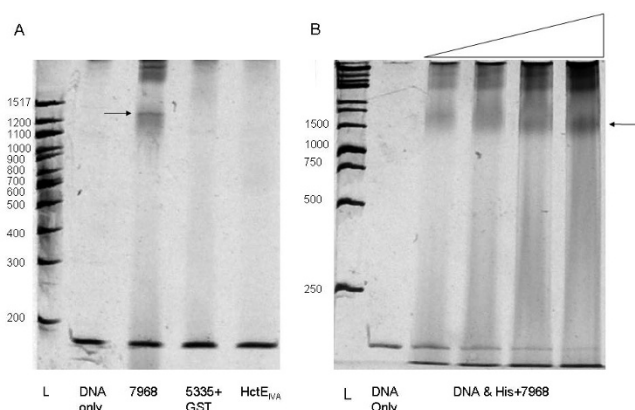
The sequences encoding the 5335 and 7968 proteins in JHB were used in creating constructs for recombinant expression in *E. coli* (Figure 8). After expression and purification of each protein, both were used in Electromobility Shift Assays (EMSAs). In these assays, protein and a fragment of DNA amplified from a region that included both the sequence of the primary jamaicamide promoter and the region upstream from the original probe (1000 -

832 bp upstream of *jamA*) were incubated and visualized on native PAGE gels. Recombinant 7968 was found to bind this putative transcription factor binding region upstream of *jamA* after His tag removal with thrombin cleavage (Figure 9a), although promiscuous binding was also observed with other control DNA fragments (data not shown). A serial titration of 7968 with the N-terminal His tag still attached showed increased DNA binding with larger amounts of protein (Figure 9b). Recombinant protein 5335 was expressed and purified with a GST-tag on the N-terminus of the protein. However, attempts to remove the GST tag were unsuccessful, and thus we assayed protein 5335 with the GST tag still attached (Figure 8c). This version of 5335 did not bind to the *upjamA*-

**Figure 8**

**Recombinant expression of JHB proteins.** A: Protein expression from *L. majuscula* JHB 7968 (His+protein: ~37 kDa). Arrow indicates eluted protein. B: Protein 7968 after thrombin His tag cleavage and concentration. Arrow indicates cleaved protein. C: Protein expression from *L. majuscula* JHB 5335 GST fusion vector (GST+protein: ~60 kDa). Arrow indicates eluted GST+5335 protein.



**Figure 9**

**Electromobility shift assays.** A) EMSA gel shift assay with DNA region -1000 - -832 bp upstream of *jamA*. DNA [270 fmol (= 30 ng)] was assayed with (from left to right) no protein, 7.3 pmol of 7968, 8.4 pmol of GST+5335, or 31 pmol of HctE<sub>IVA</sub>. Arrow indicates DNA + protein shift for 7968. B) Serial titration experiment with 45 fmol (= 5 ng) of the same DNA region with (from left to right) no protein, 6.8 pmol, 13.7 pmol, 27.3 pmol, or 54.8 pmol His+7968. Arrow indicates DNA + protein shift.

1000 - -832 bp region (Figure 9a), even with elevated protein concentrations (Additional file 3: Figure S2).

## Discussion

In this study, we explored the transcriptional machinery associated with the jamaicamide biosynthetic gene cluster in *Lyngbya majuscula*. The jamaicamide cluster was chosen because it possesses a number of features commonly seen in other secondary metabolites isolated from marine cyanobacteria [3]. The jamaicamides are produced by the most prolific cyanobacterial natural product producer yet known (*L. majuscula*), are bioactive (ichthyotoxic, neurotoxic), are composed of mixed PKS/NRPS derived subunits, and contain unusual structural features such as a vinyl chloride and alkynyl bromide rarely seen in natural products from other organisms.

The first description of the jamaicamides [6] demonstrated that the cluster is composed of 17 ORFs, with 16 transcribed in the same direction. The cluster is flanked on the 5' and the 3' ends by transposases and hypothetical proteins. From the results of our RT-PCR experiments, it appears that the gene cluster is preceded by an unusually long untranslated leader region (at least 844 bp), one that may be unprecedented in size for a secondary metabolite gene cluster. The function of having such a long region between the TSS and the start codon of *jamA* is unclear at this time, but may be important for overall regulation of the pathway. In *Synechococcus* PCC 7942, the psBAIL and

psBAIII genes encoding the photosystem II reaction center D1 protein have *cis* regulatory elements in addition to basal promoters. Contained in the untranslated leader region downstream of the psB TSS are light responsive elements that were found to be responsible for increased expression of the genes under high light conditions [37]. In the jamaicamide pathway, the fact that another region of DNA immediately upstream of *jamA* can function as a strong promoter indicates that although transcription may initiate well before the ORF start site, there could be a supplemental means of boosting transcription closer to the first protein in the cluster.

The amplification of second strand cDNA from JHB RNA corresponding to all of the intergenic regions between the jamaicamide ORFs tested indicated that the pathway is transcribed in at least two pieces. The first, *jamABCDEFGHIJKLMNOP*, is sufficiently large (~55 kb) to assume that multiple transcripts could be needed to process this portion of the gene cluster. A similar situation was found with the microcystin gene cluster [22], in which all of the intergenic regions of the pathway aside from the bidirectional promoter were transcribed, and RACE experiments with several of these regions detected variations in intergenic TSS locations. As with microcystin, the jamaicamide pathway could contain internal promoters which, while not representing true breaks in the transcription of the pathway, can function independently if not overwritten by RNAP acting from an upstream promoter (promoter occlusion; [38]). Indeed, several of these regions were able to function as promoters in a reporter assay (see below).

