

Cell Biology and Gene Regulation of Ultraviolet Radiation Mitigation Strategies in the
Cyanobacterium Nostoc punctiforme

by

Kevin Klicki

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2021 by the
Graduate Supervisory Committee:

Ferran Garcia-Pichel, Chair
Rajeev Misra
Aindrila Mukhopadhyay
Melissa Wilson

ARIZONA STATE UNIVERSITY

August 2021

ABSTRACT

I studied the molecular mechanisms of ultraviolet radiation mitigation (UVR) in the terrestrial cyanobacterium *Nostoc punctiforme* ATCC 29133, which produces the indole-alkaloid sunscreen scytonemin and differentiates into motile filaments (hormogonia). While the early stages of scytonemin biosynthesis were known, the late stages were not. Gene deletion mutants were interrogated by metabolite analyses and confocal microscopy, demonstrating that the *ebo* gene cluster, was not only required for scytonemin biosynthesis, but was involved in the export of scytonemin monomers to the periplasm. Further, the product of gene *scyE* was also exported to the periplasm where it was responsible for terminal oxidative dimerization of the monomers. These results opened questions regarding the functional universality of the *ebo* cluster. To probe if it could play a similar role in organisms other than scytonemin producing cyanobacteria, I developed a bioinformatic pipeline (Functional Landscape And Neighbor Determining gEnomic Region Search; FLANDERS) and used it to scrutinize the neighboring regions of the *ebo* gene cluster in 90 different bacterial genomes for potentially informational features. Aside from the scytonemin operon and the *edb* cluster of *Pseudomonas spp.*, responsible for nematode repellence, no known clusters were identified in genomic *ebo* neighbors, but many of the *ebo* adjacent regions were enriched in signal peptides for export, indicating a general functional connection between the *ebo* cluster and biosynthetic compartmentalization. Lastly, I investigated the regulatory span of the two-component regulator of the scytonemin operon (*scyTCR*) using RNAseq of *scyTCR* deletion mutants under UV induction. Surprisingly, the knockouts had decreased

expression levels in many of the genes involved in hormogonia differentiation and in a putative multigene regulatory element, *hcyA-D*. This suggested that UV could be a cue for developmental motility responses in *Nostoc*, which I could confirm phenotypically. In fact, UV-A simultaneously elicited hormogonia differentiation and scytonemin production throughout a genetically homogenous population. I show through mutant analyses that the partner-switching mechanism coded for by *hcyA-D* acts as a hinge between the scytonemin and hormogonia based responses. Collectively, this dissertation contributes to the understanding of microbial adaptive responses to environmental stressors at the genetic and regulatory level, highlighting their phenomenological and mechanistic complexity.

DEDICATION

To my friends, my family, and my love Bianca

“I pulse ov existence
The law ov nature undenied
I hold the torch ov Heraclitus
So I can shake the earth and move the suns”
-Behemoth

“Even as we get older
We can do it all over.
Make our own plans and score the music to our own lives.”
-Set Your Goals

“It takes grindin' to be a king, it takes grindin' to be a king.”
-Mike Jones

ACKNOWLEDGEMENTS

Thank you Ferran for facilitating my first-hand participation in the scientific enterprise, for sharing your passion for science in the most pragmatic of ways, and for always keeping my focus on the bigger picture when I got lost in the details.

Thank you to my committee members Melissa, Rajeev, and Aindrila for providing pivotal feedback and guidance throughout this journey.

Thank you to everyone who graced the FGP lab: Daniela, Brandon, Ana, Sergio, Daniel, Vanessa, Blake, Corey, Sam, Kira, Julie, Huansheng, Luis, and Thuong for the cathartic camaraderie and spirited lunchtime exchanges.

Thank you to all of the friends I've made at ASU, especially the MCEats crew, the ol' MGSA, and the church of Rajeevsus. I'll forever cherish the stories, laughs, and breakthroughs we shared over beers.

Thank you to all of my friends from the previous chapters of my life, especially the A-town click, the Mandudes, and the SIU bois, I wouldn't have made it this far if not for the eras of delinquency we went through.

Thank you to my family for being consistently supportive through my lengthy educational career and providing for me well into adulthood.

Thank you Bianca. Your companionship has been both cathartic and pragmatic through this academic adventure, and I can't wait to see what life holds next for us and our dear Caesar.

TABLE OF CONTENTS

	Page
LIST OF TABLES	viii
LIST OF FIGURES	xi
CHAPTER	
1- INTRODUCTION	1
Cyanobacteria and the Genus <i>Nostoc</i>	1
Bacterial Secondary Metabolites	4
Bacterial Responses to Ultraviolet Radiation: Microbial Sunscreens	9
Approach	13
Dissertation Structure	14
References	15
Tables and Figures	21
2- THE WIDELY CONSERVED <i>EBO</i> CLUSTER IS INVOLVED IN PRECURSOR TRANSPORT TO THE PERIPLASM DURING SCYTONEMIN SYNTHESIS IN <i>NOSTOC PUNCTIFORME</i>	24
Abstract	25
Introduction	26
Materials and Methods	28
Results	33
Discussion	38
References	43

Figures	47
3- THE FUNCTIONAL LANDSCAPE AND NEIGHBOR DETERMINING GENOMIC REGION SEARCH (FLANDERS) BIOINFORMATIC DATA MINING PACKAGE/PIPELINE AND ITS USE TO PROBE FUNCTIONS OF THE <i>EBO</i> GENE CLUSTER AMONG BACTERIA	53
Abstract	54
Introduction	54
Materials and Methods	59
Results	64
Discussion	67
References	73
Tables and Figures	76
4- BET HEDGING SUNSCREEN PRODUCTION AND MOTILITY RESPONSES AGAINST UV EXPOSURE IN A CYANOBACTERIUM	85
Abstract	86
Introduction	87
Materials and Methods	90
Results	96
Discussion	103
References	110
Tables and Figures	116
5- CONCLUSIONS	124
References	127

CHAPTER	Page
REFERENCES	128

APPENDIX

A. CELLULAR AND MOLECULAR CHARACTERIZATION OF <i>EBO</i> GENE CLUSTER KNOCKOUT MUTANTS IN THE SOIL BACTERIUM <i>PSEUDOMONAS FLUORESCENS</i> NZ17	141
B. SUPPLEMENTARY MATERIALS FOR THE WIDELY CONSERVED <i>EBO</i> CLUSTER IS INVOLVED IN PRECURSOR TRANSPORT TO THE PERIPLASM DURING SCYTONEMIN SYNTHESIS IN <i>NOSTOC PUNCTIFORME</i>	154
C. SUPPLEMENTARY MATERIALS FOR THE FUNCTIONAL LANDSCAPE AND NEIGHBOR DETERMINING GENOMIC REGION SEARCH (FLANDERS) BIOINFORMATIC DATA MINING PACKAGE/PIPELINE AND ITS USE TO PROBE FUNCTIONS OF THE <i>EBO</i> GENE CLUSTER AMONG BACTERIA	163
D. SUPPLEMENTARY MATERIALS FOR BET HEDGING SUNSCREEN PRODUCTION AND MOTILITY RESPONSES AGAINST UV EXPOSURE IN A CYANOBACTERIUM	165

LIST OF TABLES

Table	Page
1. Selected Bacterial Secondary Metabolites	21
2. Selected Results of Multigene BLAST Search for <i>ebo</i> Cluster Homologs and FLANDERS Assessment of Neighboring Gene Clusters	78
3. <i>Nostoc punctiforme</i> Strains Utilized in This Work	116
4. Primers Used in This Work for Mutant Construction and Validation	159
5. Primers Used in This Work for RT-PCR Assessment of Polar Transcriptional Effects	160
6. Complete Results of the Present Work	164
7. PCR Primers Utilized in This Work	166
8. Plasmids Utilized in this Work	168

LIST OF FIGURES

Figure		Page
1.	Microscopy Images of <i>Nostoc punctiforme</i>	23
2.	The Scytonemin Operon in <i>Nostoc punctiforme</i> ATCC 29133	23
3.	Structures of the Scytonemin Monomer, and Scytonemin	47
4.	Absorbance Spectra of Wild-Type and <i>scy</i> and <i>ebo</i> Mutant Strains	48
5.	Genomic Organization of the Scytonemin Operon in <i>N. punctiforme</i> and Other Cyanobacteria	49
6.	Separation and Characterization of a Compound Accumulated After UVA Induction by $\Delta eboC$ Strain	49
7.	Confocal fluorescence Imaging and Quantification of the Scytonemin Monomer Accumulation <i>in-vivo</i>	50
8.	Intracellular Localization of the Scytonemin Monomer in Induced $\Delta scyE$ and $\Delta eboC$ Cells	51
9.	Partitioning of Core Biosynthetic Proteins of the Scytonemin Operon between Cytoplasm and Periplasm by Osmotic Shock Lysate Proteomics	52
10.	Arrangement of the <i>ebo</i> Gene Cluster in Known Biosynthesis Context.....	76
11.	Outline of the FLANDERS Pipeline	77
12.	Number of Signal Peptide Containing ORFs Upstream and Downstream of <i>ebo</i> Cluster Homologs Surveyed	83
13.	Frequency Distribution of Complete Clusterblast Hits	84

Figure	Page
14. Numerical Overview of Differentially Regulated Genes Between <i>N. punctiforme</i> and Scytonemin Two-Component Regulatory System Deletion Mutants	117
15. Subset of Differentially Expressed Genes from RNA Sequencing of <i>N. punctiforme</i> Wild Type (UCD 153), ΔRR , and ΔReg Under UV-A and Standard White Light Conditions	118
16. Phenotypic Responses of Wild Type, $\Delta hcyA$, $\Delta hcyC$ After 5 Days in Dark, High Light ($100 \mu\text{mol m}^{-2} \text{s}^{-1}$) + UV, Low Light ($10 \mu\text{mol m}^{-2} \text{s}^{-1}$) + UV, and Dark + UV Conditions	119
17. Co-induction of Scytonemin Production and Hormogonia Differentiation in Wild Type, $\Delta hcyA$, and $\Delta hcyC$ Mutants	120
18. Expression of Scytonemin Biosynthetic (<i>scyA</i>) and Regulatory (<i>scyHK</i>) Genes, as Well as <i>hcyA</i> and <i>hcyD</i> , Over a Phototactic Hormogonium Induction Time Course	121
19. Expression Changes Relative to Time 0 of Hormogonia-Associated Genes <i>gvpA</i> and <i>pilM</i> in Wild Type (WT), ΔRR , ΔReg , $\Delta hcyA$, and $\Delta hcyC$ Mutants Under Standard Scytonemin Inductive Conditions	122
20. Working Model of The Scytonemin-Hormogonium Gene Regulatory Network. The Extent of The Canonical Networks for Each is Shown Over Colored Background	123

Figure	Page
21. <i>C. elegans</i> Predation on Mixed Cultures of Wild-Type (WT) <i>P. fluorescens</i> NZI7, Δ <i>eboC</i> (<i>eboC</i>), and Δ <i>edbA</i> (<i>edb</i>)	144
22. <i>C. elegans</i> Predation on Killed Wild-Type (WT), Δ <i>eboA</i> , Δ <i>eboC</i> , and Δ <i>edbA</i> Cells	146
23. Chromatograms of HPLC Separation of Organic Solvent-Based Extractions of Wild-Type (NZI7), <i>edbA</i> , <i>eboD</i> , <i>eboC</i> , and <i>eboA</i> Mutants	150
24. Chromatograms of HPLC Separation of Methanol-Based Extractions of Wild-Type (NZI7) and <i>edbA</i> and <i>eboD</i> Mutants	151
25. Infrared Spectra of Wild-Type (NZI7), <i>edbA</i> and <i>eboA</i> Mutant Dried Cell Pellets	151
26. SDS PAGE Gel of Wild-Type (NZI7), <i>edbA</i> , <i>eboC</i> , and <i>eboA</i> Mutants ..	152
27. HPLC Chromatograms of <i>ebo</i> and <i>scyE</i> Deletion Mutant Acetone Extracts, After UVA Induction, Exhibiting Novel Compound at 7.8 Minutes and HPLC Chromatograms of Acetone Extracts from Uninduced Cells of <i>ebo</i> and <i>scyE</i> Deletion Mutants	155
28. Electrospray Ionization Time-of-Flight Mass Spectroscopy Analysis in Negative Ion Mode of Blank HPLC Solvent, Δ <i>eboC</i> Compound Eluting at 7.8 Minutes in HPLC, After Eluent Collection	156
29. Intracellular Localization of the Scytonemin Monomer in UVA Induced <i>ebo</i> Mutants by Overlay of 665 nm Over 410 nm Emission Images	156

Figure	Page
30. Overlay Images of 665 nm and 410 nm Emission Channels and DIC Channel	157
31. Fluorescence Intensity of the Scytonemin Monomer in Cytoplasm and Periplasm of <i>ΔscyE</i> cells	157
32. I-TASSER Predicted Structures of EboA and EboC	158
33. Genes Upregulated in Δ Reg Compared to Wild type (WT) and Δ R _{RR} Under Scytonemin Inductive and Non-Inductive Conditions	169
34. Plate Spreading Motility Assay of Wild Type, Δ <i>hcyA</i> , and Δ <i>hcyC</i>	170
35. Bacterial Two-Hybrid Interaction Assay Between HcyA and HcyC	171
36. HPLC Chromatograms at 384 nm to Assess Scytonemin Production Capability of Mutant Strains Used in This Study	172

1 - INTRODUCTION

I present a dissertation whose body of work aimed to elucidate particular molecular mechanisms behind the synthesis of scytonemin, a secondary metabolite used by cyanobacteria to protect themselves from harmful UV-A radiation. A review of the literature showed important gaps in our knowledge of its biosynthesis and regulation. Some of these were experimentally addressed here, guided by specific objectives. The objectives of the present dissertation were to i) investigate the genetic and cellular basis for the yet unclear latter stages of scytonemin biosynthesis, beyond the formation of the scytonemin monomer, ii) to determine the role of a cluster of conserved genes in the cyanobacterial scytonemin operon known as the *ebo* gene cluster, both regarding scytonemin and also more generally among bacteria, and, iii) to investigate the breadth of impact of a two-component regulatory system associated with the regulation of scytonemin biosynthetic pathway as a more general sensory/regulatory system in the adaptation of *Nostoc* to UV-A irradiation.

To attain these goals, I deployed a variety of genetic and biochemical techniques on the model organism *Nostoc punctiforme* ATCC 29133, which was chosen due to its tractability to genetic manipulations, its capability to differentiate into specialized cell types, and its production of the UV-A protective sunscreen scytonemin.

Cyanobacteria and the genus *Nostoc*

Few organisms have shaped the biosphere quite like the cyanobacteria, a phylum of oxygenic photosynthetic prokaryotes. Over their 3.5 billion year history on Earth,

through a phototrophic metabolism that utilizes water as an electron donor to produce molecular oxygen, they have

oxygenated the atmosphere, allowing for the evolution of the multicellular organisms that inhabit the world today (1). Presently cyanobacteria occupy nearly all terrestrial and aquatic niches and some can be found in highly inhospitable environments such as hot springs, the surface of desert soils, and the dry valleys of Antarctica (2, 3).

Cyanobacteria are a morphologically highly diverse phylum that includes forms ranging from unicellular cocci to long branching filaments composed of linked rods, with differentiation of specialized cells (4), and many forming communal, sometimes macroscopic, colonies. Cyanobacteria are also integral to biogeochemical cycling of macronutrients; while all species are capable of fixing carbon dioxide to produce glucose for aerobic respiration or growth, some can fix atmospheric nitrogen towards the synthesis of amino acids and other nitrogenous compounds, making these species crucial nodes in the biotic nitrogen cycle (5). Perhaps the hallmark of cyanobacterial biochemistry is the wealth of pigments produced by necessity of their photosynthetic lifestyle (6). Chlorophyll *a* and other chlorophylls are directly involved in the harvesting of light energy in the blue light range, while phycobilins absorb in the red, orange, yellow, and green wavelengths. In addition to the central light harvesting pigments, cyanobacteria also possess an array of carotenoids that act to regulate light-harvesting as well as in a photoprotective capacity (7–9). Still more pigments are involved solely in mitigation of ultraviolet radiation (UVR), such as scytonemin (10).

Perhaps the genus best endowed with all of the uniquely cyanobacterial traits and capabilities is *Nostoc* (11). Cells of *Nostoc* grow into unbranched filaments surrounded by a dense exopolysaccharide sheath (12). *Nostoc* cells are capable of differentiating into nitrogen fixing heterocysts (Fig. 1) that contain the nitrogenase complex which is capable of reducing molecular nitrogen to ammonia (13). Heterocysts lack photosystem II machinery, but do contain photosystem I to facilitate photophosphorylative ATP synthesis, but because heterocyst cannot actively photosynthesize they must obtain carbon sources from neighboring cells in exchange for fixed nitrogen (14). *Nostoc* filaments are also capable of differentiating into motile hormogonia (Fig. 1): short filaments composed of smaller cells that can seek out more favorable growth conditions when cells are faced with insufficient light or nutrients (15). If conditions become so inhospitable as to require more drastic measures, *Nostoc* cells can differentiate into akinetes, a semi-quiescent cell type resistant to cold and desiccation (16). This wide array of physiological capabilities allow *Nostoc* to inhabit a plethora of terrestrial niches, from the surface of desert soils (17) to near vertical limestone faces in long streaking growths known colloquially as tintenstriche (18) where wetting is extremely periodic, as well as aquatic environments that would be prohibitive to most species such as the Aldabra Atoll, where *Nostoc* can be found among the coconut groves where daily temperatures can reach over 40°C (12). Of all the species of *Nostoc*, *Nostoc punctiforme*, particularly strain ATCC 29133, has become the go-to laboratory strain for genetic and biochemical studies due to its ease of maintenance and tractability to genetic manipulations (19). The present dissertation utilized *N. punctiforme* ATCC 29133 and mutant strains derived from it to investigate the production of the UV protective pigment

scytonemin, in turn shedding light onto the process of secondary metabolite synthesis in bacteria at large.