A second transcript in the direction complementary to the large transcript in the jamaicamide pathway is probably needed to include *jamQ*, a gene encoding a condensation like protein that is likely involved with the creation of the pyrrolinone ring of the molecule. According to our RT-PCR experiments, the regions between *jamQ* and the three genes closest upstream (*ORF5* and *ORF6*, both transposases, and *ORF7*, a hypothetical protein), are all transcribed. In addition, the upstream region of *jamQ* does not appear to serve as a strong promoter in  $\beta$ -galactosidase reporter assays (see below), despite the presence of possible conserved promoter domains (Table 1). From these data, it appears that *jamQ* could be part of a larger transcript including these transposases. A larger intergenic region (approximately 1070 bp) lies upstream of *ORF7*, which could contain the TSS and a promoter for this transcript. The reason for including at least one transposase in the *jamQ* transcript is unclear, but this may be a way of ensuring transposable elements have remained associated with the cluster so as to facilitate horizontal gene transfer and pathway evolution. The hectochlorin biosynthetic gene cluster from *L. majuscula* JHB [39] contains a trans-

posase (*hctC*) located between two of the initial genes (*hctB* and *hctD*) in the pathway, which is also thought to contribute to the plasticity of the cluster.

Biosynthetic investigations using *Lyngbya majuscula* strains have been highly successful in identifying secondary metabolite gene clusters, in part because *L. majuscula* readily incorporates isotopically labeled precursors in feeding studies [5,6]. However, further experimentation by way of gene knockout or overexpression in *L. majuscula* is not yet possible because a viable means of genetic transformation has not been developed. Due to this limitation, we used genetic constructs in *E. coli* to determine whether the promoters identified in this study, including the primary pathway promoter upstream of the TSS and those predicted in intergenic regions, were functional. Although some differences exist in the structure of RNAP between the two bacteria [40], promoter structures in cyanobacteria are often compared to consensus sequences in *E. coli* [22,41]. Furthermore, a strong *E. coli* promoter has been shown to function in the cyanobacterium *Synechococcus* [37] and the *psb2* promoter from *Microcystis* can be used in *E. coli* to drive  $\beta$ -galactosidase production [42]. The reporter assay proved effective in verifying the promoter identified upstream of the jamaicamide pathway TSS, as well as several internal promoters located at various regions throughout the gene cluster (Figures 4, 5 and 6). Previous studies with  $\beta$ -galactosidase reporter assays have been useful in demonstrating how enhancer elements within untranslated leader regions and intergenic promoters are important in driving transcription [37,43]. The strongest promoter in the assay was that identified upstream of the *jamA* TSS, but several other promoters were either equal to or greater in strength than the positive control in the assay. One of the regions predicted to contain a strong promoter (*upjamI*) is located in front of a large set of ORFs. The ORF *jamI*, encoding an enoyl-CoA hydratase/isomerase, forms a di-domain with *jamJ*, which encodes for an enoyl reductase and a large PKS [6]. In addition, the subsequent ORFs in the pathway (*jamK* - *M*) are separated by small intergenic regions and do not appear to contain promoters. If *jamI* - *M* form one contiguous transcript (~30 kb), a promoter in front of *jamI* could be needed for efficient transcription. The identification of functional promoters in several other intergenic regions suggests that they could also be used to boost transcription beyond the capacity of the initial promoter located before the TSS upstream of *jamA*.

One intriguing finding from using truncated intergenic regions in the  $\beta$ -galactosidase assay was the detection of strong activity immediately upstream of *jamA* (-76 - 0) and *jamI* (-67 - 0) (Figure 5). An additional promoter was predicted in a region of *upjamI* (-269 - -203) farther upstream in the 5' direction (Table 1), but this region was not active when used in truncated form (Figure 5). If these

active regions upstream of *jamA* (-76 - 0) and *jamI* (-67 - 0) are able to act as internal promoters to supplement overall transcription of the jamaicamide pathway, their close proximity to *jamA* and *jamI* may compromise the ability of transcripts initiating at these positions to subsequently allow for proper translation of the JamA and JamI proteins (although transcription could take place normally downstream of each location). This could occur as a result of insufficient room for a ribosome binding site, although translation of mRNA in cyanobacteria may not require the use of Shine-Dalgarno sequences [44] and some evidence exists for translation of leaderless mRNA in bacteria [45]. It is possible that our heterologous use of these *upjamA* (-76 - 0) and *upjamI* (-67 - 0) regions in *E. coli* could lead to false positive identification of promoters in some instances. However, as previously discussed, the organization of the gene cluster supports the utility of functional promoters in both locations. The untranslated leader region of *jamA* is long enough for the presence of additional regulatory elements, and *upjamI* is a probable location for a promoter because of the long *jamI* - *M* transcript. Further evaluation of these two possible promoters will be necessary to determine how transcription from their locations could affect subsequent protein translation.

Of particular interest in this study was the successful isolation of proteins using "pulldown" experiments that could be involved in the regulation of jamaicamide expression. Gene clusters of marine cyanobacterial PKS/NRPS secondary metabolites identified to date lack any associated regulatory proteins that are imbedded in or in proximity to the main cluster, in contrast to antibiotic pathways in other prokaryotes such as actinobacteria [27]. This absence has led to the suggestion that secondary metabolite pathways from *L. majuscula* could be constitutively expressed [6]. By using the upstream region of *jamA* as a DNA probe, we hoped to isolate putative regulatory proteins from the soluble protein fraction of JHB. This was predicted on the hypothesis that if the jamaicamide pathway does have associated regulatory proteins, they are located elsewhere in the genome. A biotinylated DNA sequence from the jamaicamide pathway (1000 bp upstream of *jamA* to 20 bp into the *jamA* gene) was incubated with protein lysate from *L. majuscula* JHB. The probe was long enough to encompass the entire untranslated leader region of the pathway, as well as the primary promoter and an additional 123 bp upstream of the promoter -35 hexamer. Because transcription factors commonly bind at either the -35 box of the promoter itself, or within 90 bp of the -35 box [46], it is probable that the probe was long enough to capture proteins that might associate with the promoter. The probe also allowed for binding of regulatory proteins with affinity to the untranslated leader region [37,47]. Analysis of protein samples isolated from both an excised SDS-PAGE gel band and

elution fractions of several repeated pulldown assays consistently identified two proteins in three separate data sets using LC-MS/MS. These proteins were partially identified using sequence data from the unfinished *L. majuscula* 3L genome (unpublished), a strain from Curaçao responsible for the production of the anticancer compound curacin A [5,48]. The two proteins (5335 and 7968) displayed strongest sequence identity to hypothetical proteins found in other cyanobacteria, but could not immediately be assigned a function. BLAST searches with both proteins resulted in hits with RcaD, a protein involved in complementary chromatic adaptation (CCA) in another species of cyanobacteria [34]. Interestingly, although the level of sequence identity of the two proteins with RcaD was quite different (Table 2), both proteins (in the 3L genome) had a similar gene neighborhood to RcaD, indicating probable synteny. The *L. majuscula* 3L proteins downstream of each (5336 and 7969) both had BLAST hits with RcaG, the ATPase associated with RcaD, although 7969 (49% identity) had significantly more identity than 5336 (23% identity).

Complementary chromatic adaptation has been identified in a number of freshwater [36] and marine cyanobacteria [49]. In the cyanobacterial CCA model organism *Fremyella* (= *Calothrix*, *Tolypothrix*), a photoreceptor circuit involving the Rca receptors and response regulators (RcaC, RcaE, RcaF, and RcaD) has been found to be responsible for pigment modifications under red and green light [35]. RcaD appears to affect several operons during the acclimation phase of CCA [34]. Although *Lynghya majuscula* strains have not been observed to undergo CCA in culture, there are several color morphotypes known (for example, in our culture collection *L. majuscula* 3L is red under 16 h light/8 h dark cycles, while *L. majuscula* JHB is dark green). In addition, a microarray analysis of cyanobacteria undergoing CCA found that over 80 genes were upregulated, including many not involved in photosynthesis [50]. Considering the widespread effects that CCA regulatory proteins play in cyanobacteria, it is plausible that secondary metabolite production is regulated by homologous proteins. Regulation by light could also be in accordance with the mechanisms previously described for the microcystin biosynthetic pathway [21,22].

To further evaluate the two possible regulatory proteins isolated in the pulldown assay, we overexpressed both proteins in *E. coli* to evaluate their respective binding affinities for the jamaicamide primary promoter region. Protein 7968 was found to bind to the proposed transcription factor binding region of the jamaicamide pathway (1000-832 bp upstream of *jamA*; Figure 9a), and this DNA binding activity was supported with serial protein titration (Figure 9b). Although we demonstrated that a control protein would not bind under the same conditions,

we also found that protein 7968 was able to bind nonspecifically to several other unrelated pieces of DNA. Thus, we were unable to assign a specific sequence for 7968 binding. Attempts to cleave the GST tag from the 5335 protein were unsuccessful, and binding assays indicated that the GST+5335 fusion protein was not able to bind to the same intergenic region as 7968 (Figure 9a; Additional File 3: Figure S2). Because of its strong affinity with DNA, 7968 is the better candidate protein for providing transcriptional regulation of the jamaicamide pathway. The presence of multiple intergenic promoters in the pathway could also offer other binding locations for additional regulation.

It is difficult to predict how the binding affinity of recombinant forms of 5335 or 7968 compares quantitatively with the native proteins. Noubir et al. [34] found that native RcaD bound much more effectively to the phycocyanin 2 promoter than a recombinant version, and hypothesized that the reduced affinity may be from lack of ATPase RcaG, which facilitates binding, or from lack of phosphorylation. We attempted a dual-shift experiment with 7968 and the GST tagged 5335, but no shift differences compared to 7968 alone were observed (data not shown). It will be intriguing to determine whether 5335 and 7968 work in tandem to regulate the jamaicamide pathway, or if they require downstream neighbors (5336 or 7969) to assist in binding. Alternatively, it is possible that 7968 is the true regulator of the pathway, and 5335 was "pulled down" in the magnetic bead assay due to its sequence identity being minimally sufficient for recognition. Interestingly, protein 7968 was found to form dimers by PAGE analysis. Transcription factors often function as dimers in their association with DNA and RNAP [46], and thus, this finding also supports 7968 as the best candidate regulatory protein identified in this study.

If transcription factors are in fact regulating the expression of secondary metabolites such as jamaicamide, it is useful to consider the potential pleiotropic role of proteins such as 7968 in regulating more than one biosynthetic pathway in *L. majuscula* JHB. There are a number of similarities in the secondary metabolite gene clusters of *L. majuscula*, such as those encoding for the jamaicamides, hectochlorin (also produced by the JHB strain; [39] and curacin A [5,51]. For example, the genes *jamA* and *hctA* are both ACP synthetases and are 58% identical, which might indicate that similar regulatory proteins associate with the upstream regions of each gene. If jamaicamide and hectochlorin are both used in defense of *L. majuscula* against predation or infection, their co-regulation would enhance the defense of the strain. It is also interesting to speculate that proteins in *L. majuscula* 3L homologous to jamaicamide regulatory proteins could be used to regulate production of curacin A. A comparison of the approximately 1700 bp that separate *jamA* from its upstream neighbor-

ing gene (a transposase) with the upstream region of *curA* from the curacin A pathway reveals that approximately 1550 bp of the *upjamA* region is 95% identical with the *upcurA* region. Moreover, proteins 5335 and 7968 are 99.6% and 89.5% identical with their respective homologs in *L. majuscula* 3L (the curacin A producer). If either of these two proteins functions as a pleiotropic regulator for natural products biosynthesis in *L. majuscula*, their use in overexpression efforts would be valuable in unlocking the full biosynthetic potential of these filamentous marine cyanobacteria.

Ultimately, quantitative co-transcription analyses of the two proteins with the rest of the jamaicamide pathway and gene knockouts will be necessary to conclusively link these proteins with jamaicamide regulation. Current efforts are evaluating transcription levels of the two proteins with both jamaicamide transcription and compound production, and the effect of variable light wavelengths on jamaicamide production in culture. Because targeted gene manipulation techniques in *L. majuscula* have not yet been developed, we are also in the process of conducting methodology experiments to disrupt or overexpress 5335 and 7968 to better understand their functions, including their roles in global regulation.