Bacterial Secondary Metabolites

Biological systems are defined in part by their ability to transform organic compounds into cellular structures and energy through the process of central metabolism. While central metabolism and its constituent compounds, known as central metabolites, are universal in all life forms, the diversity of ecological niches that specific organisms inhabit necessitates adaptations that go beyond mere energy generation and synthesis of cellular mass. The suite of compounds synthesized that are not necessarily required for growth and reproduction through central metabolism, but rather serve a purpose of increasing fitness in a given niche are called secondary metabolites (20). Secondary metabolites are found commonly in plants, fungi, and bacteria, and generally play a role in antagonistic organismal or environmental interactions, although they may also support mutualistic symbioses (21). Some groups of bacterial are particularly rich in secondary metabolites, like actinobacteria and cyanobacteria. Perhaps the best-known example of the former category is antibiotics, anti-microbial compounds produced by bacteria or fungi in order to outcompete species that share their ecological niche. Antibiotics have been repurposed by humans to usher in modern medicine as we know it today (22). Antibiotics often target common, central processes of cellular machinery such as protein or cell wall synthesis. So as to avoid self-immolation, antibiotic producing strains also maintain antibiotic resistance mechanisms to safeguard against the very compounds they

produce (22). Perhaps the best known bacterially synthesized antibiotic is ampicillin (Table 1), which interferes with cell wall synthesis, like all beta-lactams, preventing bacterial growth (23). Conversely, the most conspicuous use of secondary metabolites in a mutualistic context comes from the field of plant-microbe interactions that dominate the rhizosphere; the soil microenvironment occurring in sub millimeter proximity to plant roots (24, 25). A common scenario involves exudation a secondary metabolites such as polyamines like arginine (Table 1) and putrescine (26) by the plant, that will lure nitrogen fixing bacteria to the root surface, to eventually engulf them in so-called root nodules, wherein the bacteria will be supplied with sources of carbon in return for fixed atmospheric nitrogen for the plant to assimilate (27). Secondary metabolites can also be used to improve fitness in the face of abiotic stressors by acting as antioxidants, such as beta-carotene and xanthophylls (28), iron-scavenging siderophores such as bacillibactin and enterobactin (29), and ultra-violet absorbing microbial sunscreens such as mycosporine-like amino acids and scytonemin (30). Still other secondary metabolites are involved in intraspecies signaling among cells, helping to regulate growth cycles and biofilm formation. These soluble messenger compounds, such as N-acyl homoserine lactone (Table 1), are synthesized consistently during exponential growth and diffuse away from their producing cell but eventually accumulate to such a degree that concentration gradients favor diffusion back into the cell and levels build up intercellularly (31). Once a threshold is passed, the second messenger acts as a signal to modulate cell activity often leading to initiation of pathogenicity or biofilm formation, among other processes (31).

Secondary metabolites take on a variety of structures owing to their diversity of roles, but can be broadly categorized by their constituent chemical moieties (32). The largest of secondary metabolites are non-ribosomal peptides (33). As their name suggests, these are oligo-peptides whose constituent amino acids are polymerized by specialized non-ribosomal peptide synthase (NRPS) enzymes rather than by the ribo-protein machinery of canonical protein synthesis (33). The genes coding for non-ribosomal peptide synthases can be arranged modularly allowing for the synthesis of a great diversity of products from a limited number of conserved enzymes (34). Polyketides are also large secondary metabolites composed of polymerized ketones isomerized to produce a variety of polycyclic compounds (35). Like NRPSs genes, polyketide synthase genes can also be arranged modularly in the genome to produce a wide array of polyketides (36). Many members of these two classes have antibiotic activity such as vancomycin and erythromycin respectively (37). Others (Table 1) have immunosuppressive (38) or toxic such as ciclosporin and aflatoxin B₁ respectively (39) properties, when administered to animals, although their original role in the producing cells is clearly a different one. In many cases, the roles that secondary metabolites play in the species of origin remains to be ascertained.

The diversity of small molecule based secondary metabolites is even greater, and includes phenazine derivatives, indolic compounds, terpenoids, and alkaloids (25). Phenazines (Table 1) are nitrogen containing heterocyclic compounds that inhibit bacterial growth by the production of intracellular reactive oxygen species (40, 41). Example phenazines include phenazine-1-carboxylic acid (42), phenazine-1-carboxamide

(43) produced by select rhizosphere pseudomonads to inhibit the growth of phytopathogenic fungi. Biosynthesis of phenazines begins by shunting two moles of chorismic acid from the aromatic amino acid synthesis pathway, and proceeds through a series of amination, decarboxylation, and condensation steps to form phenazine's polycyclic structure (40). Indolic compounds are most often used as signaling compounds to coordinate motility, biofilm formation, and virulence (44). Like phenazine, indole synthesis begins by shunting tryptophan from amino acid metabolism, also yielding pyruvate and ammonia as byproducts (44). Terpenoids are another broad class of secondary metabolites with a diversity of biological functions such as antibacterial, antifungal, and anti-nematodal activity (45). Terpenoid biosynthesis begins by diverting the ubiquinone precursor farnesyl diphosphate or the mevalonate pathway intermediary geranyl diphosphate into a series of cyclization reactions to produce polycyclic terpenes such as geosmin, pentalenene, and epicubenol (46). Lastly, alkaloids represent an additional class of polycyclic nitrogenous compounds with a wide range of activities, including antibacterial and anti-tumor (47). Still more diversity is possible through the hybridization of two or more of these categories, for example the indole-alkaloid microbial sunscreen scytonemin (48). Each of these classes originate from precursors redirected from central metabolism by action of enzymes that are highly specialized to the biosynthesis of their product secondary metabolite. The genes encoding these biosynthetic pathways are generally arranged in contiguous operons to allow for their coregulation and the maintenance of optimal stoichiometric ratios (49, 50). The enzymatic reactions catalyzed by these gene products are as varied as their resultant products often including redox, hydrolysis, and condensation transformations of

compounds diverted from central metabolism into secondary metabolite synthesis (49). Unlike central metabolites, which are relegated to the cytoplasmic space, many secondary metabolites are secreted and function extracellularly (51). To this end, the synthesis of extracellularly active compounds is often partitioned into cytoplasmic and periplasmic phases in gram-negative bacteria (51). Furthermore, because secondary metabolites perform such specialized functions, expression of their biosynthetic enzymes must be tightly controlled to avoid expending cellular energy and resources on a product irrelevant to an organism's present environmental context (22). To this end, secondary metabolite production is often controlled by bacterial two component regulatory systems; systems able to sense environmental stimuli and relay a signal to DNA-binding transcription factors to precisely affect the transcription of relevant biosynthetic machinery. For instance, the plant commensal rhizosphere bacterium *Pseudomonas fluorescens* PF-5 utilizes the two component system composed of the sensor kinase ApdA and response regulator GacA to control the synthesis of a range of anti-fungal compounds in response to plant signaling compounds (52). Another means of controlling secondary metabolite synthesis is catabolite repression wherein the cell uses ambient concentrations of some metabolite to modulate gene expression. The antibiotic actinomycin is synthesized by several *Streptomyces* species, and expression of its producing enzyme phenoxine synthase is strongly repressed by presence of glucose in the media (53). In a case where secondary metabolites themselves can regulate the synthesis of other secondary metabolites, quorum sensing utilizing second messengers can achieve a population density dependent initiation of secondary metabolite synthesis (54). For example, in several species of the plant pathogen *Erwinia*, the synthesis of the antibiotic

carbapenem is controlled by the quorum sensing of ambient levels of the auto-inducer N-(3-oxohexanoyl)-L-homoserine lactone (31). Together with a vast diversity of central metabolic pathways, secondary metabolites allow near universal niche exploitation by bacteria and represent a rich source of pharmaceutically and industrial useful natural products.

Bacterial Responses to Ultraviolet Radiation: Microbial Sunscreens

Microorganisms are subject to a range of environmental insults including temperature, pH, and osmotic stressors, but terrestrial surface microbes especially the cyanobacteria are particularly prone to damage by ultra-violet radiation (UVR) due to their necessary exposure to solar radiation(55, 56). UVR can bring about significant collateral damage, and cyanobacteria have thus developed a suite of adaptations to mitigate its deleterious effects. UVR can be subdivided into three main categories based on wavelength and type of biological effects: UVC from 100-279 nm, UVB from 280-314 nm, and UVA from 315-399 nm. UV-A is of particular detriment to cyanobacteria, because it can form reactive oxygen species, and because cyanobacteria are consistently evolving molecular oxygen as a metabolic byproduct. Hence, UV-A mitigation is of great importance.

Because UVB and UVC can cause direct damage to crucial cellular components such as DNA and proteins, cyanobacteria have developed a suite of techniques for DNA repair by action of photolyases and exonucleases as well as recombinant strategies, and heat shock protein-like proteins for protein repair (57). Cyanobacteria also engage in behavioral avoidance of UVR exposure by migrating into the substrate in which they are growing (58). The synthesis of antioxidants is a secondary means of UVR mitigation, as UVR

itself produces reactive oxygen species (ROS) readily in cyanobacteria due to their oxygenic metabolism (59, 60). Among cyanobacterial antioxidants are carotenoids tetraterpenoids conjugated to specialized proteins called carotenoproteins (61). The orange carotenoid protein is a carotenoprotein that plays a photoprotective role by binding to the phycobilisome, the light harvesting antennae complex, preventing it from transferring energy to photosystems in high light conditions (8). By decoupling the light harvesting complex from photosynthetic machinery under high light, ROS production can be mitigated under circumstances when electron transport in the thylakoid exceeds the rate at which electrons can be consumed through CO₂ fixation (62). A more general bacterial response to ROS is the production of specialized enzymes like superoxide dismutase that deactivates superoxide radicals (63), and catalase that cleaves H₂O₂ (64).

A simpler, more direct means of UVR mitigation is the production of UV absorbing sunscreens, which absorb incident UVR before it can reach cellular structures and cause damage (30). The mycosporine-like amino acids (MAA) are sunscreens synthesized by select cyanobacteria, as well as other bacteria and fungi based on 5-hydroxy-5-hydroxymethyl-cyclohex-1,2-ene ring with a methoxy-substituent in C2 (58, 60). The broad diversity of MAAs stem from the variable substitution of an amino compound at C3 as well as an oxo or imino moiety at C1 (67). MAAs originate from products shunted from central metabolism through the action of a sugar-phosphate cyclase and an O-methyltransferase to form the 6 carbon ring which is then conjugated with an amino acid moiety, for example glycine through the action of carbamoyl phosphate synthetase in the

case of mycosporine-Gly (68). MAAs have punctate absorbance spectra with maxima ranging between 309 and 362 nm, and their production is induced by exposure to UVB (280–315 nm). MAAs are generally relegated to the cytoplasm, but in some cases are covalently bound to extracellular polysaccharides (69). Extracellular polysaccharides are a physiological necessity for many terrestrial and freshwater cyanobacteria, such as *Nostoc commune*, and therefore it is energetically favorable to couple sunscreen and polysaccharide secretion. This extracellular positioning of MAAs also increases their effectiveness in deflecting UVR from cellular targets (70). In addition to their UVR protective capabilities, some MAAs such mycosporine-Gly can also play an antioxidative role by scavenging ROS (60). Another sunscreen compound type distributed across multiple kingdoms is the melanins, a loosely defined collection of black and brown polymeric pigments that are highly insoluble with a broad absorption spectrum from far UV to infrared (71). They can be divided into three main categories: eumelanins derived from the oxidized tyrosine product dihydroxyphenylalanine (DOPA) prior to polymerization, pheomelanins whose monomeric unit is based on a cysteinated DOPA, and allomelanins based on dihydroxynaphthalene or homogenistic acid-based monomers (71). In bacterial systems melanin mediated UVR mitigation has been experimentally determined in *Streptomyces* and *Bacillus* species (72, 73), however they may not play a solely UV-protective role because like MAAs melanins are also capable of antioxidant activity by scavenging ROS (71).

While MAAs are found in bacteria, algae and fungi that are exposed to high UVR fluxes, the sunscreen compound scytonemin is relegated exclusively to select terrestrial and

intertidal cyanobacteria (30). Scytonemin is a brownish-yellow lipid-soluble pigment that is excreted and accumulated in the extracellular matrix in response to UVA radiation exposure (Table 1) (74). Structurally unique among natural products, scytonemin is a homodimeric indole-alkaloid, composed of two heterocyclic units that are symmetrically connected through a carbon-carbon bond (48). The complex ring structure allows for strong absorption in the UVA-violet-blue range (325-425 nm), with a maximum around 370 nm *in vivo* (48, 74). A genomic region comprising 18 contiguous ORFs, (Npun_R1276 to Npun_R1259) was found to be responsible for scytonemin biosynthesis in *Nostoc punctiforme* ATCC 29133 (75). Six consecutive genes, named *scyABCDEF* (76) conform the core biosynthetic locus. *In vitro* studies confirmed that ScyA, ScyB and ScyC carry out the early stages of the scytonemin assembly: ScyB first catalyzes the oxidative deamination of L-tryptophan to yield indole-3 pyruvic acid, while ScyA mediates the acyloin coupling of the indole-3 pyruvic acid to *p*-hydroxyphenylpyruvic acid, producing a labile β -ketoacid. Subsequently, ScyC catalyzes the cyclization and decarboxylation of the previous compound to form a ketone, which is one oxidation state away from the potential scytonemin monomeric precursor (77). The aromatic precursors are supplied by a set of redundant orthologues coding for enzymes in the aromatic amino acids biosynthetic and shikimic acid pathways, which are located downstream of the core biosynthetic genes in *Nostoc punctiforme* (75, 78). A two-component regulatory system upstream controls the expression of the entire operon (79). While it would logically follow that the rest of the core genes (*scyDEF*) could catalyze the final oxidative dimerization of scytonemin monomers, two lines of evidence indicated that this is not the case: (i) of the three genes, only *scyE* is essential for scytonemin synthesis (80), and (ii)

expression of the *scyA-E* locus as well as the entire 18-gene cluster in *Escherichia coli* is insufficient to induce heterologous scytonemin production (81). Comparative genomics revealed the presence of an additional group of five highly conserved genes of unknown function within the *scy* operon among scytonemin-producing cyanobacteria. These genes are annotated as a hypothetical protein, a TatD like hydrolase, a UbiA prenyltransferase, a 5-dehydroquinate synthase, an AP endonuclease, and an alkaline phosphatase respectively. In the genome of *N. punctiforme* ATCC 29133, however, these were found at a distal locus and dubbed “satellite” genes (Fig. 2).

Surprisingly, this satellite cluster was found to not only be associated with scytonemin synthesis clusters, but also in over 100 species of bacteria as well as the plastid genomes of some eustigmatophyte algae (82). Denoted as the Eustigmatophyte-bacterial operon, or *ebo* gene cluster, the discovery of the former ‘satellite’ cluster across non scytonemin producing bacteria prompted a flurry of questions. A literature search for clusters bearing similar annotation to the *ebo* cluster revealed an instance of its conservation in the soil bacterium *Pseudomonas fluorescens* NZI7, where it is involved in the repellence of nematode predation (83). The *ebo* cluster and its role in scytonemin production was the nucleation point of the present thesis, and its broader role in secondary metabolite synthesis became evident through a series of serendipitous if not for confounding experiments, analyses, and correspondences.

Approach

Previous studies had determined the role of genes *scyA-C* in-vivo and heterologously, but none had satisfactorily described a role for *scyDEF*. In the interest of furthering understanding about the scytonemin synthesis pathway, I utilized knock out mutants of the latter scytonemin synthesis genes (*scyDEF*) as well as knock out mutants of each of the five satellite genes along with a complete cluster knockout. HPLC, LC-MS, and microscopic analyses of these knock out mutants were used to elucidate a potential mechanism for the penultimate steps in scytonemin synthesis. To assess the reach of influence of the scytonemin two-component regulatory system, RNA sequencing of two-component system knock-out mutants under UV irradiation was conducted and analyzed for transcriptional derivations from wild-type. The association of the *ebo* gene cluster with non-scytonemin biosynthetic gene clusters was conducted through the development and application of the Functional Landscape And Neighbor Determining gEnomic Region Search (FLANDERS) analysis pipeline.

Dissertation structure

The present work is comprised of this introductory chapter, followed by three research chapters in the format of published or published manuscripts respectively, as well as concluding remarks. Appendix A is an overview of further experiments conducted which did not fit into any of the three chapters, and appendices B, C, and D are supplemental materials for chapters 2, 3, and 4 respectively.

References

1. Whitton BA, Potts M. 2012. Introduction to the cyanobacteria. *Ecol Cyanobacteria II Their Divers Sp Time*. Springer Netherlands.
2. Rossi F, Potrafka RM, Pichel G, De Philippis R. 2012. The role of the exopolysaccharides in enhancing hydraulic conductivity of biological soil crusts. *Soil Biol Biochem* 46:33–40.
3. Quesada A, Vincent WF. 1997. Strategies of adaptation by antarctic cyanobacteria to ultraviolet radiation. *Eur J Phycol* 32:335–342.
4. Castenholz RW. 1992. Species Usage, Concept, and Evolution in the Cyanobacteria (Blue-Green Algae). *J Phycol* 28:737–745.
5. Bryant DA. 1994. *The Molecular Biology of Cyanobacteria*. Springer Netherlands.
6. Saini DK, Pabbi S, Shukla P. 2018. Cyanobacterial pigments: Perspectives and biotechnological approaches. *Food Chem Toxicol* 120:616–624.
7. Niedzwiedzki DM, Swainsbury DJK, Canniffe DP, Neil Hunter C, Hitchcock A. 2020. A photosynthetic antenna complex foregoes unity carotenoid-to-bacteriochlorophyll energy transfer efficiency to ensure photoprotection. *Proc Natl Acad Sci U S A* 117:6502–6508.
8. Zhang H, Liu H, Niedzwiedzki DM, Prado M, Jiang J, Gross ML, Blankenship RE. 2014. Molecular mechanism of photoactivation and structural location of the cyanobacterial orange carotenoid protein. *Biochemistry* 53:13–19.
9. Hayashi Y, Ito T, Yoshimura T, Hemmi H. 2018. Utilization of an intermediate of the methylerythritol phosphate pathway, (*E*)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate, as the prenyl donor substrate for various prenyltransferases. *Biosci Biotechnol Biochem* 82:993–1002.
10. Garcia-Pichel F, Castenholz RW. 1991. Characterization And Biological Implications of Scytonemin, A Cyanobacterial Sheath Pigment. *J Phycol* 27:395–409.
11. Potts M. 2006. Nostoc, p. 465–504. *In The Ecology of Cyanobacteria*. Kluwer Academic Publishers.
12. Dodds WK, Gudder DA, Mollenhauer D. 1995. The Ecology of Nostoc. *J Phycol* 31:2–18.
13. Wolk CP, Ernst A, Elhai J. 1994. Heterocyst Metabolism and Development, p. 769–823. *In The Molecular Biology of Cyanobacteria*. Springer Netherlands.
14. Chailakhyan LM, Glagolev AN, Glagoleva TN, Murvanidze G V., Potapova T V., Skulachev VP. 1982. Intercellular power transmission along trichomes of

- cyanobacteria. *BBA - Bioenerg* 679:60–67.
15. Marsac NT. 1994. Differentiation of Hormogonia and Relationships with Other Biological Processes, p. 825–842. *In The Molecular Biology of Cyanobacteria*. Springer Netherlands.
 16. Adams DG, Duggan PS. 1999. Heterocyst and akinete differentiation in cyanobacteria. *New Phytol*. Cambridge University Press.
 17. Mazor G, Kidron GJ, Vonshak A, Abeliovich A. 2006. The role of cyanobacterial exopolysaccharides in structuring desert microbial crusts. *FEMS Microbiol Ecol* 21:121–130.
 18. Lüttge U. 1997. Cyanobacterial Tintenstrich Communities and their Ecology. 526 *Naturwissenschaften* 84:526–534.
 19. Meeks JC, Elhai J, Thiel T, Potts M, Larimer F, Lamerdin J, Predki P, Atlas R. 2001. An overview of the Genome of *Nostoc punctiforme*, a multicellular, symbiotic Cyanobacterium. *Photosynth Res* 70:85–106.
 20. Bentley R. 1999. Secondary metabolite biosynthesis: The first century. *Crit Rev Biotechnol* 19:1–40.
 21. Pichersky E, Gang DR. 2000. Genetics and biochemistry of secondary metabolites in plants: An evolutionary perspective. *Trends Plant Sci* 5:439–445.
 22. Thapa SS, Grove A. 2019. Do Global Regulators Hold the Key to Production of Bacterial Secondary Metabolites? *Antibiotics* 8:160.
 23. Rafailidis PI, Ioannidou EN, Falagas ME. 2007. Ampicillin/sulbactam: Current status in severe bacterial infections. *Drugs* 67:1829–1849.
 24. Vining LC. 1990. Functions of Secondary Metabolites. *Annu Rev Microbiol* 44:395–427.
 25. Tyc O, Song C, Dickschat JS, Vos M, Garbeva P. 2017. The Ecological Role of Volatile and Soluble Secondary Metabolites Produced by Soil Bacteria. *Trends Microbiol* 25:280–292.
 26. Jiménez-Bremont JF, Marina M, Guerrero-González M de la L, Rossi FR, Sánchez-Rangel D, Rodríguez-Kessler M, Ruiz OA, Gárriz A. 2014. Physiological and molecular implications of plant polyamine metabolism during biotic interactions. *Front Plant Sci* 5:1–14.
 27. Smercina DN, Evans SE, Friesen ML, Tiemann LK. 2019. To fix or not to fix: Controls on free-living nitrogen fixation in the rhizosphere. *Appl Environ Microbiol* 85:e02546-18.
 28. Guan N, Li J, Shin H dong, Du G, Chen J, Liu L. 2017. Microbial response to environmental stresses: from fundamental mechanisms to practical applications.

- Appl Microbiol Biotechnol 101:3991–4008.
29. Hider RC, Kong X. 2010. Chemistry and biology of siderophores † ‡. *Nat Prod Rep* 27:637–657.
 30. Gao Q, Garcia-Pichel F. 2011. Microbial ultraviolet sunscreens. *Nat Rev Microbiol* 9:791–802.
 31. Barnard AML, Bowden SD, Burr T, Coulthurst SJ, Monson RE, Salmond GPC. 2007. Quorum sensing, virulence and secondary metabolite production in plant soft-rotting bacteria. *Philos Trans R Soc B Biol Sci* 362:1165–1183.
 32. Stone MJ, Williams DH. 1992. On the evolution of functional secondary metabolites (natural products). *Mol Microbiol* 6:29–34.
 33. Challis GL, Naismith JH. 2004. Structural aspects of non-ribosomal peptide biosynthesis. *Curr Opin Struct Biol* 14:748–756.
 34. Stachelhaus T, Marahiel MA. 1995. Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis. *FEMS Microbiol Lett* 125:3–14.
 35. Zhang L, Hashimoto T, Qin B, Hashimoto J, Kozono I, Kawahara T, Okada M, Awakawa T, Ito T, Asakawa Y, Ueki M, Takahashi S, Osada H, Wakimoto T, Ikeda H, Shin-ya K, Abe I. 2017. Characterization of Giant Modular PKSs Provides Insight into Genetic Mechanism for Structural Diversification of Aminopolyol Polyketides. *Angew Chemie* 129:1766–1771.
 36. Chen H, Du L. 2016. Iterative polyketide biosynthesis by modular polyketide synthases in bacteria. *Appl Microbiol Biotechnol* 100:541–557.
 37. Baltz RH. 2006. Molecular engineering approaches to peptide, polyketide and other antibiotics. *Nat Biotechnol* 24:1533–1540.
 38. Forsythe P, Paterson S. 2014. Ciclosporin 10 years on: Indications and efficacy. *Vet Rec* 174:13–21.
 39. Smela ME, Currier SS, Bailey EA, Essigmann JM. 2001. The chemistry and biology of aflatoxin B1: From mutational spectrometry to carcinogenesis. *Carcinogenesis* 22:535–545.
 40. Mavrodi D V., Peever TL, Mavrodi O V., Parejko JA, Raaijmakers JM, Lemanceau P, Mazurier S, Heide L, Blankenfeldt W, Weller DM, Thomashow LS. 2010. Diversity and evolution of the Phenazine Biosynthesis Pathways. *Appl Environ Microbiol* 76:866–879.
 41. Blankenfeldt W, Parsons JF. 2014. The structural biology of phenazine biosynthesis. *Curr Opin Struct Biol* 29:26–33.
 42. Saleh O, Gust B, Boll B, Fiedler HP, Heide L. 2009. Aromatic prenylation in

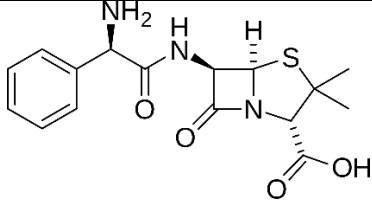
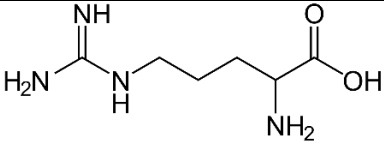
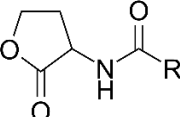
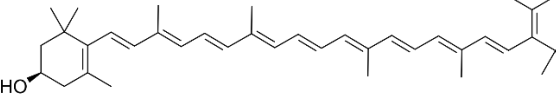
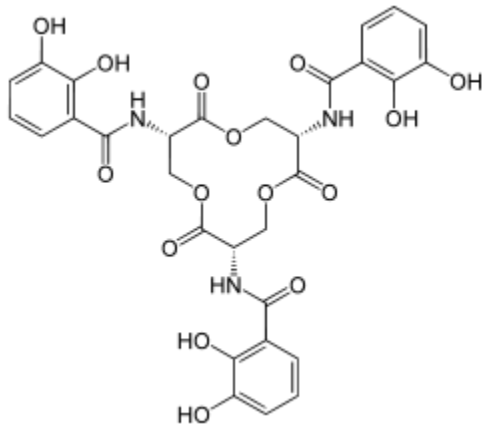
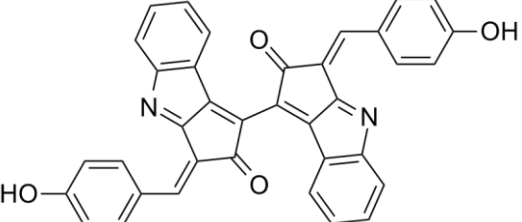
- phenazine biosynthesis: Dihydrophenazine-1-carboxylate dimethylallyltransferase from streptomyces anulatus. *J Biol Chem* 284:14439–14447.
43. Narsing Rao MP, Xiao M, Li WJ. 2017. Fungal and bacterial pigments: Secondary metabolites with wide applications. *Front Microbiol* 8:1113.
 44. Melander RJ, Minvielle MJ, Melander C. 2014. Controlling bacterial behavior with indole-containing natural products and derivatives. *Tetrahedron* 70:6363–6372.
 45. Gershenzon J, Dudareva N. 2007. The function of terpene natural products in the natural world. *Nat Chem Biol* 3:408–414.
 46. Dickschat JS. 2016. Bacterial terpene cyclases. *Nat Prod Rep* 33:87–110.
 47. Rathbone DA, Bruce NC. 2002. Microbial transformation of alkaloids. *Curr Opin Microbiol* 5:274–281.
 48. Proteau PJ, Gerwick WH, Garcia-Pichel F, Castenholz R. 1993. The structure of scytonemin, an ultraviolet sunscreen pigment from the sheaths of cyanobacteria. *Experientia* 49:825–829.
 49. Walsh CT, Fischbach MA. 2010. Natural Products Version 2.0: Connecting Genes to Molecules Introduction: Why Do Natural Products Still Matter? *J Am Chem Soc* 132:2469–2493.
 50. Gokulan K, Khare S, Cerniglia C. 2014. METABOLIC PATHWAYS | Production of Secondary Metabolites of Bacteria, p. 561–569. *In Encyclopedia of Food Microbiology*. Elsevier.
 51. Sidebottom AM, Carlson EE. 2015. A reinvigorated era of bacterial secondary metabolite discovery. *Curr Opin Chem Biol* 24:104–111.
 52. Corbell N, Loper JE. 1995. A global regulator of secondary metabolite production in *Pseudomonas fluorescens* Pf-5. *J Bacteriol* 177:6230–6236.
 53. Gallo M, Katz E. 1972. Regulation of Secondary Metabolite Biosynthesis: Catabolite Repression of Phenoxazinone Synthase and Actinomycin Formation by Glucose. *J Bacteriol* 109:659–667.
 54. Rutherford ST, Bassler BL. 2012. Bacterial quorum sensing: Its role in virulence and possibilities for its control. *Cold Spring Harb Perspect Med* 2:a012427.
 55. Blatchley ER, Dumoutier N, Halaby TN, Levi Y, Laîné JM. 2001. Bacterial responses to ultraviolet irradiation, p. 179–186. *In Water Science and Technology*. IWA Publishing.
 56. Rastogi RP, Sinha RP, Moh SH, Lee TK, Kottuparambil S, Kim Y-J, Rhee J-S, Choi E-M, Brown MT, Häder D-P, Han T. 2014. Ultraviolet radiation and cyanobacteria. *J Photochem Photobiol B Biol* 154–169.

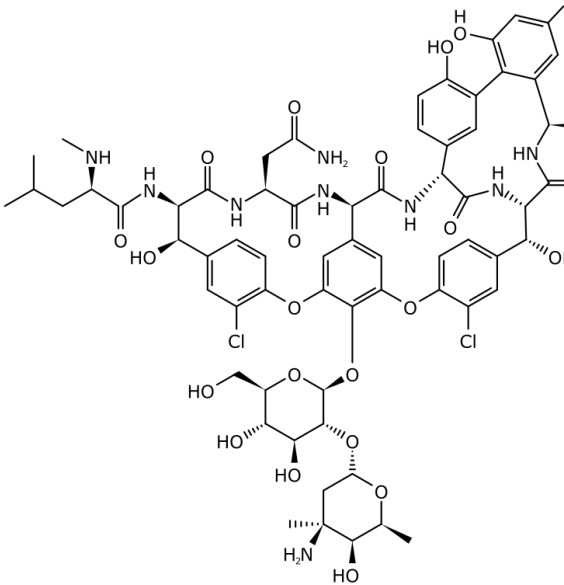
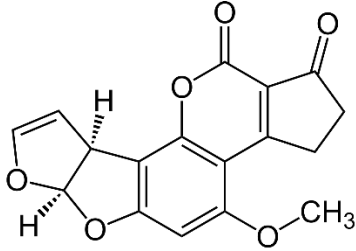
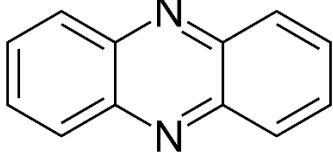
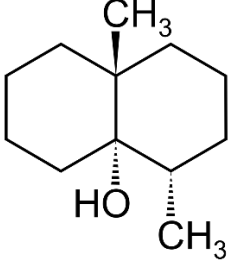
57. Rastogi RP, Richa, Kumar A, Tyagi MB, Sinha RP. 2010. Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. *J Nucleic Acids* 2010:592980.
58. Nadeau T, Howard-Williams C, Castenholz R. 1999. Effects of solar UV and visible irradiance on photosynthesis and vertical migration of *Oscillatoria* sp. (Cyanobacteria) in an Antarctic microbial mat. *Aquat Microb Ecol* 20:231–243.
59. Singh SP, Hä Der D-P, Sinha RP. Cyanobacteria and ultraviolet radiation (UVR) stress: Mitigation strategies. *Ageing Res Rev* 9:79–90.
60. Oren A, Gunde-Cimerman N. 2007. Mycosporines and mycosporine-like amino acids: UV protectants or multipurpose secondary metabolites? *FEMS Microbiol Lett* 269:1–10.
61. Britton G, Armit GM, Lau SYM, Patel AK, Shone CC. 1982. Carotenoproteins, p. 237–251. *In* Carotenoid Chemistry and Biochemistry. Elsevier.
62. Latifi A, Ruiz M, Zhang CC. 2009. Oxidative stress in cyanobacteria. *FEMS Microbiol Rev* 33:258–278.
63. McCord JM, Fridovich I. 1969. Superoxide Dismutase. *J Biol Chem* 244:6049–6055.
64. Fita I, Rossmann MG. 1985. The active center of catalase. *J Mol Biol* 185:21–37.
65. Karsten U, Garcia-Pichel F. 1996. Carotenoids and Mycosporine-like Amino Acid Compounds in Members of the Genus *Microcoleus* (Cyanobacteria): A Chemosystematic Study. *Syst Appl Microbiol* 19:285–294.
66. Garcia-Pichel F, Wingard CE, Castenholz RW. 1993. Evidence regarding the ultraviolet sunscreen role of the Mycosporin-like compounds in the cyanobacterium *Gloeocapsa* sp. *Appl Environ Microbiol* 59:170–176.
67. Favre-Bonvin J, Bernillon J, Salin N, Arpin N. 1987. Biosynthesis of mycosporines: Mycosporine glutaminol in *Trichothecium roseum*. *Phytochemistry* 26:2509–2514.
68. Takanoa S, Uemurab D, Hiratab Y. 1977. Isolation and Structure of A Mycosporine from the Zoanthid *Palythoa Tuberculosa*. *Tetrahedron Lett* 28:2429–2430.
69. Bohm G, Pflleiderer W, Schere S. Structure of a Novel Oligosaccharide-Mycosporine-Amino Acid Ultraviolet A/B Sunscreen Pigment from the Terrestrial Cyanobacterium *Nostoc commune*. *J Biol Chem* 270:8536–8539.
70. Ehling-Schulz M, Scherer S. 1999. UV protection in cyanobacteria. *Eur J Phycol* 34:329–338.
71. Plonka PM, Grabačka M. 2006. Melanin synthesis in microorganisms-

- biotechnological and medical aspects *. *Acta Biochim Pol* 53:429–443.
72. Romero-Martinez R, Wheeler M, Guerrero-Plata A, Rico G, Torres-Guerrero H. 2000. Biosynthesis and functions of melanin in *Sporothrix schenckii*. *Infect Immun* 68:3696–3703.
 73. Funa N, Funabashi M, Ohnishi Y, Horinouchi S. 2005. Biosynthesis of hexahydroxyperylenequinone melanin via oxidative aryl coupling by cytochrome P-450 in *Streptomyces griseus*. *J Bacteriol* 187:8149–8155.
 74. Garcia-Pichel F, Castenholz RW. 1991. Characterization and Biological Implications of Scytonemin, a Cyanobacterial Sheath Pigment. *J Phycol* 27:395–409.
 75. Soule T, Stout V, Swingley WD, Meeks JC, Garcia-Pichel F. 2007. Molecular genetics and genomic analysis of scytonemin biosynthesis in *Nostoc punctiforme* ATCC 29133. *J Bacteriol* 189:4465–72.
 76. Soule T, Garcia-Pichel F, Stout V. 2009. Gene expression patterns associated with the biosynthesis of the sunscreen scytonemin in *Nostoc punctiforme* ATCC 29133 in response to UVA radiation. *J Bacteriol* 191:4639–4646.
 77. Balskus EP, Walsh CT. 2009. An Enzymatic Cyclopentyl[b]indole Formation Involved in Scytonemin Biosynthesis. *J Am Chem Soc* 131:14648–14649.
 78. Soule T, Palmer K, Gao Q, Potrafka R, Stout V, Garcia-Pichel F. 2009. A comparative genomics approach to understanding the biosynthesis of the sunscreen scytonemin in cyanobacteria. *BMC Genomics* 10:336.
 79. Naurin S, Bennett J, Videau P, Philmus B, Soule T. 2016. The response regulator Npun_F1278 is essential for scytonemin biosynthesis in the cyanobacterium *Nostoc punctiforme* ATCC 29133. *J Phycol* 52:564–571.
 80. Ferreira D, Garcia-Pichel F. 2016. Mutational Studies of Putative Biosynthetic Genes for the Cyanobacterial Sunscreen Scytonemin in *Nostoc punctiforme* ATCC 29133. *Front Microbiol* 7:1–10.
 81. Malla S, Sommer MOA. 2014. A sustainable route to produce the scytonemin precursor using *Escherichia coli* †. *Green Chem* 16:3255–3265.
 82. Yurchenko T, Ševčíková T, Strnad H, Butenko A, Eliáš M. 2016. The plastid genome of some eustigmatophyte algae harbours a bacteria-derived six-gene cluster for biosynthesis of a novel secondary metabolite. *Open Biol* 6.
 83. Burlinson P, Studholme D, Cambray-Young J, Heavens D, Rathjen J, Hodgkin J, Preston GM. 2013. *Pseudomonas fluorescens* NZI7 repels grazing by *C. elegans*, a natural predator. *ISME J* 7:1126–1138.