## Conclusion

Understanding the regulation of natural product pathways that encode compounds with pharmaceutical potential is important to overcoming the "supply issue" that is so prevalent in natural products research [8]. While marine cyanobacteria are recognized as prolific producers of bioactive compounds, natural product yields from field collections are low, and slow culture growth severely limits the amount of compound that can be produced in this manner. The transcriptional profile of the jamaicamide biosynthetic gene cluster presented here provides insight into the mechanisms by which these pathways are transcribed and potentially regulated. Future advances in classifying promoters and transcription factors for cyanobacterial gene clusters will be important to diverse applications in biotechnology, such as combinatorial biosynthesis and the heterologous expression of entire natural product pathways. Additionally, this information should also benefit ongoing efforts attempting to regulate the expression of cyanobacterial toxins with deleterious environmental impacts.

## Methods

### Bacterial strains, culture conditions, PCR reactions, and DNA measurements

*Lyngbya majuscula* JHB was originally collected from Hector's Bay, Jamaica [6] and was maintained in a culture facility at Scripps Institution of Oceanography. Cultures were grown in BG-11 saltwater media at 29°C under a light intensity of approximately 5  $\mu\text{E m}^{-2} \text{s}^{-1}$  and under 16

h light/8 h dark cycles. *E. coli* TOP-10 and BL-21 (DE3) were grown in Luria-Bertani (LB) media. *E. coli* cultures were grown with ampicillin (100  $\mu\text{g ml}^{-1}$ ), or kanamycin (50  $\mu\text{g ml}^{-1}$ ) when necessary. PCR reactions were conducted using either PCR Master Mix (Promega) or Pfx50 proofreading Taq Polymerase (Invitrogen). DNA concentrations were measured using either Beckman-Coulter DU800 or NanoDrop 1000 (Thermo Scientific) spectrophotometers. Protein concentrations for recombinant JHB proteins were determined using the BCA assay (Pierce). Ladders for DNA (Fermentas and New England Biolabs) and protein (Bio-Rad) were used for size estimations when necessary.

### RT-PCR using *L. majuscula* RNA to search for the transcription start site (TSS) and promoter regions in the jamaicamide pathway

Cyanobacterial filaments (approximately 2 g wet weight) from a culture of the jamaicamide producing strain of *L. majuscula* JHB were harvested and subjected to RNA isolation using TRIzol reagent (Invitrogen) and procedures based on those recommended by the manufacturer with minor modifications. RNA was treated with TURBO DNase (Ambion) for 2 h at 38°C before use in cDNA reactions. To verify that genomic DNA contamination was not present, in selected cases negative control reactions were run in parallel with cDNA reactions in which reverse transcriptase enzyme was omitted. For the primer extension experiment, first strand cDNA was synthesized from the RNA using the primer *upjamA* 20-0 R (Sigma Genosys; Additional file 1: Table S1) and the Superscript III Reverse Transcriptase Protocol (Invitrogen) with minor modifications. Second strand reactions were conducted with primers ranging from 500-902 bp upstream in 50 bp increments to determine where RNA transcription upstream of *jamA* initiated. For cDNA synthesis of jamaicamide intergenic regions, first strand cDNA was generated using either random (Invitrogen and [52]; Additional file 1: Table S1) or specific jamaicamide upstream intergenic region reverse primers (Additional file 1: Table S1). Forward and reverse oligonucleotide primers for each upstream intergenic region of the cluster were used to PCR amplify regions from the first strand cDNA to create second strand cDNA. In some instances, sequencing was used to confirm the correct amplification of second strand cDNA, either by direct sequencing of PCR products or by sequencing of TOPO TA cloning vectors (Invitrogen) containing the intergenic region. All cDNA PCR products were visualized on agarose gels.

### Use of promoter prediction and -galactosidase reporter gene assays to search for promoter activity

Each intergenic region upstream of the genes in the jamaicamide pathway was examined for conserved binding regions (in comparison to the  $\sigma^{70E. coli}$  consensus promoter) using the BPROM predictor (<http://www.soft>

[berry.com](#); Table 1). The upstream (up-) regions of genes predicted to contain a promoter (*jamA*, *jamB*, *jamC*, *jamD*, *jamG*, *jamI*, *jamN*, and *jamQ*) were amplified with specific primers (Additional file 1: Table S1) from fosmids produced previously [6]. Each upstream section was individually cloned into the pBLUE TOPO vector (Invitrogen) and transformed into TOP-10 *E. coli*. Plasmid purification (Qiagen) and sequencing (Seqxcel, Inc., La Jolla, CA) were used to confirm the sequence and direction of the inserts. Verified clones were used to measure relative promoter activity using the  $\beta$ -galactosidase reporter gene assay (Invitrogen), standardized against total soluble protein content as measured by BCA assay (Pierce). For this protocol, colonies from freshly streaked selective (100  $\mu$ g ml<sup>-1</sup> ampicillin) plates were grown overnight in LB media (5 ml). Cultures were pelleted at 4000 RPM, and the pellets were washed once with 3 ml of chilled PBS. After an additional centrifugation step, the pellets were resuspended in 1 ml PBS. The tube contents were centrifuged at 14,000 RPM for 5 min, and subsequently resuspended in 110  $\mu$ l lysis buffer (0.25 mM Tris, pH 8.0). The cells were lysed by freeze-thawing the suspensions 4 times with dry ice and a 37°C water bath, followed by another centrifugation period. The supernatant was removed from each tube for use in the  $\beta$ -galactosidase and BCA assays in a 96 well plate. O-nitrophenol measurements from ONPG cleavage in the assay were taken at 420 nm, while BCA readings were taken at 570 nm (Thermo-Electron Multiskan Ascent plate reader). Serial dilutions of lysis buffer from each culture were made to find an optimal range for colorimetric detection of o-nitrophenol and for a comparison of relative activity (identified as a 3-fold dilution). Relative activity calculations were made by determining the average nmol ONPG hydrolyzed min<sup>-1</sup> mg soluble protein<sup>-1</sup> value for each vector insert and dividing each value by the highest overall average value to obtain a percentage. A fragment of the pBAD TOPO vector (Invitrogen) containing the arabinose promoter and vector ribosome binding site (upstream of the TOPO cloning site) from *E. coli* was found to consistently produce  $\beta$ -galactosidase in the pBLUE TOPO vector in preliminary experiments, and was used as a positive control. Because the arabinose operator was not included in the positive control, the addition of arabinose was not required to produce  $\beta$ -galactosidase. A 49 bp segment of the jamaicamide *jamG* gene was used as a negative control. [Note: the pBLUE vector contains a cryptic promoter that is reported to possibly limit the efficacy of assaying other promoter fragments in a prokaryotic host (Invitrogen). However, a series of preliminary assays indicated significant and repeatable differences in promoter activity between possible promoter regions, and baseline activity in the negative control was sufficiently low as to not conflict with the assay results. The BPROM prediction software was used to verify that the vector constructs did not introduce any artificial promoters]. Those regions found to have promoter activity were assayed