Tables and Figures

Table 1: Selected bacterial secondary metabolites

Secondary Metabolite	Functional Class	Mechanism of action	Structure
Ampicillin	Antibiotic	Inhibits bacterial transpeptidase, halting cell wall synthesis	
Arginine	Plant symbiosis initiation	Promote biofilm formation during plant symbiosis	
N-acyl homoserine lactone	Second messenger	Diffusion and receptor binding	
Cryptoxanthin	Antioxidant	Non-photochemical quenching under high light intensity	
Enterobactin	siderophore	Chelation of ferric iron in center of polycycle	
Scytonemin	Indole-alkaloid UV sunscreen	Absorption of energy from UV-A photon	

Vancomycin	Non-ribosomal peptide antibiotic	Intercalation of NAM/NAG peptides preventing cell wall synthesis	
Aflatoxin B1	Polyketide carcinogen	Mutation of codon 249 in <i>p53</i> tumor suppressor gene	
Phenazine	Phenazine based bacterial control for plant symbiont	Suppress growth of <i>Erwinia amylovaora</i> , preventing blight disease	
Geosmin	Terpenoid	Creates earth or musty smell of moist soil	

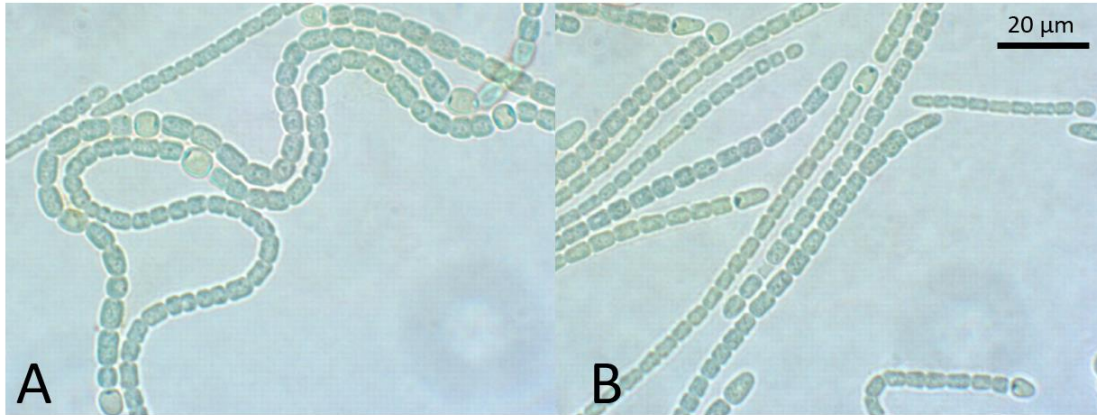


Fig. 1: Microscopy images of *Nostoc punctiforme* A. Vegetative cells of *Nostoc punctiforme* interspersed with lightly pigmented rotund heterocysts B. *Nostoc punctiforme* differentiated into hormogonia filaments

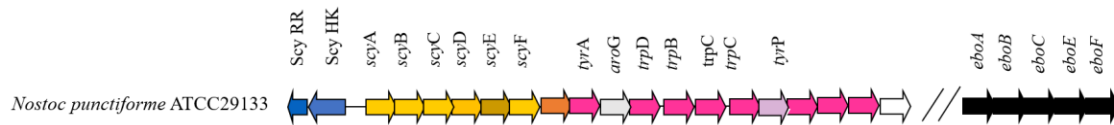


Fig. 3: Arrangement of the scytonemin operon in *Nostoc punctiforme* ATCC 29133. The scytonemin core synthesis genes are yellow, aromatic amino acid synthesis genes are pink, the scytonemin two component regulatory system composed of a histidine kinase (Scy HK) and response regulator (Scy RR) is blue, and the *ebo* gene cluster is black.

2 - THE WIDELY CONSERVED *EBO* CLUSTER IS INVOLVED IN PRECURSOR
TRANSPORT TO THE PERIPLASM DURING SCYTONEMIN SYNTHESIS IN
NOSTOC PUNCTIFORME.

Published in mBio

Klicki K, Ferreira D, Hamill D, Dirks B, Mitchell N, Garcia-Pichel F. 2018. The Widely
Conserved ebo Cluster Is Involved in Precursor Transport to the Periplasm during Scytonemin
Synthesis in *Nostoc punctiforme*. mBio 9: e02266-18

Coauthors have acknowledged the use of this manuscript in my dissertation.

Kevin Klicki^{1,2}, Daniela Ferreira¹, Demetra Hamill¹, Blake Dirks^{1,2}, Natalie Mitchell¹,
Ferran Garcia-Pichel^{1,2}

¹School of Life Sciences, Arizona State University, Tempe, Arizona 85281

²Center for Fundamental and Applied Microbiomics, Biodesign Institute, Arizona
State University, Tempe, Arizona 85281

Corresponding Author: Ferran Garcia-Pichel ferran@asu.edu

Abstract

Scytonemin is a dimeric indole-phenol sunscreen synthesized by some cyanobacteria under exposure to UVA radiation. While its biosynthetic pathway has been only partially elucidated, comparative genomics reveals that the scytonemin operon often contains a cluster of five highly conserved genes (*ebo* cluster) of unknown function that is widespread and conserved among several bacterial and algal phyla. We sought to elucidate the function of the *ebo* cluster in the cyanobacterium *Nostoc punctiforme* by constructing and analyzing in-frame deletion mutants (one for each *ebo* gene, one for the entire cluster). Under UVA induction, all *ebo* mutants were scytoneminless, and all accumulated a single compound, the scytonemin monomer, Fig. 1, clearly implicating all *ebo* genes in scytonemin production. We showed that the scytonemin monomer also accumulated in an induced deletion mutant of *scyE*, a non-*ebo* scytonemin gene whose gene product is demonstrably targeted to the periplasm. Confocal autofluorescence microscopy revealed that the accumulation was confined to the cytoplasm in all *ebo* mutants, but not so in the *scyE* deletion, with an intact *ebo* cluster, where the scytonemin monomer was also excreted to the periplasm. The results implicate the *ebo* cluster in the export of the scytonemin monomer to the periplasm for final oxidative dimerization by ScyE. By extension, the *ebo* gene cluster may play a similar role in metabolite translocation across many bacterial phyla. We discuss potential mechanisms for such role based on structural and phylogenetic considerations of the *ebo* proteins.

Introduction

The cytoplasm is an ideal environment for the synthesis of secondary metabolites, being highly regulated and rich in energetic compounds, soluble enzymes, and cofactors.

Often, however, products synthesized there must function in the periplasm or outside of the cell, necessitating transmembrane systems to facilitate their transport through the inner (or cytoplasmic) membrane. Among secondary metabolites that are synthesized in the cytoplasm but later excreted are some sunscreen compounds produced by cyanobacteria to cope with excess energy collected for photosynthesis. Some species growing on exposed surfaces are able to produce UV-absorbing sunscreens, e. g. mycosporine-like amino acids and scytonemin. These compounds intercept UV radiation thus decreasing damage to cellular machinery (1). Scytonemin, found exclusively among cyanobacteria, is a brownish-yellow, lipid-soluble pigment that is excreted and accumulated in the extracellular matrix in response to UVA radiation exposure (315-400 nm) (2-4). Structurally unique among natural products, it is a homodimeric indole-alkaloid, with a molecular mass of 544 g mol^{-1} , in its oxidized form, composed of two heterocyclic units that are symmetrically connected through a carbon-carbon bond (5). The complex ring structure allows for strong absorption in the UVA-violet-blue range (325-425 nm), with a maximum at 384 nm in acetone and around 370 nm *in vivo* (2,5) (Fig. 2A).

A genomic region comprising 18 contiguous ORFs—Npun_R1276 to Npun_R1259—was found to be responsible for scytonemin biosynthesis in *Nostoc punctiforme* ATCC 29133 (PCC 73102), their transcription being induced by UVA exposure (Fig. 3)

(6,8,14). Six consecutive genes (ORFs Npun_R1276 to Npun_R1271; named *scyABCDEF* (7)) form the core biosynthetic locus. *In-vitro* studies confirmed that ScyA, ScyB, and ScyC carry out the early stages of the scytonemin assembly: ScyB first catalyzes the oxidative deamination of L-tryptophan to yield indole-3 pyruvic acid, while ScyA mediates the acyloin coupling of the indole-3 pyruvic acid and *p*-hydroxyphenylpyruvic acid, producing a labile β -ketoacid compound (10). Subsequently, ScyC catalyzes the cyclization and decarboxylation of the previous compound to form a ketone, which is one (auto)oxidation state away from the scytonemin monomer, Fig. 1 (11). The precursors are supplied by a set of redundant orthologues coding for enzymes in the aromatic amino acid biosynthetic and shikimic acid pathways (6-8).

A two-component regulatory system controls the expression of the entire operon (9). While it would logically follow that the rest of the core genes (*scyDEF*) catalyze the final oxidative dimerization of the scytonemin synthesis, two lines of evidence indicate that this is not the case: (i) of the three, only *scyE* is essential for scytonemin synthesis (12), and (ii) expression of the *scyA-E* locus as well as the entire 18-gene cluster in *Escherichia coli* was insufficient to attain heterologous scytonemin production (13). Comparative genomics revealed an additional group of five highly conserved genes (*ebo* genes, see below) of unknown function within the *scy* operon among several cyanobacteria. In the genome of *N. punctiforme* ATCC 29133, however, these are found at a distal locus, but are also up-regulated with the scytonemin synthesis operon under UVA exposure (1,7,14) (Fig. 3). For these reasons, it was suggested that the *ebo* genes may be involved in the synthesis of scytonemin, perhaps responsible for the later biosynthetic steps (12).

Intriguingly, this five gene cluster was found to be highly conserved in synteny and sequence homology among many bacteria, across several phyla, as well as in the plastid genomes of some eustigmatophyte algae, and hence named the “eustigmatophyte/bacterial operon” or *ebo* (15). Despite an intensive bioinformatic analysis, a clear potential function for the *ebo* genes could not be put forth. However their widespread presence in an array of bacterial and plastid genomes, the overwhelming majority of which do not produce scytonemin, weakens the hypothesis that they have a dedicated role in scytonemin biosynthesis (15).

We sought to elucidate the function of the *ebo* gene cluster in *N. punctiforme* by constructing in-frame deletion mutants in relevant open reading frames, and investigating their resulting phenotypes.

Materials and Methods

Cultures and culture conditions. All experiments were conducted with a wild-type *Nostoc punctiforme* strain ATCC 29133 (PCC 73102) derivative, UCD 153, that displays a dispersed growth under shaking culture conditions and a higher frequency of gene replacement by conjugal transfer compared to the original isolate (16) and with mutants derived from it (supplementary Table S1). Wild-type and mutant strains of *N. punctiforme* were grown under standard culture conditions as previously described (12) in liquid Allen and Arnon medium (17), diluted fourfold (AA/4), and on solidified AA medium plates. When necessary, the medium was supplemented with 2.5 mM NH₄Cl buffered with 5 mM MOPS (pH 7.8). Neomycin at 25 µg ml⁻¹ was used for the selection and maintenance of transformed single recombinants. *Escherichia coli* strains and

derivatives were grown in liquid or solid Lysogeny Broth (LB)(33) supplemented with kanamycin at 25 $\mu\text{g ml}^{-1}$ and, when required, chloramphenicol at 30 $\mu\text{g ml}^{-1}$.

Construction of mutants. All chromosomal mutations in this study were in-frame deletions of individual genes, with the exception of Δebo , where the entire five-gene cluster was deleted. The deletions were generated by PCR using *N. punctiforme* genomic DNA and primers designed to amplify DNA upstream and downstream of the deletion (2.0 kb – 3.0 kb on each side to allow for homologous recombination), with the primers adjacent to the deletion containing overlapping sequences (see supplementary materials and methods and supplementary Table S1 for oligonucleotide details and gene deletions).

Screening for polar transcriptional effects. In order to verify that the *ebo* gene in-frame deletion mutations did not deleteriously effect transcription of downstream *ebo* genes, RT-PCR was used. For each mutant, we targeted the transcript of the next gene downstream under inductive conditions except for $\Delta eboF$, wherein downstream effects would be irrelevant. For this, total RNA was extracted using Molbio Powersoil total RNA extraction kit from which cDNA was synthesized using Superscript III reverse transcriptase (Thermo Fisher Scientific). Primers specific to *eboB*, *eboC*, *eboE*, and *eboF* (Supplementary table S2) were used to amplify cDNA fragments corresponding to their parent mRNA. Resulting PCR products were analyzed by gel electrophoresis against reactions run with wild-type genomic DNA as a synthesis template.

Biochemical characterization of mutant strains. Cells from *N. punctiforme* wild-type and derived deletion mutants were tested for their ability to produce scytonemin upon

induction by UVA radiation. Following UVA exposure, the cells were harvested and the lipid-soluble pigments were extracted in equal volumes of 100% acetone. Extracts were initially analyzed spectrophotometrically between 330 nm to 730 nm, a strong absorption peak at 384 nm indicating that scytonemin had accumulated in the cells (2). Following UVA exposure, water-soluble compounds were also extracted from whole cells in equal volumes of 25% aqueous methanol. 50 μ l of concentrated acetone or methanol extracts from cells exposed to UVA radiation were also analyzed by high pressure liquid chromatography (see supplementary materials and methods). Unknown peaks of interest were sourced from five independent biological replicates and collected with a Gilson FC 205 Fraction Collector. Exact masses of collected compounds were analyzed by electrospray ionization mass spectrometry using a Bruker Daltonics micrOTOF-Q in the positive and negative ion modes. Fractions collected from the wild-type at the same retention time and solvents utilized to run the HPLC were used as negative controls for MS, so as to only consider peaks found exclusively in the fractions collected for the mutants. Authentic standards for the scytonemin monomer (Fig. 1) were obtained by heterologous expression of *scy* genes in *E. coli* following Malla and Sommer (13) (see supplementary materials and methods). The scytonemin monomer (Fig. 1) was purified, characterized by spectra, retention time and mass to serve as an authentic standard. Additionally, steady-state fluorescence spectra of the scytonemin monomer (Fig. 1) were obtained using a LS-55 Fluorescence Spectrometer (PerkinElmer Inc., Waltham, MA) equipped with a red-sensitive R928 photomultiplier tube detector (Hamamatsu Corporation, Bridgewater, NJ). The collected compound was diluted in HPLC-grade acetonitrile in a 1 cm quartz cuvette to an optical density at 292 nm of 0.2. Spectra were

determined between 300 and 555 nm at 0.5 nm increments with 292 nm excitation. One hundred spectra were collected and averaged to reduce noise and the spectrum of acetonitrile alone was collected to confirm that solvent did not contribute to the sample spectrum. Finally, the spectrum was corrected for the sensitivity profile of the detector using a manufacturer-supplied correction file.

Cellular characterization of mutant strains. To determine the intracellular localization of accumulated the scytonemin monomer, we used fluorescence confocal microscopy. UVA induced wild-type, $\Delta scyE$, and all *ebo* gene deletion strains were cultured and treated as described above. After five days of exposure to UVA, cells were collected, wet mounts prepared, and imaged on a Leica TCS SP5 AOBS Spectral Confocal System, using both bright field and fluorescence microscopy. Laser excitation was at 405 nm. Emission at 665 nm was used to visualize photosynthetic pigment fluorescence (chlorophyll *a* and phycobilin emission), and emission at 410 nm was used to visualize scytonemin monomer fluorescence. All images were taken at 400x magnification. Fluorescence quantification and image analyses were performed using ImageJ (26). Additional imaging was carried out using a Zeiss LSM800 laser scanning confocal microscope equipped with a Plan-Apochromat 63x 1.40 NA oil immersion objective with fluorescence excitation wavelengths as above. Further contrast adjustment for presentation purposes was done using Zen 2.3 (© Carl Zeiss Microscopy GmbH, 2011) image analysis software.

Cellular localization of ScyE. Bioinformatic analysis of *scyE* revealed that it contains an N-terminal Sec pathway signal peptide, suggesting that it is

periplasmically localized. To confirm this prediction, UVA induced and uninduced wild-type *N. punctiforme* cultures were harvested and subjected to lysis by osmotic shock as described by Ross *et. al.* 2007 (28) with the following revisions: 300 mM sucrose lysis buffer was used to preferentially lyse the outer-membrane, while increasingly efficient whole cell lysates were obtained using 400 and 500 mM lysis buffers. Proteomic analyses of lysate preparations were then conducted as previously described in Mitchell *et. al.* 2018 (31). In order to ascertain cellular localization of gene products of interest, we compared the ratio of abundance (normalized spectral abundance factor, NSAF) for each protein (ScyA, ScyC, ScyE, and ScyF) to the corresponding abundances (NSAF) of three known periplasm targeted peptides. (S-layer Domain Protein, Peptidase S8, and L-sorbose dehydrogenase; Uniprot accessions B2J6K1, B2JAL9, and B2IVQ5 respectively). The ratios were plotted against the molar concentration of lysis buffer and exponential regressions assessed by ANOVA.

Presence of the scytonemin monomer in periplasmic extracts. To confirm the differences in cellular periplasmic localization of the scytonemin monomer in $\Delta scyE$ vs. Δebo mutants we assessed the susceptibility of its extraction from cell lysates induced by mild osmotic shock. Induced $\Delta scyE$ and Δebo cells were shocked as described above (28) and the lysates (3 ml) extracted by mixing in equal volumes of ethyl acetate, followed by phase separation. The lipid-soluble extracts were analyzed by HPLC as described above and scytonemin monomer quantified fluorometrically, excited at 293 nm with emission

detected at 407 nm. Percentage yields of lysate fractions were compared to the total extracts determined from cell pellets in the corresponding preparations.

Results

To determine whether the product of any gene within the *ebo* cluster (here defined as ORFs from Npun_F5232 to Npun_F5236) was involved in the production of scytonemin, we deleted the entire cluster in *N. punctiforme* (Fig. 3). The mutant strain, Δebo , was then tested for its ability to produce scytonemin, against the wild-type (WT) and a previously obtained scytoneminless mutant, $\Delta scyE$, as controls. The Δebo strain presented a clear scytoneminless phenotype, like that of $\Delta scyE$, and only the WT produced scytonemin under inductive conditions (Fig. 3 Top). Subsequently, to determine which gene products in the *ebo* cluster were responsible for the Δebo scytoneminless phenotype, and to assess their specific roles, we constructed five in-frame deletion mutants: $\Delta eboA$, $\Delta eboB$, $\Delta eboC$, $\Delta eboE$, and $\Delta eboF$ (NpunF5232, NpunF5233, NpunF5234, NpunF5235, NpunF5236, respectively; Fig. 3), in which the rest of the operon was conserved in its proper reading frame (see supplemental Table S1 for details on construction). Recombinant plasmids were sequenced to ensure that no other mutations were created during construction. Chromosome segregation of the deletion mutants was confirmed by PCR using different combinations of primers (Table S1), as previously described (12). We found no polar transcriptional effects by RT-PCR targeting transcripts of the *ebo* gene downstream of each mutation. All mutant strains were tested for scytonemin production as described for Δebo . Each one of the five in-

frame deletion mutants was scytoneminless, indicating that all five of the *ebo* genes are essential for scytonemin production (Fig. 3 Bottom). In addition to lacking scytonemin, the *ebo* deletion mutants also showed enhanced susceptibility to UVA damage under inductive conditions; within the five-day induction period all *ebo* mutants exhibited chlorotic phenotypes, while the wild-type strain remained unaffected. This is not necessarily a result of lack of sunscreen, given that other scytoneminless mutants do not exhibit increased sensitivity to UVA (6).