again with additional dilution (10 fold) to quantify promoter strength, expressed as specific activity (nmol ONPG hydrolyzed min<sup>-1</sup> mg soluble protein<sup>-1</sup>).

#### **Isolation of possible transcription factors from a pulldown assay**

Protein pulldown experiments were based on methods similar to [53]. A DNA probe that extended from 1000 bp upstream of *jamA* to 20 bp into the *jamA* gene was amplified by PCR from the jamaicamide fosmid described above using the primers up*jamA* 1000 biotin (biotinylated at the 5' end; Invitrogen) and up*jamA* 20 - 0 R (Additional file 1: Table S1). The PCR product was purified (MinElute PCR Purification Kit, Qiagen) and 10 pmol of the biotinylated DNA were incubated with 1 mg of magnetic M-270 streptavidin Dynabeads (Invitrogen), according to the manufacturer's instructions. *L. majuscula* JHB tissue was obtained from pan cultures that had been growing for 1-2 months. Approximately 2-3 ml of culture was measured by displacement in sterile, chilled binding buffer [10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT, 150 mM NaCl, and 5% (w/v) glycerol]. The binding buffer was also treated with a broad range protease inhibitor (Complete, EDTA free; Roche). The tissue was sonicated and kept on ice using a probe sonicator with six 10-s pulses, and insoluble material was pelleted at 13,200 RPM for 10 minutes. The soluble protein fraction (750  $\mu$ l) was added to each mg of DNA coated beads. One  $\mu$ g of Poly DI-DC was also added to inhibit non-specific binding of protein to the DNA. Magnetic beads that were not treated with biotinylated DNA were incubated with JHB soluble protein as a negative control. The beads and soluble protein were incubated for 1 h using an end-over-end rotator at 4°C. The beads were subsequently washed twice using 200  $\mu$ l of binding buffer containing 100  $\mu$ l sheared salmon sperm DNA (Invitrogen; 5 mg ml<sup>-1</sup>), three times with binding buffer, and eluted with 50  $\mu$ l of binding buffer containing 1.0 M NaCl. Wash fractions and elutions were flash frozen in liquid N<sub>2</sub> and concentrated on a freeze dryer. After concentration, aliquots of each were mixed with protein sample buffer, denatured for 3 minutes at 95-100°C, and analyzed by SDS-PAGE. The gels were stained with either silver (Silverquest Kit, Invitrogen) or colloidal Coomassie brilliant blue G-250.

#### **Identification of DNA binding proteins**

Once gel bands were visible in the elution fraction from the binding assay, the assay was repeated on a larger scale using additional replicates of the procedure described above to isolate sufficient protein for mass spectrometry (visible by colloidal Coomassie staining). Both gel bands (excised using a scalpel) and whole elution fractions were submitted to The Scripps Research Institute (La Jolla, CA) Center for Mass Spectrometry for nano-LC MS/MS analysis. Raw spectrum data (mzdata format) was obtained and



analyzed at UCSD by a DOS common-line version of InsPecT 20070712 [31].

InsPecT search parameters for the mzdata files were the following: (i) *Lyngbya majuscula* 3L common database (unpublished data), common contaminants database, reverse or "phony" database, and NCBI nr database; (ii) parent ion  $\Delta m = 1.5$  Da; (iii) b and y-ion  $\Delta m = 0.5$  Da. Top protein identifications were verified by using two different database searches: (i) *Lyngbya majuscula* 3L genome alone; (ii) NCBI nr with *L. majuscula* 3L genome inserted. The mass spectral identifications of 5335 and 7968 were further verified by manual annotation of the N-terminal and C-terminal peptides, as well as the most abundant peptide identified.

#### **Characterization of putative transcription factors from a pulldown assay**

Protein sequences detected using InsPecT were compared with raw nucleotide sequences from the *L. majuscula* 3L genome to identify their corresponding ORFs. Forward and reverse primers (5335 F & R, 7968 F & R, Additional file 1: Table S1) were designed from each sequence and used to amplify the corresponding genes from *L. majuscula* JHB. The blunt PCR products were cloned (Z-Blunt TOPO vector, Invitrogen) and transformed into *E. coli* for sequencing to compare the gene sequences from JHB with those of 3L. Additional gene boundary primers (5335 FB, 5335 RB; 7968 FB, 7968 RB; Additional file 1: Table S1) were used to amplify the JHB genes with priming sites 25 bp upstream and downstream in order to verify the sequences covered by 5335 and 7968 forward and reverse primers and avoid inclusion of sequences from *L. majuscula* 3L. Bioinformatic analyses of each gene sequence were conducted using BLAST programs available through the National Center for Biotechnology Information (NCBI; <http://blast.ncbi.nlm.nih.gov/>).