No compounds unique to the *ebo* mutants were found to accumulate in the aqueous extract preparations in any of the mutants by HPLC analysis. By contrast, acetone extracts of all *ebo* mutants (single-gene and cluster mutant) contained a single compound that accumulated when induced by UVA exposure, but was not present in the wild-type cell extracts. This compound was not detected in any of the *ebo* mutants when non-induced cells were extracted. It exhibited a retention time of 7.8 minutes and a characteristic absorbance spectrum with a visible maximum at 407 nm (Fig. 4A). Upon isolation and collection, it was determined to have a molecular mass of 275 Da, by mass spectroscopy (274 Da fragment due to deprotonation in negative ion mode)(Supplementary fig S2). The retention time, absorbance spectra, and mass are consistent with those of the compound produced by the expression of *scyA-C* in *E. coli* (13), the structure of which has been fully resolved by nuclear magnetic resonance ((3Z)-3-[(4-hydroxyphenyl)methylidene]-1H,2H,3H,4H-cyclopenta[b]indol-2-one; scytonemin monomer in Fig. 1). In order to confirm the identity of the accumulated compound, we obtained an authentic standard for HPLC by constructing an *E. coli* strain containing

scyA-C, described in Materials and Methods, and isolating the main compound produced. HPLC co-elution of the standard and each of the compounds collected from *ebo* mutants produced single peaks in all cases, thus confirming identity. Having determined that the scytonemin monomer accumulated in all of the *ebo* mutants, we also re-analyzed mutants Δ *scyD*, Δ *scyE*, and Δ *scyF*. We could confirm that neither Δ *scyD*, nor Δ *scyF* (both with a scytonemin positive phenotype) accumulated the monomer, but we could clearly detect and identify it in Δ *scyE* under conditions of UV induction (but not in non-induced cells); this had not been detected in previous studies (12). All *ebo* mutants and Δ *scyE* were biochemically identical with respect to a lack of scytonemin production and the accumulation of the scytonemin monomer. The presence of the scytonemin monomer in Δ *scyE* cells indicates that absence of *ebo* genes is not an absolute requirement for the production of the scytonemin monomer.

Once the identity of the compound produced by *ebo* deletion mutants and Δ *scyE* was confirmed, we sought to determine if, as can be predicted by its molecular structure, it would emit fluorescence upon excitation, allowing for the investigation of its intracellular localization via fluorescence microscopy. Indeed, the scytonemin monomer exhibited a wide fluorescence emission in the blue with maxima around 407 nm, in accordance with its absorbance spectra (Fig. 4A inlay, excitation at 292 nm). We then tested if the levels of accumulation and the fluorescence yield would be sufficient for microscopy imaging. We used laser excitation at 405 ± 1 nm and collecting emission at 410 ± 2 nm, rather than at the absolute maximum of 407 nm to avoid excitation bleeding. Indeed, it was possible to visualize the accumulation of the scytonemin monomer in the

mutants, as shown exemplary with confocal microscopy images for the wild-type, Δebo , and $\Delta scyE$ mutants, as well as in a quantitative manner for all mutants (Fig 5). Wild-type *N. punctiforme* autofluorescence levels at 410 nm were low, both in induced and non-induced cells, but all mutants showed several-fold increases in fluorescence, as expected, upon induction and accumulation of the scytonemin monomer. In the absence of UVA induction, the fluorescence at 410 nm in all mutants was as basal as that of the wild-type.

All microscopic images shown in Fig. 5 were conducted under identical microscope settings and, to avoid variability, stem from a single concurrent induction experiment. Inductions and microscopic imaging were replicated independently three times for each mutant. Autofluorescence of photosynthetic pigments at 665 nm is also shown for comparison. Some signs of chlorosis (lower content of photosynthetic pigments compared to the wild-type) were present in *ebo* and $\Delta scyE$ mutants under exposure to UVA radiation. In these experiments, periods of induction longer than five days resulted in obvious cellular damage, generalized loss of autofluorescence both in the blue and in the red.

Because photosynthetic pigments are part of macromolecular complexes localized in the intracytoplasmatic thylakoid membranes (sometimes also on the inner leaflet of the cytoplasmic membrane), cell autofluorescence at 665nm, which has its origin largely from Chl *a* and phycobiliproteins, can be used to visualize the bounds of the cytoplasm. Having concurrent photopigment fluorescence to define the bounds of the cytoplasm and bright field images to establish the boundaries of the whole cell, we used comparative overlay images to determine the cellular localization of scytonemin monomer

accumulation in the various mutants. The images clearly show the presence of the scytonemin monomer only in the cytoplasm in each of the Δebo cells, but in both the cytoplasm and periplasm of $\Delta scyE$ cells (Fig. 6, Supplementary Fig. S3). Additionally, differential interference contrast / 410 nm fluorescence image overlays indicate that the scytonemin monomer is contained within the cell (Supplementary Fig. S4).

In order to verify the periplasmic localization of ScyE suggested by its N-terminal Sec signal peptide, induced cultures were subjected to periplasmic fractionation. ScyE could be detected in periplasmic lysate, as could other periplasmic proteins (S-layer Domain Protein, Peptidase S8, and L-sorbose dehydrogenase). The S-layer domain containing proteins have been identified in the peptidoglycan layer of gram negative bacteria (29), necessitating their translocation to the periplasmic space after synthesis. L-sorbose dehydrogenase plays a role in L-sorbose assimilation, a periplasmic process (30). Ratios of abundances for cytoplasmic proteins compared to the three standard periplasmic proteins should increase exponentially with increasing strength of osmotic shock due to dilution of the periplasmic contents by increased release of cytoplasm. This was demonstrably the case for ScyA and ScyC (Fig. 7). However, the abundance ratio of ScyE and ScyF to each of these periplasmic proteins either remained statistically invariant or decreased slightly with increasing strength of osmotic shock, which indicates that they partition to the periplasm with equal to or greater consistency than the standard proteins (Fig. 7) and confirms the periplasmic localization of ScyE and ScyF.

In order to confirm the periplasmic localization of the scytonemin monomer with an alternative approach to microscopy, we assessed its differential release in periplasmic

fractions of induced $\Delta scyE$ vs. those in Δebo cells. We found that $\Delta scyE$ yielded on average 21 % of the total cellular content of scytonemin monomer to periplasmic preparations obtained with mild osmotic shock (0.2-0.3 M sucrose shock) whereas Δebo yielded only about 6%, consistent with a periplasmic localization of the scytonemin monomer in $\Delta scyE$ mutants and a cytoplasmic restriction in Δebo cells.

Discussion

The fact that every *ebo* deletion mutant resulted in a scytoneminless phenotype strongly suggested that the *ebo* genes are involved in the process of scytonemin production in *N. punctiforme*, despite their widespread occurrence in non-scytonemin producing bacteria. This conclusion is consistent with their conserved placement within the scytonemin operon in most cyanobacteria (Fig. 3), and their up-regulation in *N. punctiforme* cells exposed to UVA radiation (7,14). However, gene-specific phenol-indolic metabolites did not accumulate in *ebo* mutants, at least not to detectable levels, which does not fit the narrative that the *ebo* genes simply code for enzymes carrying out a sequential set of reactions involved in the late stages of scytonemin formation, as had been hypothesized (12). Rather, a single intermediary, the scytonemin monomer, accumulates in each of the *ebo* mutants under UVA induction (Fig. 1). This compound is identical to the combined products of ScyA-C when expressed heterologously in *E. coli* (13) or, just one oxidation step away from the product of ScyC *in vitro* (11), suggesting that the action of the Ebo proteins must be other than a fundamental biochemical transformation of the ScyC product. That all *ebo* knockout mutants resulted in the accumulation of the very same intermediary suggests that a coordinated action of all of

gene products is required. Thus, in entertaining a potential role for the *ebo* cluster that would satisfy the available biochemical evidence, one must envision an ancillary process, such as the synthesis of a cofactor necessary in the scytonemin synthesis pathway. Such a generic role is attractive in that the Ebo proteins could potentially serve a similar role in the various biochemical pathways of the bacterial and algal species in which they are found. But, as noted, the bioinformatics evaluation of the *ebo* gene products did not immediately suggest viable hypotheses as to what that role may be (15).

Because several of the scytonemin operon gene products had been predicted to be targeted to the periplasm (ScyDEF) since they possess canonical N-terminal signal peptides utilized by the Sec transport system (12), the late portion of the scytonemin biosynthesis was predicted to take place there, *en route* to eventual secretion (7,1). Determining the cellular localization of the scytonemin monomer could provide support for these predictions. Fortunately, accumulation of the scytonemin monomer in *ebo* deletions and Δ *scyE* could be visualized by fluorescence microscopy. These investigations (Fig. 6 and Supplementary Fig. S4) clearly showed that in the absence of any one (or all) of the *ebo* genes studied, the scytonemin monomer remained restricted to the cytoplasm. By contrast, in Δ *scyE*, where all *ebo* genes were intact, the monomer was also found in the periplasm, where it reached concentrations similar to those in the cytoplasm (Supplementary Fig. S5). These results were confirmed by comparison of extraction yields of Δ *scyE* and Δ *ebo* periplasmic fractions. From this we deduce that the Ebo proteins together play a role in the translocation of the scytonemin monomer across the inner membrane, wherein the absence of any one of them prevents this translocation,

leading to cytoplasmic accumulation. These findings also lead us to posit that ScyE is responsible for the oxidative dimerization of the scytonemin monomer, to form scytonemin, the process taking place in the periplasm, to where ScyE is demonstrably targeted. In $\Delta scyE$, the scytonemin monomer is excreted to the periplasm, only to find no enzyme target, and accumulates there, eventually causing feedback accumulation in the cytoplasm as well. In other mutants, such as those bearing deletions of ScyD or ScyF, proteins that are not central to biosynthesis, the scytonemin monomer is produced, exported and further processed, thus leading to scytonemin positive phenotypes. Together, our observations lead to a new model of scytonemin synthesis in which the *ebo* gene products act together for the translocation of the scytonemin monomer from the cytoplasm to the periplasm for final oxidative dimerization.

A potential mechanism of action. To entertain potential mechanisms for this translocation, we briefly review structural and phylogenetic traits of the proteins involved. *eboC* (Npun_F5234) presents strong homologies to the UbiA superfamily of prenyltransferases (15), transmembrane proteins that catalyze key prenylation steps in the production of various metabolites (21). EboC does indeed contain seven transmembrane domains, being thus integral to the cytoplasmic membrane, and holds closest homology to both the digeranylgeranylgeranylphosphate (DGGGP) synthases involved in the synthesis of archaeal membrane lipids and to archaeal UbiA homologues involved in the synthesis of ubiquinone. Its predictive active site (Supplementary Fig. S6) contains both the cluster of 4 asparagine residues that interact with the prenyl-group

donor, as well as the residues (Arg142 and Asp145) responsible for prenyl acceptance (25). We did not analyze *eboD* because it is not present in the *Nostoc punctiforme ebo* cluster proper, but a homologue is found nearby in the opposite reading direction, and it is generally integrated within the *ebo* operon in other bacteria. EboD presents clear homology with sugar phosphate cyclases and has been posited to act on sedoheptulose-7-P (15). EboA has no close sequence homologs, but a structural prediction reveals several repeating alpha helices (Supplementary Fig. S6), consistent with the Tetra-Trico-Peptide Repeat (TPR) domain responsible for protein-protein complex stabilization (24, 27). Of the *ebo* gene products, only EboC and EboB contain parallel alpha-helices reminiscent of a TPR domain, thus EboA may facilitate formation of an EboCAB complex anchored to the cell membrane; the *eboA* knockout mutant fails to organize the correct assembly. Assigning a potential role to the other *ebo* genes becomes much more speculative. *eboB* is annotated as a putative hydrolase, an enzyme class with such diversity as to preclude further predictions. *eboE* (Npun_F5235) codes for a TIM-barrel containing enzyme annotated as a putative xylose isomerase, and *eboF* (Npun_F5236) shares homology with pyrophosphatases. These considerations suggest that the enzymatic reactions carried out by the *ebo* cluster include the modification of sugars, likely cyclitol formation from heptose precursors, and the prenylation of an undetermined substrate. Prenylated molecules, so-called lipid carriers, are indeed known as integral components of two other periplasmic metabolite translocation processes in bacteria: the synthesis of cell wall peptidoglycan (34), and the excretion of capsular polysaccharides (CPS) / lipopolysaccharides (LPS) (35). In the latter case, the carrier is a glycolipid that is translocated along with the polysaccharide chain through a transmembrane ATP-binding-

cassette (ABC) transporter protein complex (35). The carrier is composed of a lyso-phosphatidylglycerol moiety and the linker sugar 3-deoxy-D-manno-oct-2-ulosonic acid (or keto-deoxyoctulosonate, KDO)(35). Thus, one can see some parallels between a potential *ebo*-cluster product composed of a lipid and a sugar cyclitol, and a lipid carrier similar to that described for CPS systems, as this would enable the translocation of the scytonemin monomer using the existing CPS ABC permease. Indeed, scytonemin synthesis and CPS excretion are related processes in *Nostoc*: UVA elicits the production of both (36), and *Nostoc* mutants deficient in scytonemin production display enhanced capsular polysaccharide production under UVA induction relative to the wild type (6). In fact, the *Nostoc* gene with the closest homology to well-described CPS ABC permease (Npun_R5235) is among those upregulated by UVA exposure in global transcriptomic studies (37). Considering that no lipid-conjugated scytonemin monomer intermediate was found in the HPLC-MS analyses of this study it is not likely that covalent bonding is responsible for association of the putative *ebo* lipid carrier and the scytonemin monomer. Perhaps non-covalent interactions such as polar- π bonding (38) between hydroxyls of the sugar headgroup of the carrier and the aromatic rings of the monomer facilitate this attachment. In any event, investigating the accuracy of this proposed mechanistic model will constitute a long term, challenging task that may benefit from a molecular characterization of the *ebo* proteins and their products as they pertain to glycolipid biosynthesis rather than indole-phenol intermediaries, as done here. We find this model of *ebo* function an attractive hypothesis for future testing, given its integration with a well-defined and ubiquitous membrane translocation system as well as its consistency with CPS formation.

Roles of ebo genes beyond cyanobacteria. As a recognizable genomic element in the genomes of over 150 microbes that occupy diverse ecological niches, the impact of the *ebo* cluster clearly exceeds the biology of scytonemin and cyanobacteria. The present findings represent the first description of *ebo* gene function in any organism, potentially shedding light onto the reason for their high evolutionary conservation across disparate phyla of bacteria and algae. In the literature to date there has been only one tangential study of the *ebo* gene cluster function. Burlinson *et. al.* 2013 (32) produced transposon insertion mutants in a cluster of *Pseudomonas fluorescens* NZ17 that rendered the bacterium vulnerable to predation by the nematode *Caenorhabditis elegans*; they hence named this genetic region the EDB cluster (for ‘edible’) and conjectured that it codes for the synthesis of a nematode-repellent compound. The EBD cluster in fact contains an full *ebo* cluster, in which transposon interrupted *eboA*, *eboD*, and *eboE* homologues result in the edible phenotype. Interestingly, some of the other non-*ebo* EBD open reading frames contain signal peptides for excretion to the periplasm, and genes homologous to the CPS ABC permease systems are found just upstream of the EDB cluster. With these coincidences and the hindsight of our study, it is fair to postulate that the *ebo* homologues within the EDB cluster may very well facilitate the excretion of such hypothetical nematode repellent in *Pseudomonas*. A generic role in metabolite excretion for the *ebo* genes may also offer explanations for their relative incidence in symbiotic interactions between bacteria and microalgae (15, 39).

References

1. Gao Q, Garcia-Pichel F. 2011. Microbial ultraviolet sunscreens. *Nat Rev Microbiol* 9:791–802.

2. Garcia-Pichel F, Castenholz RW. 1991. Characterization and biological implications of scytonemin, a cyanobacterial sheath pigment. *J Phycol* 27:395-409.
3. Garcia-Pichel F, Sherry ND, Castenholz RW. 1992. Evidence for an ultraviolet sunscreen role of the extracellular pigment scytonemin in the terrestrial cyanobacterium *Chlorogloeopsis* sp. *Photochem Photobiol* 56:17-23.
4. Rastogi RP, Sinha RP, Incharoensakdi A. 2013. Partial characterization, UV-induction and photoprotective function of sunscreen pigment, scytonemin from *Rivularia* sp. HKAR-4. *Chemosphere* 93:1874–1878.
5. Proteau PJ, Gerwick WH, Garcia-Pichel F, Castenholz R. 1993. The structure of scytonemin, an ultraviolet sunscreen pigment from the sheaths of cyanobacteria. *Experientia* 49:825-829.
6. Soule T, Stout V, Swingley WD, Meeks JC, Garcia-Pichel F. 2007. Molecular genetics and genomic analysis of scytonemin biosynthesis in *Nostoc punctiforme* ATCC 29133. *J Bacteriol* 189:4465-4472.
7. Soule T, Palmer K, Gao Q, Potrafka R, Stout V, Garcia-Pichel F. 2009. A comparative genomics approach to understanding the biosynthesis of the sunscreen scytonemin in cyanobacteria. *BMC Genomics* 10:336.
8. Sorrels CM, Proteau PJ, Gerwick WH. 2009. Organization, evolution, and expression analysis of the biosynthetic gene cluster for scytonemin, a cyanobacterial UV-absorbing pigment. *Appl Environ Microbiol* 75:4861–4869.
9. Naurin S, Bennett J, Videau P, Philmus B, Soule T. 2016. The response regulator Npun_F1278 is essential for scytonemin biosynthesis in the cyanobacterium *Nostoc punctiforme* ATCC 29133. *J Phycol* 52:564–571.
10. Balskus EP, Walsh CT. 2008. Investigating the initial steps in the biosynthesis of cyanobacterial sunscreen scytonemin. *J Am Chem Soc* 130:15260-15261.
11. Balskus EP, Walsh CT. 2009. An enzymatic cyclopentyl[b]indole formation involved in scytonemin biosynthesis. *J Am Chem Soc* 131:14648-14649.
12. Ferreira D, Garcia-Pichel F. 2016. Mutational studies of putative biosynthetic genes for the cyanobacterial sunscreen scytonemin in *Nostoc punctiforme* ATCC 29133. *Front Microbiol* 7:1–10.