#### **Recombinant expression of identified proteins**

Genes corresponding to identified proteins in the JHB protein pulldown assay were amplified from JHB genomic DNA using the primers 5335 Nco1F and 5335 Not1R or 7968 Nde1F and 7968 Xho1R (Additional file 1: Table S1). Start codons in both gene sequences were changed from either leucine (5335) or valine (7968) to methionine for improvements in *E. coli* overexpression efforts. PCR products from both reactions were purified and digested with their corresponding restriction enzymes (New England Biolabs). Gene 5335 was ligated into the pGS21a vector (Genscript), which contains both an N-terminal 6 $\times$  His tag and GST tag. The 5335 construct was verified via transformation, plasmid isolation from TOP-10 *E. coli*, and sequencing. The vector containing the gene sequence was then transformed into BL-21 (DE3) *E. coli*. Four liters of *E. coli* harboring the 5335 pGS21a vector were grown (using starter cultures) for 4 h at 37°C to an

OD<sub>600</sub> between 0.6 and 1.0, induced with 0.7 mM IPTG, and then grown at 18°C overnight. Cultures were centrifuged at 4500 RPM for 15 min at 4°C, and the ensuing pellets from each liter of culture were resuspended in 5 ml protein lysis buffer (20 mM Tris, pH 8.0, 500 mM NaCl, and 20 mM imidazole) and lysed on ice using sonication (6-7 10-s pulses). Resuspended lysate was centrifuged, and supernatant containing soluble protein was incubated on an end-over-end rotator at 4°C with Nickel-Superflow resin (Qiagen) for 2 h. Following incubation, the recombinant GST+5335 fusion protein was purified using polypropylene columns (Qiagen). The nickel slurry from the incubation was washed twice with protein wash buffer (20 mM Tris, pH 8.0, 500 mM NaCl, and 50 mM imidazole), and protein was eluted with 5  $\times$  1 mL aliquots of protein elution buffer (20 mM Tris, pH 8.0, 500 mM NaCl, and 750 mM imidazole). Purified protein was dialyzed against binding buffer (the same buffer used for the pulldown assay, but not containing DTT) overnight using a 50 kDa MWCO dialysis membrane (Spectrum Labs, Rancho Dominguez, CA). To express 7968, the corresponding gene sequence was ligated into the N-terminal His tag containing pET28b vector (Novagen), and grown and purified similarly to 5335 (although the 7968 wash buffer contained 40 mM imidazole). Spectra/Por Float-A-Lyzer G2 dialysis membranes (20 kDa MWCO; Spectrum Labs) were used to dialyze protein 7968. Concentrations of each protein for use in assays were determined using the BCA assay (Pierce).

#### **Electromobility shift assays (EMSAs)**

Gel shift EMSAs were performed to verify binding of 5335 and 7968 to jamaicamide promoter regions. The region upstream of the jamaicamide TSS (1000 - 832 bp upstream of *jamA*) was amplified from a jamaicamide fosmid using Pfx50 Taq Polymerase (Invitrogen). Each PCR product was purified (MinElute kit, Qiagen) before being used in the assay. For the comparative binding assay (Figure 9a), the N-terminal His tag was cleaved from protein 7968 using the Thrombin Cleavage Capture Kit (Novagen). The cleaved 6 $\times$  His tag was subsequently removed by concentrating the protein sample over a Microcon 10,000 MWCO column (Millipore). SDS-PAGE gels and western blotting were conducted to confirm the success of the cleavage reaction. Serial titration of the 7968 protein to verify DNA binding was performed using the recombinant version of the protein without His tag cleavage. GST+5335 elutions were also concentrated over Microcon 10,000 MWCO columns prior to use in shift assays. For a negative control, the purified recombinant protein HctE<sub>IVA</sub> from the hectochlorin pathway (purification described in [39]) was used. The concentrations of HctE<sub>IVA</sub> protein used in the EMSA experiments were measured using a Bradford assay, and the purified HctE<sub>IVA</sub> included a 6 $\times$  N-terminal His tag from its expression vector (pET28b; Novagen). Each gel shift assay reaction was per-

formed with the indicated quantities of DNA and purified protein (Figure 9) in EMSA binding buffer adapted from the DIG Gel Shift Kit, 2<sup>nd</sup> generation protocol (Roche) [20 mM HEPES, pH 7.6, 1 mM EDTA, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1 mM DTT, Tween 20, 0.2% (w/v), 30 mM KCl] and water (total volume 20 µl) for 30 min at room temperature. Following the incubation period, 5 µl of native loading dye containing bromophenol blue was added to each reaction, and the reaction contents were immediately transferred to a 10% native PAGE gel. The gels were electrophoresed at 85 V for ~3.0 h in 0.5× TBE buffer (44.5 mM Tris, 44.5 mM boric acid, 1 mM EDTA), followed by staining for at least 10 min in SYBR Gold Nucleic Acid Gel Stain (Molecular Probes/Invitrogen) and visualization on a UV transilluminator.

### Sequence information

DNA and amino acid sequences of the proteins identified in this study have been deposited in Genbank under the accession numbers [GQ860962](#) and [GQ860963](#).

### Authors' contributions

ACJ, LG, and WHG conceived of the study and designed experiments, ACJ performed experiments and drafted the manuscript, and DG and PCD performed protein mass spectrometry analyses. All authors contributed to, read, and approved the final manuscript.

### Additional material

#### Additional file 1

**Table S1: Primers used in this study.** This Excel file (.xls) is a complete list of all primers used for RT-PCR experiments, -galactosidase reporter assays, and protein identification, recombinant expression, and EMSA experiments.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2180-9-247-S1.XLS>]

#### Additional file 2

**Figure S1: Sequence alignment with *Lyngbya majuscula* JHB protein 7968 and 5 proteins with highest identity matches from NCBI BLAST analyses.** This TIFF file (.tiff) shows an alignment of these 6 protein sequences performed in ClustalX2.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2180-9-247-S2.TIFF>]

#### Additional file 3

**Figure S2: EMSA with DNA region -1000 - -832 bp upstream of *jama* and protein GST+5335.** This TIFF file (.tiff) shows, from left to right: 270 fmol DNA only, 8.4 pmol, 16.4 pmol, 33.5 pmol, and 67.0 pmol of GST+5335 combined with 270 fmol DNA.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2180-9-247-S3.TIFF>]

### Acknowledgements

The authors wish to thank Carla M. Sorrels for assistance with RNA extraction procedures, Sheila Podell for assistance with bioinformatics, and R. Cameron Coates and Tara Byrum for maintenance of JHB cultures. This work was supported by grants from the NIH (GM075832 and NS053398), and NOAA Grant NA08OAR4170669, California SeaGrant College Program Project SG-100-TECH-N, through NOAA's National Sea Grant College Program, US Department of Commerce. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of California Sea Grant or the US Department of Commerce.