13. Malla S, Sommer MOA. 2014. A sustainable route to produce the scytonemin precursor using *Escherichia coli*. *Green Chem* 16:3255–3265.
14. Soule T, Garcia-Pichel F, Stout V. 2009. Gene expression patterns associated with the biosynthesis of the sunscreen scytonemin in *Nostoc punctiforme* ATCC 29133 in response to UVA radiation. *J Bacteriol* 191:4639–4646.
15. Yurchenko T, Ševčíková T, Strnad H, Butenko A, Eliáš M. 2016. The plastid genome of some eustigmatophyte algae harbours a bacteria-derived six-gene cluster for biosynthesis of a novel secondary metabolite. *Open Biol* 6. doi:10.1098/rsob.160249
16. Campbell EL, Summers ML, Christman H, Martin ME, Meeks JC. 2007. Global gene expression patterns of *Nostoc punctiforme* in steady-state dinitrogen-grown heterocyst-containing cultures and at single time points during the differentiation of akinetes and hormogonia. *J. Bacteriol* 189:5247–5256.
17. Allen MB, Arnon DI. 1955. Studies on nitrogen-fixing blue-green algae. I. Growth and nitrogen fixation by *Anabaena cylindrica*. *Lemm. Plant Physiol* 30:366–372.
18. Cai YP, Wolk CP. 1990. Use of a conditionally lethal gene in *Anabaena* sp. strain PCC 7120 to select for double recombinants and to entrap insertion sequences. *J Bacteriol* 172:3138–3145.
19. Cohen MF, Wallis JG, Campbell EL, Meeks JC. 1994. Transposon mutagenesis of *Nostoc* sp. strain ATCC 29133, a filamentous cyanobacterium with multiple cellular differentiation alternatives. *Microbiology* 140:3233–3240.
20. Karsten U, Garcia-Pichel F. 1996. Carotenoids and Mycosporine-like Amino Acid Compounds in Members of the Genus *Microcoleus* (Cyanobacteria): A Chemosystematic Study. *Syst Appl Microbiol* 19:285–294.
21. Li W. 2016. Bringing Bioactive Compounds into Membranes: The UbiA Superfamily of Intramembrane Aromatic Prenyltransferases. *Trends Biochem Sci* 41:356–370.
22. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. 2001. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J Mol Biol* 305:567-580.
23. J Yang, R Yan, A Roy, D Xu, J Poisson, Y Zhang. 2015. The I-TASSER Suite: Protein structure and function prediction. *Nature Methods* 12: 7-8.

24. Watkins AM, Wuo MG, Arora PS. 2015. Protein-Protein Interactions Mediated by Helical Tertiary Structure Motifs. *J Am Chem Soc* 137:11622–11630.
25. Cheng W, Li W. 2014. Structural Insights into Ubiquinone Biosynthesis in Membranes 343:878–882.
26. Schneider CA, Rasband WS, Eliceiri KW. 2012. HISTORICAL commentary NIH Image to ImageJ : 25 years of image analysis. *Nat Methods* 9:671–675.
27. Zeytuni N, Zarivach R. 2012. Review Structural and Functional Discussion of the Tetra-Trico-Peptide Repeat, a Protein Interaction Module. *Struct Des* 20:397–405.
28. Ross DE, Ruebush SS, Brantley SL, Hartshorne RS, Clarke TA, Richardson DJ, Tien M. 2007. Characterization of protein-protein interactions involved in iron reduction by *Shewanella oneidensis* MR-1. *Appl Environ Microbiol* 73:5797–5808.
29. Breitwieser A, Gruber K, Sleytr UB. 1992. Evidence for an S-layer protein pool in the peptidoglycan of *Bacillus stearothermophilus*. *J Bacteriol* 174:8008–15.
30. Shinjoh M, Tazoe M, Hoshino T. 2002. NADPH-dependent L-sorbose reductase is responsible for L-sorbose assimilation in *Gluconobacter suboxydans* IFO 3291. *J Bacteriol* 184:861–863.
31. Mitchell NM, Sherrard AL, Dasari S, Magee DM, Grys TE, Lake DF. 2018. Proteogenomic Re-Annotation of *Coccidioides posadasii* Strain Silveira. *Proteomics* 18: 1700173-1700173
32. Burlinson P, Studholme D, Cambray-Young J, Heavens D, Rathjen J, Hodgkin J, Preston GM. 2013. *Pseudomonas fluorescens* NZI7 repels grazing by *C. elegans*, a natural predator. *ISME J* 7:1126–1138.
33. Bertani G. 1951. Sensitivities of Different Bacteriophage Species to Ionizing Radiation. *J. Bacteriol*
34. Bouhss A, Trunkfield AE, Bugg TDH, Mengin-Lecreulx D. 2008. The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol Rev* 32:208–233.
35. Willis LM, Whitfield C. 2013. Structure, biosynthesis, and function of bacterial capsular polysaccharides synthesized by ABC transporter-dependent pathways. *Carbohydr Res* 378:35–44.

36. Ehling-Schulz M, Bilger W, Scherer S. 1997. UV-B-induced synthesis of photoprotective pigments and extracellular polysaccharides in the terrestrial cyanobacterium *Nostoc commune*. *J Bacteriol* 179:1940–1945.
37. Soule T, Gao Q, Stout V, Garcia-pichel F. 2013. The Global Response of *Nostoc punctiforme* ATCC 29133 to UVA Stress , Assessed in a Temporal DNA Microarray Study. *Photochem Photobiol* 89: 415-423.
38. Du Q-S, Wang Q-Y, Du L-Q, Chen D, Huang R-B. 2013. Theoretical study on the polar hydrogen- π (Hp- π) interactions between protein side chains. *Chem Cent J* 7:92.
39. Yurchenko T, Tereza Š, Pavel P, Karkouri E, Klime V, Zbránková V, Kim E, Raoult D, Santos LMA, Eliá M. 2018. A gene transfer event suggests a long-term partnership between eustigmatophyte algae and a novel lineage of endosymbiotic bacteria. *ISME J* <https://doi.org/10.1038/s41396-018-0177-y>

Figures

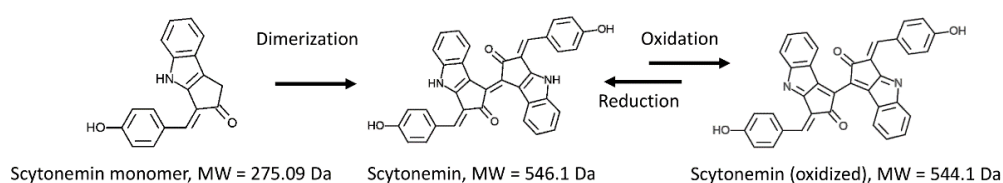


Fig. 1: Structures of the scytonemin monomer, and scytonemin. The likely final step in scytonemin biosynthesis involves oxidative dimerization of the scytonemin monomer.

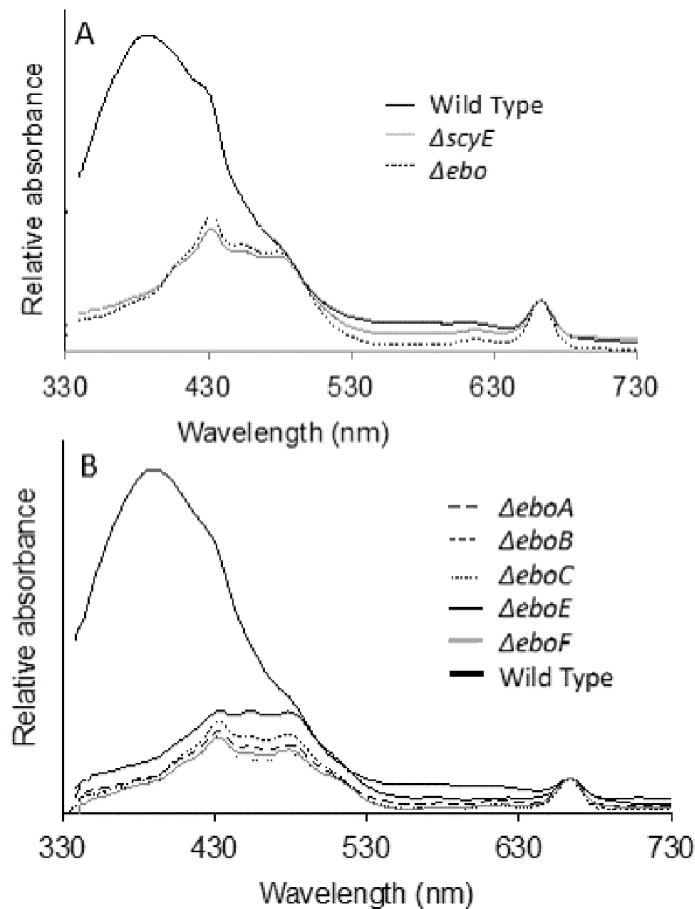


Fig. 2: Absorbance spectra of wild-type and *scy* and *ebo* mutant strains. Top: Absorbance spectra of acetone cell extracts from wild-type (solid black), *scyE* mutant (solid grey), and *ebo* mutant (dotted black), after UVA induction of the scytonemin operon. The wild-type produced scytonemin, as indicated by a large absorbance maximum at 384 nm. Bottom: Absorbance spectra of acetone cell extracts of individual *ebo* gene deletion mutants after UVA induction of the scytonemin operon, all displaying a scytoneminless phenotype.

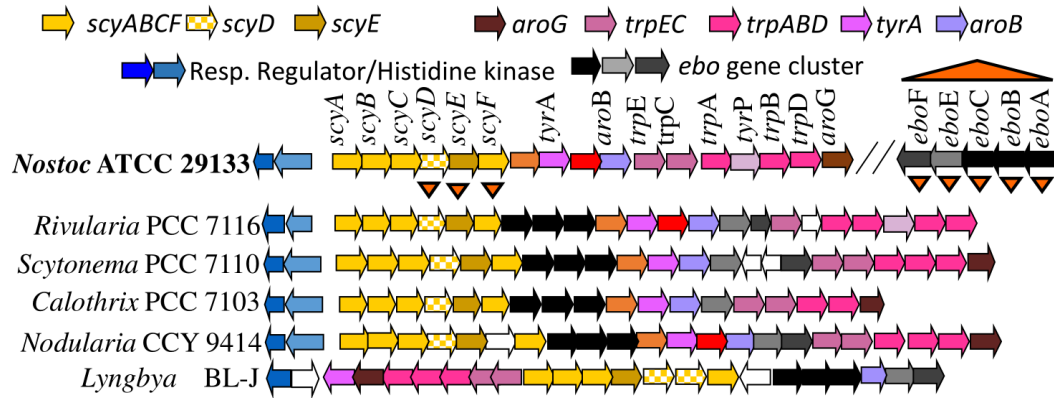


Fig. 3. Genomic organization of the scytonemin operon in *N. punctiforme* and other cyanobacteria, including the *ebo* genes of unknown function found within the *scy* operon of most cyanobacteria (but distally in *N. punctiforme*). Triangles indicate the genes whose deletion mutants were examined in this study.

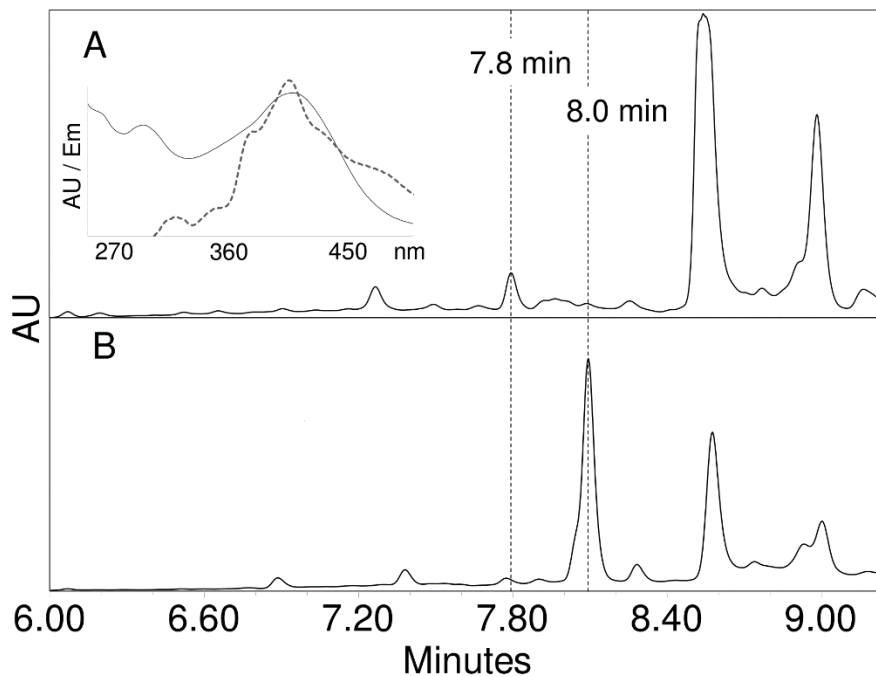


Fig. 4: Separation and characterization of a compound accumulated after UVA induction by $\Delta eboC$ strain. (A) HPLC chromatogram of acetone extract showing production of a novel compound eluting at 7.8 minutes and the absence of scytonemin peak at 8 minutes. This pattern was found in all *ebo* mutants and in $\Delta scyE$ (see Supplementary Fig. S1), and none produced the compound without an induction of the scytonemin operon (see Supplementary Fig. S1). The inlay shows the UV-Vis absorbance spectrum (solid line) and the fluorescence emission spectrum (dotted line) of the newly accumulated compound after collection from HPLC eluent. (B) Chromatogram of wild-type extract

after UVA induction, indicating presence of scytonemin at 8 minutes and the absence of the 7.8 m peak. $\lambda = 407\text{nm}$

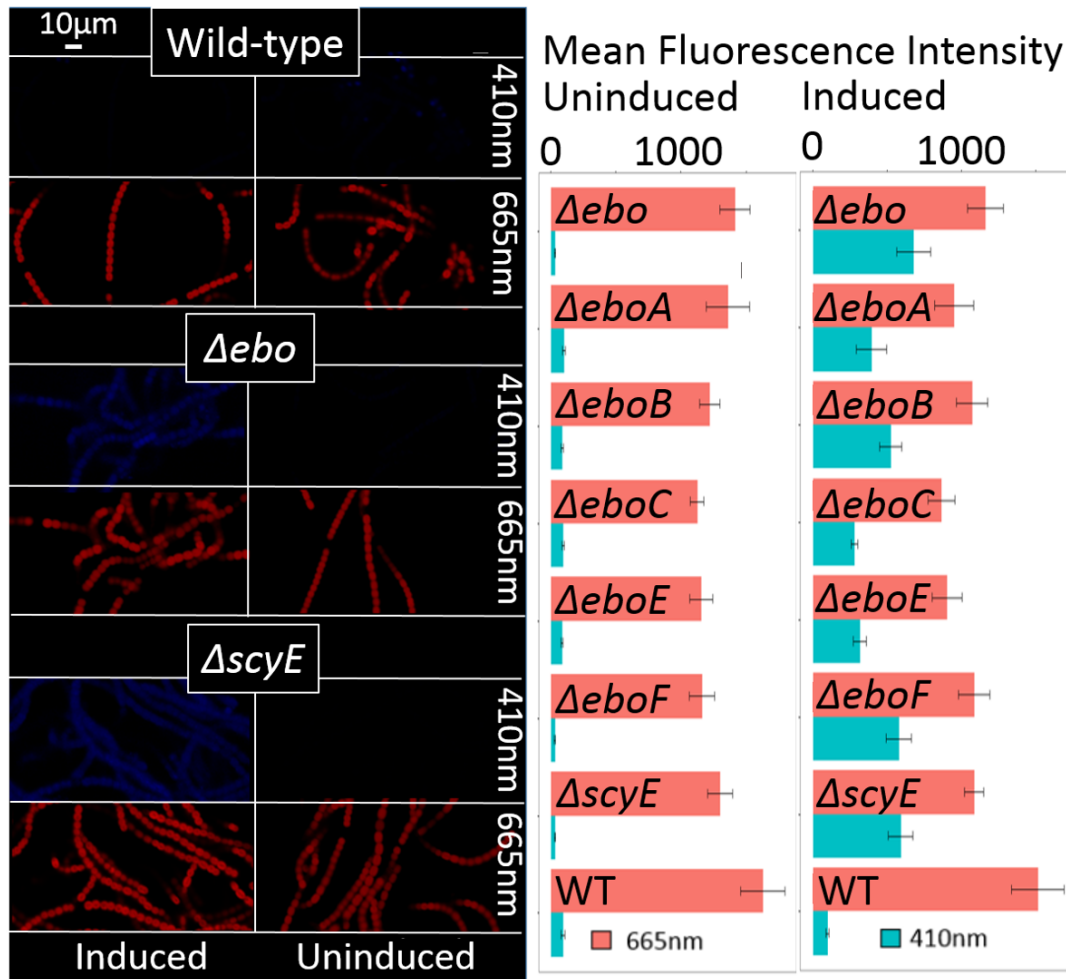


Fig. 5: Confocal fluorescence imaging and quantification of the scytonemin monomer accumulation *in-vivo*. Left: Fluorescence images of wild-type (top), Δebo (middle), and $\Delta scyE$ (bottom) with emission at 410 nm (to visualize the scytonemin monomer) and 665 nm (to visualize photopigments), under conditions of induction and without induction by UVA. Right: Fluorescence intensity quantification within cells of wild-type and mutants at 410 and 665 nm under inductive and non-inductive conditions, respectively (n = 10; bars indicate standard error of the mean).