### References

- Blunt JW, Copp BR, Hu W-P, Munro MHG, Northcote PT, Prinsep MR: **Marine natural products.** *Nat Prod Rep* 2008, **25**:35-94.
- Tan LT: **Bioactive natural products from marine cyanobacteria for drug discovery.** *Phytochem* 2007, **68**:954-979.
- Tidgewell K, Clark BR, Gerwick WH: **The natural products chemistry of cyanobacteria.** In *Comprehensive Natural Products Chemistry* Pergamon Press in press.
- Chang Z, Flatt P, Gerwick WH, Nguyen V-A, Willis CL, Sherman DH: **The barbamide biosynthetic gene cluster: A novel cyanobacterial system of mixed polyketide synthase (PKS)-non-ribosomal peptide synthetase (NRPS) origin involving an unusual trichlorooleucyl starter unit.** *Gene* 2002, **296**:235-247.
- 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.
- Edwards DJ, Marquez BL, Nogle LM, McPhail K, Goeger DE, Roberts MA, Gerwick WH: **Structure and biosynthesis of the jamaicamides, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium *Lyngbya majuscula*.** *Chem Biol* 2004, **11**:817-833.
- Gu L, Geders TW, Wang B, Gerwick WH, Håkansson K, Smith JL, Sherman DH: **GNAT-like strategy for polyketide chain initiation.** *Science* 2007, **318**:970-974.
- Cragg GM, Newman DJ, Snader KM: **Natural products in drug discovery and development: The United States National Cancer Institute Role.** In *Phytochemicals in Human Health Protection, Nutrition, and Plant Defense* Edited by: Romeo JT. New York: Kluwer Academic/Plenum Publishers; 1999:1-29.
- Suyama TL, Gerwick WH: **Stereospecific total synthesis of somocystinamide A.** *Org Lett* 2008, **10**:4449-4452.
- Pfeifer BA, Wang CCC, Walsh CT, Khosla C: **Biosynthesis of yersiniabactin, a complex polyketide-nonribosomal peptide, using *Escherichia coli* as a heterologous host.** *Appl Environ Microbiol* 2003, **69**:6698-6702.
- Schmidt EW, Nelson JT, Rasko DA, Sudek S, Eisen JA, Haygood MG, Ravel J: **Patellamide A and C biosynthesis by a microcin-like pathway in *Prochloron didemni*, the cyanobacterial symbiont of *Lissoclonium patella*.** *Proc Natl Acad Sci USA* 2005, **102**:7315-7320.
- Watanabe K, Hotta K, Praseuth AP, Koketsu K, Migita A, Boddy CN, Wang CCC, Oguri H, Oikawa H: **Total biosynthesis of antitumor nonribosomal peptides in *Escherichia coli*.** *Nat Chem Biol* 2006, **2**:423-428.
- Wilkinson B, Micklefield J: **Mining and engineering natural-product biosynthetic pathways.** *Nat Chem Biol* 2007, **3**:379-386.
- Galm U, Shen B: **Expression of biosynthetic gene clusters in heterologous hosts for natural product production and combinatorial biosynthesis.** *Exp Op Drug Discov* 2006, **1**:409-437.
- 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.
- Ishiyama M, Kutsuna S, Aoki S, Iwasaki H, Andersson CR, Tanabe A, Golden SS, Johnson CH, Kondo T: **Expression of a gene cluster *kaiABC* as a circadian feedback process in cyanobacteria.** *Science* 1998, **281**:1519-1523.
- Yoon H-S, Golden JW: **PatS and products of nitrogen fixation control heterocyst pattern.** *J Bacteriol* 2001, **183**:2605-2613.