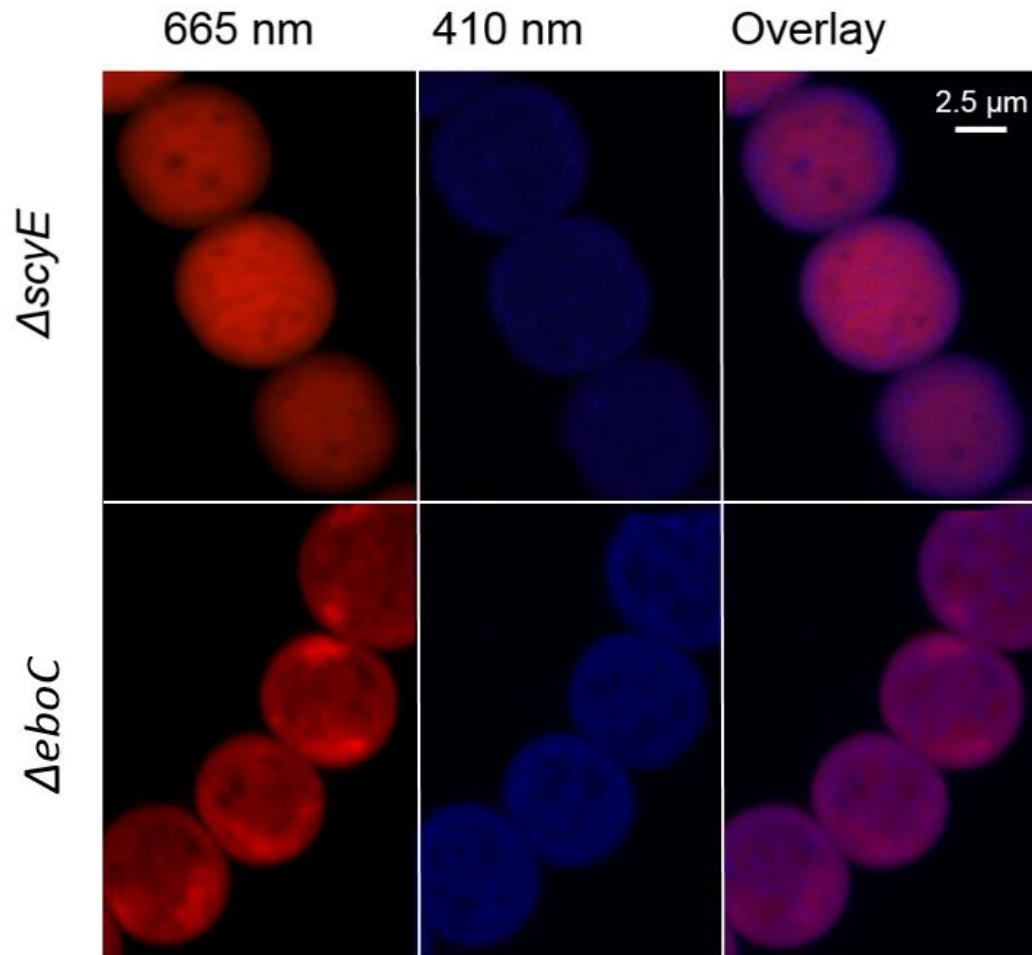


Fig. 6: Intracellular localization of the scytonemin monomer in induced $\Delta scyE$ and $\Delta eboC$ cells. Overlay of the 665 nm over the 410 nm images demonstrate the localization of the scytonemin monomer the cytoplasm of $\Delta eboC$ cells (this is the case for each and all of the *ebo* mutants, as can be seen in Supplementary Fig. 2). In $\Delta scyE$ cells, however, the scytonemin monomer accumulates both in the cytoplasm and the periplasm.

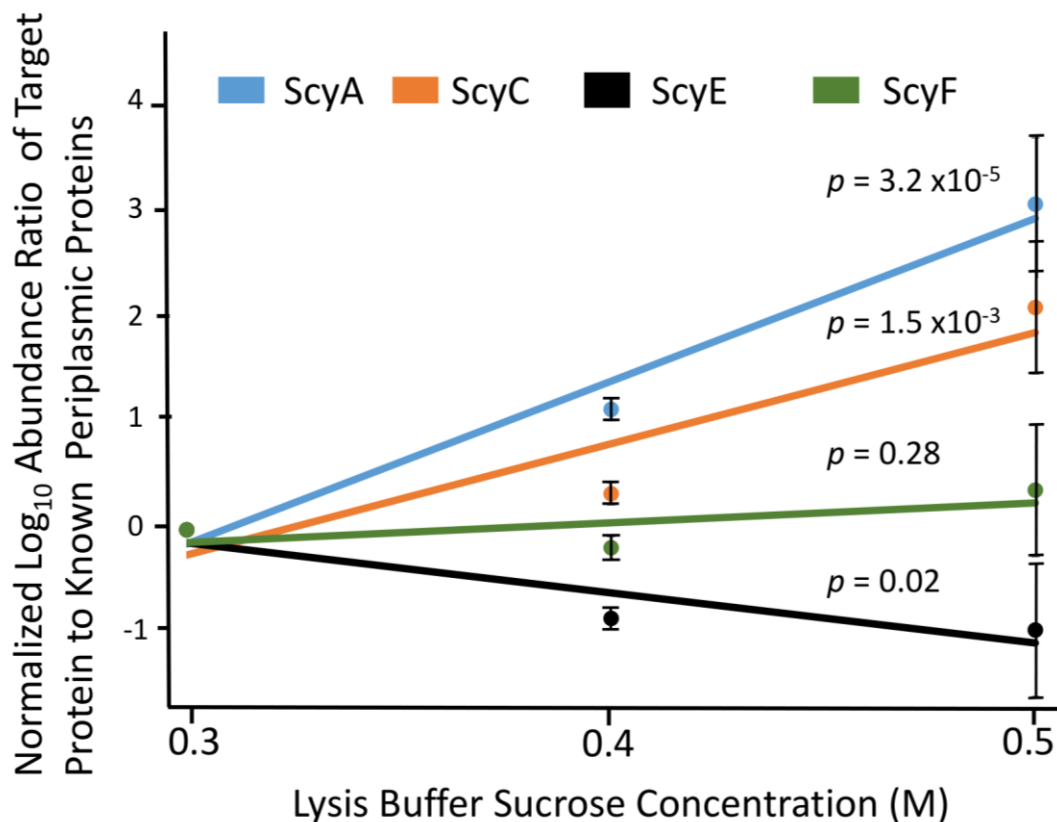


Fig. 7: Partitioning of core biosynthetic proteins of the scytonemin operon between cytoplasm and periplasm by proteomic analyses of osmotic shock lysates. Target protein relative abundance ratios to three different known periplasm-targeted proteins are plotted against the strength of lysis buffer used. Proteins localized to the cytoplasm should show an increase in ratio with buffer strength, whereas ratios of proteins partitioning preferentially to the periplasm should remain invariant or decrease. All ratios were normalized to 1 at 0.3 M sucrose for ease of graphing, and the probabilities that the slope of exponential regressions are significantly different from zero are including for each target protein dataset.

3 - THE FUNCTIONAL LANDSCAPE AND NEIGHBOR DETERMINING GENOMIC
REGION SEARCH (FLANDERS) BIOINFORMATIC DATA MINING
PACKAGE/PIPELINE AND ITS USE TO PROBE FUNCTIONS OF THE *EBO* GENE
CLUSTER AMONG BACTERIA

Kevin Klicki and Ferran Garcia-Pichel

Abstract

The coupling of transcription and translation through polycistronic mRNAs in bacteria sets the stage for coregulation of genetic elements by colocalization. By examining the genes upstream or downstream of a bacterial operon, one may garner insights about the biochemical function of the genomic region at large. The Functional Landscape And Neighbor Determining gEnomic Region Search (FLANDERS) pipeline was developed to extract the neighboring regions of a given query gene cluster and perform sequence and annotation based analyses on these regions to surmise a functional context. FLANDERS uses multigene BLAST, antiSMASH, and SignalP as dependencies. The neighboring regions of the widely conserved *ebo* gene cluster were investigated as a proof-of-concept case in 92 *ebo* containing genomes. No known biosynthetic gene clusters were detected in any *ebo* adjacent gene, but many were highly conserved putative biosynthetic clusters. Most *ebo* adjacent gene clusters had at least 2-3 signal peptide containing genes, consistent with the hypothesis that the *ebo* genes are involved in precursor transport during secondary metabolite biosynthesis, and are a modular genetic element that can plug and play into many genomic contexts. The source code and documentation for the FLANDERS pipeline can be found at <https://github.com/kklicki/FLANDERS>.

Introduction

Prokaryotes are distinguished from eukaryotes by the absence of a membrane bound nucleus containing genetic material, and thus they are able to directly couple the

processes of transcription and translation (1). The linkage of the processes makes advantageous the physical colocalization of functionally interdependent genes into operons, thus allowing for their coregulation via the synthesis of polycistronic mRNA transcripts (2). The arrangement of functionally interdependent genes into operons sets the stage for passive and energetically conservative regulation of function. For example, it is more parsimonious for a single transcription factor to control the expression of a multi-gene operon than requiring separate transcription factors to control the expression of a number of functionally interdependent genes scattered throughout the genome (3). For these reasons, operons are ubiquitous across bacteria, and in the age of whole genome sequencing, can provide insights into the potential function of a gene cluster purely by the colocalization of its constituent genes.

In an evolutionary capacity, the arrangement of functionally interdependent genes into coherent operons may lead to the formation of modular genetic elements; gene clusters whose action may facilitate or enhance the function of nearby genes in a variety of genomic contexts. These 'plug and play' modules could in turn be acquired horizontally and adopt a novel role in a new genomic context. This concept has been extensively studied in polyketide synthesis clusters, sets of modular gene products responsible for the condensation of coenzyme A activated ketones to produce polyketides with antibiotic properties (4–7). The diversity of polyketides across bacteria and fungi arises from the modular nature of their synthesis wherein a number of different polyketide synthesis enzymes can be integrated into a given pathway to produce a variety of active products (8). Another example of modular gene clusters comes from the non-ribosomal peptide

synthesis wherein multimodular non-ribosomal peptide synthases that polymerize amino acid substrates to create intricate oligopeptides like vancomycin (9). The modular design of these systems is of specific interest for biotechnology; modeling synthetic gene networks with interchangeable constituents on these naturally occurring analogs could hold the key to engineering robust and diverse pathways in industrially scalable organisms (10, 11). For example, In seeking to express the type 3 secretion system (T3S) from *Xanthomonas euvesicatoria* into non-pathogenic strains to deliver cargo proteins to eukaryotic targets, Hausner et al (2019) found that by arranging the genes encoding the T3S into a modular cloning construct, they were able to express the T3S in heterologous systems (12). These results suggest that the T3S may be a modular membrane transport system whose conservation can be accounted for by the variety of protein substrates it is capable of translocating. The ‘legolization’ of modular genetic elements holds great potential for many biotechnological applications (13). Interestingly, outside of polyketide synthesis and biotechnological applications, very few other modular genetic systems have been described.

The eustigmatophyte-bacterial operon (*ebo* gene cluster) is a putative modular genetic element that has been identified in a number of bacteria as well as the plastid genomes of two eustigmatophyte algae (14). It is composed of 6 genes (*eboA-F*), annotated as follows: 1) a hypothetical protein, 2) a TatD hydrolase, 3) UbiA prenyltransferase, 4) 3-dehydroquinate synthase, 5) a TIM-barrel containing xylose isomerase, and 6) a Type I phosphodiesterase respectively (14). While the precise mechanism of *ebo* function has yet to be elucidated, two studies have demonstrated its indispensability in cellular

adaptations to environmental stressors: Burlinson et al 2013 found the *ebo* cluster to be contained within a putative biosynthetic cluster coined *edb*, for edible, because knock-out mutants of any gene in this cluster resulted in a loss of resistance to predation by the soil nematode *C. elegans* (15). While the remaining genes of the *edb* cluster showed little homology to any known product, 4 of them contain a sec signal peptide, tagging them for transport from the cytoplasm to the periplasm (16). The authors hypothesized that the *edb* cluster may be responsible for the synthesis of a novel secondary metabolite which acts as a repellent of nematodal predators, but this putative compound has yet to be identified (15) (and see Appedix A of this dissertation). The other study into the role of the *ebo* cluster took place in the filamentous cyanobacteria *Nostoc punctiforme*, wherein it was shown that the *ebo* cluster is crucial for the synthesis of the homodimeric UV sunscreen compound scytonemin (17). Klicki et al 2018 showed that the *ebo* cluster is involved in the translocation of the scytonemin monomer from the cytoplasm to the periplasm (17). As with the *P. fluorescens* example, the interruption of any one or all of the *ebo* genes resulted in cessation of scytonemin synthesis, and relegation of the scytonemin monomer to the cytoplasm. Furthermore, as with the *edb* cluster in *P. fluorescens*, the *ebo* cluster is generally co-localized with the scytonemin biosynthesis cluster in the genomes of scytonemin producing cyanobacteria, although in *Nostoc* itself it is found away from the scytonemin operon. Another similarity to the *edb* cluster is that three genes in the scytonemin operon (*scyDEF*) contain sec signal peptides, consistent with a cytoplasmic/periplasmic compartmentalization of scytonemin synthesis. A map of the *ebo* genes in relation to their adjacent biosynthetic clusters in select genomes is shown in Fig. 1.

This consistent colocalization of the *ebo* cluster with biosynthetic clusters between these two model systems suggests that the *ebo* cluster may be a modular genetic element that facilitates the biosynthesis of a variety of bacterial secondary metabolites whenever they require transport across the cell membrane. The transporter-like function of the *ebo* genes during scytonemin synthesis may thus be an example of a more general role of the *ebo* cluster, as many secondary metabolites function extracellularly. What is more, the genomic colocalization of the *ebo* genes with the scytonemin and *edb* clusters suggests a potential means for discovering novel secondary metabolite biosynthesis clusters, if indeed *ebo* clusters tend to be flanked by the biosynthetic cluster with which they interact. An analysis of regions upstream and downstream of the *ebo* cluster may provide insights about both the *ebo* cluster itself as well as point to biosynthetic clusters, of known or unknown nature with which it may cooccur. To date there has been no bioinformatic data-mining tool to extract and analyze the flanking regions of a query genomic locus in a high throughput manner. In this work we report on the development of such a tool, the Functional Landscape And gEnomic Region Search (FLANDERS) and on its application to explore these questions as they pertain to the *ebo* cluster.

FLANDERS takes a list of homologous gene clusters as input, retrieves their source genomes from the NCBI database using the ncbi-acc-download module (18), extracts the upstream and downstream flanking regions, and performs sequence and homology analyses on these flanking ORFs. For the purpose of identifying known secondary metabolite biosynthetic gene clusters (BGCs) in *ebo* adjacent regions, the Antibiotic and Secondary Metabolite Analysis Shell (antiSMASH) (19) was employed. To test the hypothesis that *ebo* is involved in transport of different pathway intermediaries during

compartmentalized biosynthesis, *ebo* flanking ORFs were analyzed for the presence of signal peptide domains using SignalP (20). Finally, to assess the overall conservation of *ebo* adjacent gene clusters, *ebo* flanking regions were submitted to Multigene BLAST (21) in order to find conserved clusters up and downstream of *ebo* genes. Utilizing FLANDERS, we were able to determine that the *ebo* gene cluster is generally flanked by putative biosynthetic clusters containing signal peptide containing genes, and sometimes is part of a genomic island of transport related genes. These results demonstrated the utility of the FLANDERS pipeline towards discovering modular genetic systems in other contexts.

Materials and Methods

Identification of query cluster homologs

Prior to the FLANDERS pipeline, a query gene cluster in protein FASTA or a list of query gene clusters in protein FASTA format, are submitted to a multigene BLAST search. In either case, the multigene BLAST cluster-blast output file or list of files can then be submitted directly to the FLANDERS pipeline. From the list of cluster homologs, a user defined cutoff can be set for what FLANDERS will consider to be a complete cluster homolog. For instance, if the query cluster contains five genes, a full cluster homolog can be defined as a cluster-blast hit which contains homologs to all five query genes, or one that contains at least four of the five, etc. In the present use case, to gather a library of *ebo* homologs for neighbor region analyses, each cluster in the list

reported by Yurchenko et al 2016 was submitted to Multigene BLAST using the genbank_mf database of non-redundant protein sequences (accessed 12/2015) with search parameters set to Multigene BLAST defaults. Full *ebo* cluster homologs were defined as cluster-BLAST hits having the same number of genes as the query cluster (five genes). Cluster homologs were amassed into a list by accession number and duplicates were removed. Hits which did not originate from complete genome or plasmid sequences were also removed.

Description of FLANDERS pipeline

The FLANDERS pipeline was written in Python 3.8, and utilizes the subprocess() module to pass arguments to its dependencies, the os() module to deal with file and directory creation and navigation, and the pandas() module to organize data internally and write final output as a .csv file. The pipeline accepts as input either a list of comma-separated accession numbers corresponding to the seed cluster and source genome to be analyzed, or the clusterblast_hits.txt document generated by Multigene BLAST. The pipeline can then be divided into two distinct phases (Fig. 1):

I. Obtaining neighbor region accessions and sequences

The FLANDERS pipeline begins by using the query cluster NCBI protein accession numbers provided by the user or scraped from the clusterblast_hits.txt output from multigene BLAST to generate NCBI protein accession numbers for ORFs upstream and downstream of the query cluster. The number of neighboring genes for which FLANDERS will generate NCBI protein accessions is determined by the user, i.e. if the

user selects 4, FLANDERS will generate accessions for the 4 genes upstream and 4 genes downstream of the query cluster for a total of 8 genes passed to further analyses. By means of the accession numbers generated, the ncbi-acc-download tool downloads corresponding protein FASTA files, while the genome accession number provided by the user for each cluster is used to download whole genome .gbk records. From these, nucleotide sequence of the seed cluster neighboring regions are extracted. For the purpose of the *ebo* cluster, the NCBI accession numbers corresponding to the first and last genes in each *ebo* cluster homolog were used to generate sequential accession numbers for ORFs upstream and downstream of the *ebo* cluster in each genome. The four upstream and four downstream flanking ORF sequences were then retrieved from NCBI using ncbi-acc-download. Nucleotide sequences for upstream and downstream regions were obtained using genomic coordinates specified in whole genome .gff files, extracting the corresponding regions from whole genome .fasta files.

II. Performing sequence-based analyses on neighboring regions

The current iteration of the FLANDRES pipeline includes a series of integrated sequence-based analysis options which can be toggled on or off by the user depending on the biological questions at hand. Presently, FLANDERS uses the Antibiotic and Secondary Metabolite Analysis Shell (antiSMASH) (19) to search query cluster neighboring regions for elements of known BGCs. The taxon rules and gene finding tool that antiSMASH will use on a given FLANDERS run can be specified by the user (FLANDERS default: taxon – bacteria, gene finding tool – prodigal). The antiSMASH

BGC prediction for the upstream and downstream regions of a given query cluster is passed to the FLANDERS output file, but full output is contained in the FLANDRES directory tree under the directory <current_genome>/<current_genome_upstream> or <current_genome_downstream>, for upstream and downstream regions, respectively. To search the *ebo* flanking genomic regions for known secondary metabolite synthesis clusters, the nucleotide sequence of the upstream and downstream regions were searched for known BGCs with antiSMASH 5.0 using bacteria as the query taxon and Prodigal as the gene finding tool. Any known BGCs identified were listed in the FLANDERS output next to the flanking region in which they were found.