18. Lindell D, Padan E, Post AF: **Regulation of ntcA expression and nitrite uptake in the marine *Synechococcus* sp. strain WH 7803.** *J Bacteriol* 1998, **180**:1878-1886.
19. de Figueiredo DR, Azeiteiro UM, Esteves SM, Goncalves F, Pereira MJ: **Microcystin-producing blooms—a serious global public health issue.** *Ecotoxicol Environ Safety* 2004, **59**:151-163.
20. 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.
21. Kaebnick M, Neilan BA, Börner T, Dittmann E: **Light and the transcriptional response of the microcystin biosynthesis gene cluster.** *Appl Environ Microbiol* 2000, **66**:3387-3392.
22. Kaebnick M, Dittmann E, Börner T, Neilan BA: **Multiple alternate transcripts direct the biosynthesis of microcystin, a cyanobacterial nonribosomal peptide.** *Appl Environ Microbiol* 2002, **68**:449-455.
23. Sevilla E, Martin-Luna B, Vela L, Bes MT, Fillat MF, Peleato ML: **Iron availability affects mcvD expression and microcystin-LR synthesis in *Microcystis aeruginosa* PCC7806.** *Env Microbiol* 2008, **10**:2476-2483.
24. Martin-Luna B, Sevilla E, Hernandez JA, Bes MT, Fillat MF, Peleato ML: **Fur from *Microcystis aeruginosa* binds in vitro promoter regions of the microcystin biosynthesis gene cluster.** *Phytochem* 2006, **67**:876-881.
25. Saito K, Sei Y, Miki S, Yamaguchi K: **Detection of microcystin-metal complexes by using cryospray ionization-Fourier transform ion cyclotron resonance mass spectrometry.** *Toxicon* 2008, **51**:1496-1498.
26. 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, **9**:965-970.
27. Bibb MJ: **Regulation of secondary metabolism in Streptomyces.** *Curr Opin Microbiol* 2005, **8**:208-215.
28. Aoyama T, Takanami M, Ohtsuka E, Taniyama Y, Marumoto R, Sato H, Ikehara M: **Essential structure of *E. coli* promoter: Effect of spacer length between the two consensus sequences on promoter function.** *Nuc Acids Res* 1983, **11**:5855-5864.
29. Hawley DK, McClure WR: **Compilation and analysis of *Escherichia coli* promoter DNA sequences.** *Nuc Acids Res* 1983, **11**:2237-2255.
30. Lissner S, Margalit H: **Compilation of *E. coli* mRNA promoter sequences.** *Nuc Acid Res* 1993, **21**:1507-1516.
31. Tanner S, Shu H, Frank A, Wang L-C, Zandi E, Mumby M, Pevzner PA, Bafna V: **InsPecT: Identification of posttranslationally modified peptides from tandem mass spectra.** *Anal Chem* 2005, **77**:4626-4639.
32. Sobczyk A, Bely A, Tandeau de Marsac N, Houmard J: **A phosphorylated DNA-binding protein is specific for the red-light signal during complementary chromatic adaptation in cyanobacteria.** *Mol Microbiol* 1994, **13**:875-885.
33. Schyns G, Jia L, Coursin T, Tandeau de Marsac N, Houmard J: **Promoter recognition by a cyanobacterial RNA polymerase: In vitro studies with the *Calothrix* sp. PCC 7601 transcriptional factors RcaA and RcaD.** *Plant Mol Biol* 1998, **36**:649-659.
34. Noubir S, Luque I, Ochoa de Alda JAG, Perewoska I, Tandeau de Marsac N, Copley JG, Houmard J: **Co-ordinated expression of phycobiliprotein operons in the chromatically adapting cyanobacterium *Calothrix* PCC 7601: a role for RcaD and RcaG.** *Mol Microbiol* 2002, **43**:749-762.
35. Kehoe DM, Gutu A: **Responding to color: The regulation of complementary chromatic adaptation.** *Ann Rev Plant Biol* 2006, **57**:127-150.
36. Li L, Alvey RM, Bezy RP, Kehoe DM: **Inverse transcriptional activities during complementary chromatic adaptation are controlled by the response regulator RcaC binding to red and green light-responsive promoters.** *Mol Microbiol* 2008, **68**:286-297.
37. Li R, Golden SS: **Enhancer activity of light-responsive regulatory elements in the untranslated leader regions of cyanobacterial psbA genes.** *Proc Natl Acad Sci USA* 1993, **90**:11678-11682.
38. Gonzalez-y-Merchand JA, Colston MJ, Cox RA: **Roles of multiple promoters in transcription of ribosomal DNA: Effects of growth conditions on precursor rRNA synthesis in mycobacteria.** *J Bacteriol* 1998, **180**:5756-5761.
39. Ramaswamy AV, Sorrels CM, Gerwick WH: **Cloning and biochemical characterization of the hecotochlorin biosynthetic gene cluster from the marine cyanobacterium *Lyngbya majuscula*.** *J Nat Prod* 2007, **70**:1977-1986.
40. Xie WQ, Jager K, Potts M: **Cyanobacterial RNA polymerase genes rpoC1 and rpoC2 correspond to rpoC of *Escherichia coli*.** *J Bacteriol* 1989, **171**:1967-1973.
41. Shibato J, Agrawal GK, Kato H, Asayama M, Shirai M: **The 5'-upstream cis-acting sequences of a cyanobacterial psbA gene: Analysis of their roles in basal, light-dependent and circadian transcription.** *Mol Genet Genom* 2002, **267**:684-694.
42. Shibato J, Asayama M, Shirai M: **Specific recognition of the cyanobacterial psbA promoter by RNA polymerases containing principal sigma factors.** *Biochim Biophys Acta* 1998, **1442**:296-303.
43. Nakano MM, Zuber P, Glaser P, Danchin A, Hulett FM: **Two-component regulatory proteins ResD-ResE are required for transcriptional activation of fnr upon oxygen limitation in *Bacillus subtilis*.** *J Bacteriol* 1996, **178**:3796-3802.
44. Mutsuda M, Sugiura M: **Translation initiation of cyanobacterial rbcS mRNAs requires the 38-kDa ribosomal protein S1 but not the Shine-Dalgarno sequence: Development of a cyanobacterial in vitro translation system.** *J Biol Chem* 2006, **281**:38314-38321.
45. Moll I, Grill S, Gualerzi CO, Blasi U: **Leaderless mRNAs in bacteria: Surprises in ribosomal recruitment and translational control.** *Mol Microbiol* 2002, **43**:239-246.
46. Browning DF, Busby SJW: **The regulation of bacterial transcription initiation.** *Nat Rev Microbiol* 2004, **2**:57-65.
47. Hobl B, Mack M: **The regulator protein PyrR of *Bacillus subtilis* specifically interacts in vivo with three untranslated regions within pyr mRNA of pyrimidine biosynthesis.** *Microbiol* 2007, **153**:693-700.
48. 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.
49. Palenik B: **Chromatic adaptation in marine *Synechococcus* strains.** *Appl Environ Microbiol* 2001, **67**:991-994.
50. Stowe-Evans EL, Ford J, Kehoe DM: **Genomic DNA Microarray Analysis: Identification of new genes regulated by light color in the cyanobacterium *Fremyella diplosiphon*.** *J Bacteriol* 2004, **186**:4338-4349.
51. Gu L, Wang B, Kulkarni A, Geders TW, Grindberg RV, Gerwick L, Håkansson K, Wipf P, Smith JL, Gerwick WH, Sherman DH: **Metamorphic enzyme assembly in polyketide diversification.** *Nature* 2009, **459**:731-735.
52. Frias-Lopez J, Bonheyo GT, Fouke BW: **Identification of differential gene expression in bacteria associated with coral black band disease by using RNA-arbitrarily primed PCR.** *Appl Environ Microbiol* 2004, **70**:3687-3694.
53. Rachid S, Gerth K, Kochems I, Müller R: **Deciphering regulatory mechanisms for secondary metabolite production in the myxobacterium *Sorangium cellulosum* So ce56.** *Mol Microbiol* 2007, **63**:1783-1796.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
http://www.biomedcentral.com/info/publishing\_adv.asp