To assess the presence of signal peptide containing ORFs upstream and downstream of the query cluster, FLANDERS utilizes SignalP. Protein sequences for each neighboring ORF are submitted to SignalP with user defined parameters (default: organism – gram-). SignalP uses a deep neural-network based method to detect the presence of Sec secretion or twin-arginine translocase (Tat) signal peptides in a query amino acid sequence. After SignalP analysis, FLANDERS parses SingalP output, and writes the signal peptide type detected for each neighbor ORF, as well as the total number of neighboring ORFs that contain a signal peptide for a given query cluster. Additionally, the full SignalP output is contained within the FLANDERS output directory tree under <current_genome>/<upstream_summary.signalp5> or <downstream_summary.signalp> , for upstream or downstream regions, respectively. In order to test the hypothesis that *ebo* associated BGCs are involved in cytoplasmic/periplasmic compartmentalized

biosynthesis, amino acid sequences of *ebo* flanking ORFs were subjected to signal peptide prediction using SignalP. Positive results for the presence of either Type 3 secretion (Sec) or Twin-arginine transport (Tat) were tallied for each ORF in the upstream and downstream *ebo* flanking regions and reported in the FLANDERS output accordingly. In order to verify the significance of finding a given number of signal peptide containing ORFs out of the 8 *ebo* adjacent ORFs, whole genomes were scanned in parallel by SignalP to obtain the total number of signal peptide containing ORFs in the genome. A binomial distribution was then utilized to obtain the probability of selecting the number of signal peptide containing ORFs found in the *ebo* adjacent clusters of each genome by randomly selecting 8 genes from therein.

Lastly, FLANDERS can subject neighboring clusters to a multigene BLAST search to find cluster homologs that may not correspond to known BGCs detected by antiSMASH but may hint towards a conserved *ebo* dependent system of some sort. To accomplish this, protein FASTA files of upstream and downstream ORFs are passed to multigene BLAST protein architecture tool using user determined parameters. The resulting clusterblast_output.txt file for each neighbor cluster is parsed for the number of complete clusters (user defined, as previously stated), and this number reported in the FLANDERS output. Additionally, the clusterblast_output.txt file is deposited in the <current_genome>/<upstream_cluster> or <downstream_cluster> directory respectively in the FLANDERS output directory tree. In the case of the *ebo* cluster, *ebo* adjacent ORFs were themselves run through Multigene BLAST to assess their conservation and distribution. The protein sequence-based architecture search was utilized with default

search parameters. Complete clusters were defined as having an equal number of ORFs to the query cluster. Complete clusters were tallied from each clusterblast hits list and reported in the FLANDERS output. *ebo* adjacent clusters with a higher number of complete cluster hits were considered to be more highly conserved, and therefore may represent putative biosynthetic clusters of unknown natural products.

The FLANDERS output file is a .csv file with columns corresponding to: 1) The source genome NCBI accession number, 2). The location of the analyzed region relative to the seed (query) cluster (upstream or downstream), 3) The NCBI accession numbers corresponding to the upstream, seed, and downstream ORFs, 4) Known BGCs identified by antiSMASH (if any), 5) Signal peptide predictions for each ORF provided by SignalP, 6) Total number of signal peptides detected upstream and downstream of the query cluster, and 7) The number of complete clusters detected by Multigene BLAST for each neighbor region among bacterial genomes.

Results

In order to amass a list of *ebo* cluster homologs beyond that of Yurchenko et al 2016, each *ebo* cluster in the Yurchenko dataset along with the *ebo* cluster from *Pseudomonas fluorescens* NZI7 was submitted to Multigene BLAST from which full *ebo* cluster homologs were obtained. While Yurchenko et al 2016 found *ebo* clusters of various lengths, only complete *ebo* clusters (containing all 6 genes) were used as seed clusters in the present search. Under a less stringent search that considered clusters containing 5 out of 6 to be complete *ebo* clusters, no further *ebo* containing genomes were discovered.

This search yielded 92 genomes containing an *ebo* cluster homolog, many of which (47)

were described by Yurchenko et al 2016 , but some that were not (45), likely due to the differing databases used for the search (the search presented in Yurchenko 2016 was performed with the Genbank non-redundant protein database accessed in 2016). The neighboring regions of these 92 *ebo* cluster homologs originating from whole genome records were used for downstream analyses.

To address the hypothesis that the *ebo* genes are generally colocalized with BGCs, *ebo* adjacent genomic regions were analyzed by antiSMASH to identify any known BGCs that may require the *ebo* genes for proper function. Somewhat surprisingly, none of the 92 species contained an *ebo* adjacent to predictable BGCs. To validate that antiSMASH was functioning properly in the FLANDERS pipeline, the full genome for each organism was searched for BGCs. The majority of genome sequences queried had at least one BGC, but none of them were nearby the *ebo* cluster. This is also the case in *N. punctiforme* ATCC 29133, which is unique amongst scytonemin producing cyanobacteria as the *ebo* gene cluster is found at a distal locus to the scytonemin operon.

Due to the fact that no known BGCs were identified by antiSMASH in any *ebo* adjacent gene clusters, an analysis of the conservation of *ebo* neighbors was performed by feeding *ebo* adjacent ORFs through Multigene BLAST under the same search parameters and database utilized elsewhere in the pipeline. Most *ebo* adjacent clusters queried only had 1-5 full cluster homologs in the genomes of the subject database. Among the *ebo* adjacent clusters that had greater than 20 cluster homologs were the scytonemin synthesis operon of cyanobacteria as well as the homologs of the *edb* cluster from soil *Pseudomonads*. Other *ebo*-adjacent clusters with more than 20 hits included that in the

genome of *Burkholderiales* sp. JOSHI_001, which contained the *ebo* cluster in a putative genomic island populated by transport related genes that including a set of four drug efflux system genes upstream as well as an Fe³⁺ ABC transporter cluster downstream. The *ebo*-adjacent regions of *Azospirillum* sp. B510 had fewer clusterblast hits than JOSHI_001, but were also in a genomic island harbored on a plasmid of transporters containing a cluster of sugar transporters along with a periplasmic binding protein. Of the *ebo*-adjacent clusters with no homology hits in other *ebo* adjacent clusters, that of *Leptospira vanthielii* ATCC 700522 contained a partial sequence of the transposase for insertion element IS1086.

While high throughput means of assessing the presence of BGCs in the present dataset did not yield significant results to known BGCs, a cursory manual examination of *ebo* adjacent gene annotations uncovered a few interesting finds. Biosynthetic gene clusters often include synthases for specific precursors of their cognate secondary metabolite, and therefore the annotations of *ebo* adjacent genes were searched for synthases. Several species of *Leptospira* had specific synthases adjacent to their *ebo* clusters. *Leptospira meyeri* serovar Hardjo str. Went 5 has a strictodisine and glutathione synthases, while *Muricauda eckloniae* and *Aliterella atlantica* CENA595 both have anthranilate synthases.

In both *N. punctiforme* and *P. fluorescens*, some *ebo* flanking biosynthetic genes contain signal peptide domains that target their gene product to the periplasmic space. This finding is consistent with the hypothesis that the *ebo* cluster is involved in transport of natural product precursors during their biosynthesis. To determine if the presence of signal peptide containing encoding ORFs proximal to the *ebo* cluster is a ubiquitous

phenomenon, flanking regions were scanned using SignalP. Indeed, the overwhelming majority of *ebo* clusters examined (85/92) had at least one signal peptide containing ORF within four ORFs upstream and downstream (Fig. 2). A large minority (28/92) of genomes had three flanking signal peptide containing ORFs, while a smaller proportion had between five and seven (11/92) collectively. To test the significance of these findings, the set of relevant whole genomes was submitted to SignalP to assess the number of signal peptide containing genes within each genome. The total number of signal peptide containing genes was divided by the total number of genes for each genome to obtain the probability of selecting a signal peptide containing gene purely at random. This probability was used in a binomial distribution to determine the significance of finding the given number of signal peptide containing ORFs in the 4 upstream and 4 downstream *ebo* adjacent ORFs (Supplementary table 1). Overall, 19 genomes had a less than 5% probability of finding the given number of *ebo* adjacent signal peptide containing ORFs by random, 15 had a probability of less than 10%, and the remaining 57 had a greater than 10% chance (Fig. 2). Probabilities for each genome are listed in supplementary table 1.

Discussion

In prokaryotic genomes, functionally interdependent genes are organized into operons colocalized to allow for their coregulation and coexpression. Given this functional clustering, it follows that the annotations and other sequence characteristics of ORFs upstream and downstream of a gene cluster of interest may be informative of its

biological role. The present work outlines the development and implementation of the FLANDERS pipeline to glean insights about the genomic neighborhood of homologs of the widely conserved *ebo* gene cluster. Previous findings of the *ebo* cluster's genomic colocalization with known BGCs in scytonemin producing cyanobacteria and soil *Pseudomonads* were recapitulated through the FLANDERS pipeline.

The initial goal in applying the FLANDERS pipeline to investigate the *ebo* gene cluster was to identify BGCs for known secondary metabolites that occur adjacent to *ebo* cluster homologs. To this end, *ebo* neighboring regions were passed through the antibiotic and secondary metabolite analysis shell (antiSMASH) to identify known BGCs.

Unexpectedly, there were no known BGCs identified by antiSMASH directly upstream or downstream of any *ebo* cluster homologs surveyed. These results may be due to the fact that the antiSMASH tool provides rule-based detection of only 55 well characterized BGC classes (19), while conservative estimates of the diversity of secondary metabolites across bacteria, fungi, and plants well into the thousands (22).

Because no known BGCs were found to be *ebo* adjacent, neighboring regions were searched for cluster homologs using Multigene BLAST. Most clusters queried had at least one other full cluster hit besides itself, but some had upwards of 20-50. The presence of the scytonemin and *edb* gene clusters among these, a) recapitulated previous findings about the distribution of the scytonemin operon and the *ebo* gene cluster in select terrestrial cyanobacteria and, b) demonstrated that the *edb* cluster is conserved among a variety of soil *Pseudomonads*. These findings prompted further investigation into conserved *ebo* adjacent clusters that did not fall into one of the former two categories.

Burkholderiales bacterium JOSHI_001's *ebo* cluster is neighbored upstream by a cluster containing an EmrB/QacA family drug resistance transporter, a multidrug resistance efflux pump, and an NodT family outer membrane lipoprotein efflux transporter, which is highly conserved in a number of *Acidovorax* and *Rhizobacter* species. Downstream of JOSHI_001's *ebo* cluster are the genes for a canonical ferric iron ABC transporter system which is not highly conserved. The involvement of both of these *ebo* flanking clusters in transport suggests that in the case of JOSHI_001, the *ebo* may be a member of a genomic island dedicated to transport function. Genomic islands are regions of a genome that may have been acquired horizontally and retained due to a resulting increase in fitness (23–25). The inclusion of the *ebo* cluster into this genomic island of transporters provides further support for its role as a generalizable translocation system. Its presence on a putative genomic island may also lend credence to the hypothesis that the distribution of the *ebo* cluster is due to horizontal gene transfer, and that it is a modular apparatus that can plug in to a variety of biosynthetic pathways. Also supportive of this hypothesis was the discovery of an *ebo* cluster homolog harbored on the plasmid pAB510b in *Azospirillum* sp. B510. Directly upstream of the *ebo* cluster on this plasmid is a simple sugar ABC transporter and a periplasmic binding protein, supporting the possibility that *ebo* may be a modular horizontally acquired transport system, in this case by plasmid uptake. Another case of possible horizontal acquisition of the *ebo* cluster was identified in *Leptospira vanthielii* ATCC 700522 in the form of a partial transposon insertion site downstream of the *ebo* cluster. Transposons are mobile genetic elements composed of a transposase gene flanked by insertion sites which are used in homologous recombination (26). They are a ubiquitous mechanism for horizontal gene transfer, and oftentimes leave

behind genomic scarring wherein partial sequences of an inserted transposase may remain. The inclusion of the *ebo* cluster on a transposon may be yet another means of its horizontal transfer across bacteria. Yet another potential means for horizontal gene transfer is phage mediated, wherein a lysogenized bacteriophage will take with it flanking genomic regions during its lytic cycle. Adjacent *ebo* cluster of *Plesiocystis pacifica* SIR-1 contains a cluster of hypothetical proteins and a putative phage derived ORF, suggesting that the *ebo* cluster may have originated in its genome via specialized transduction. While no highly conserved novel BGCs were detected using the present methodology, future studies utilizing more recent databases may allow for a broader detection threshold by which to consider a given *ebo* cluster neighboring region to be considered highly conserved. Furthermore, the presence of specific synthases adjacent to *ebo* clusters in select genomes suggests that the synthesis of these compounds may be *ebo* facilitated. Several species of *Leptospira* had specific synthases adjacent to their *ebo* clusters. *Leptospira meyeri* serovar Hardjo str. Went 5 has a strictosidine as well as a glutathione synthase upstream of its *ebo* cluster. Strictosidine is a precursor for a variety of monoterpene indole alkaloids, many of which have anticancer, anti-malarial and anti-arrhythmic activities (27). Glutathione is a potent bacterial antioxidant (28).

In both the scytonemin and *edb* systems, several ORFs contain signal peptide sequences; N-terminal amino acids that target their preproprotein to the periplasmic space for proper folding and function (16). The presence of these signal peptides in the scytonemin synthesis genes *scyDEF* is consistent with the role of the *ebo* gene cluster being translocation of the scytonemin monomer to the periplasm during synthesis (17), and

prompted the search for signal peptides in neighboring ORFs of *ebo* cluster homologs. The overwhelming majority of *ebo* adjacent clusters contained at least one signal peptide containing ORF (Fig. 2), supporting the hypothesis that the *ebo* cluster acts as a modular transport system which can plug in to a variety of biosynthetic pathways which have cytoplasmic and periplasmic stages. Given that nearly half (42) of the *ebo* adjacent clusters examined had 3 or 4 signal peptide containing genes, and that the probability of obtaining these results randomly is quite low for a majority of cases, we see the finding as supportive of the *ebo* clusters role in transport. Likewise, there were no cases where all four upstream and four downstream genes contained signal peptides, consistent with biosynthetic pathways that involve both cytoplasmic and periplasmic stages. Over one third of all genomes had a probability of finding their respective number of signal peptides at random of less than 10%, suggesting that in many cases, *ebo* adjacent ORFs are enriched in signal peptides. A small number of *ebo* neighbors queried did not contain any signal peptides, but two of the six (*Tolypothrix* sp. PCC 7601 and *Robiginitalea biformata* HTCC2501) contained multiple ABC transporters and MFS permeases respectively proximal to their *ebo* cluster. Both of these gene products are also involved in metabolite transport, and may be members of a genomic island dedicated to transport functions (29, 30). The role of the *ebo* cluster as a nascent secondary metabolite transporter (17) may support its inclusion in such a transport genomic island. To determine if this colocalization could be due to random distribution of transporters throughout the genome, the total number of transport related genes was compared to total genes to obtain the probabilities of obtaining at least 2 transporters out of 8 randomly selected genes (Supplementary table 1). For *Tolypothrix* PCC. 7601, this probability was

7% and for *Robiginitalea biformata* HTCC2501 it was 6%. These probabilities indicate that these findings may be purely due to randomness, and the total number of transporters in each genome and associated probabilities are listed in supplementary table 1. In the case of signal peptide containing ORFs, across all genomes, the average probability of finding at least 2 signal peptides out of 8 random selected genes was 26%, however the likelihoods of finding 3 or 4 were 7% and 2% respectively. Many species of *Tolypothrix* and *Lyngbya* are scytonemin producing, however strains PCC 7601 and PCC 8106 respectively do not contain the canonical scytonemin synthesis operon, and perhaps its *ebo* cluster is a remnant from a scytonemin producing ancestral strain. Lastly, *Robiginitalea biformata* HTCC2501, *Muricauda ruestringensis*, and *Zobellia galactanivorans* strain DsiJT all had a polyphosphate kinase and phosphatase proximal to their *ebo* cluster. Bacteria utilize polyphosphate as a means of sequestering inorganic phosphorous as an insurance plan against phosphorous limited conditions (31). While the synthesis of polyphosphate is generally cytoplasmic, some species have been reported to accumulate polyphosphate in both the cytoplasm and periplasm (32). In the case of these *ebo* adjacent polyphosphate synthesis clusters, cytoplasmically synthesized polyphosphate may potentially be translocated to the periplasm using the *ebo* gene products as the central or an ancillary means of transport. In this case, there would be no need for any of the *ebo* adjacent ORFs to contain signal peptides, as polyphosphate may be translocated to the periplasm fully polymerized, thus negating the need for periplasmic synthesis steps.

The application of the FLANDERS pipeline to probe the function and distribution of *ebo* adjacent gene clusters provided further evidence for both its role as a modular transport mechanism as well as its probable horizontal origins in many species. Future iterations of the pipeline will include further protein level analyses such as prediction of membrane spanning domains, which are common in well described bacterial transport systems, as well as genomic context analyses to discern if the *ebo* cluster is part of a genomic island. In a broader sense, these results indicate the utility of the FLANDERS pipeline towards gleaning insights for any conserved gene cluster which may have modular function in a variety of genomic contexts. As with any bioinformatic analyses, laboratory experimentation will be required to support or reject any hypotheses conceived by way of the FLANDERS pipeline. The FLANDERS pipeline represents a simple yet effective tool for exploring the ways in which prokaryotic gene organization and functional interdependence interact towards the synthesis and trafficking of secondary metabolites, as well as other cellular functions.

References

1. Artsimovitch I. 2018. Rebuilding the bridge between transcription and translation. Mol Microbiol. Blackwell Publishing Ltd.
2. Stent GS. 1964. The Operon: On Its Third Anniversary New Series.
3. Fondi M, Emiliani G, Fani R. 2009. Origin and evolution of operons and metabolic pathways. Res Microbiol 160:502–512.
4. Chen H, Du L. 2016. Iterative polyketide biosynthesis by modular polyketide synthases in bacteria. Appl Microbiol Biotechnol 100:541–557.
5. Wlodek A, Kendrew SG, Coates NJ, Hold A, Pogwizd J, Rudder S, Sheehan LS, Higginbotham SJ, Stanley-Smith AE, Warneck T, Nur-E-Alam M, Radzom M, Martin CJ, Overvoorde L, Samborsky M, Alt S, Heine D, Carter GT, Graziani EI, Koehn FE, McDonald L, Alanine A, Rodríguez Sarmiento RM, Chao SK, Ratni H, Steward L, Norville IH, Sarkar-Tyson M, Moss SJ, Leadlay PF, Wilkinson B,

- Gregory MA. 2017. Diversity oriented biosynthesis via accelerated evolution of modular gene clusters. *Nat Commun* 8:1206.
6. Gomez-Escribano JP, Song L, Fox DJ, Yeo V, Bibb MJ, Challis GL. 2012. Structure and biosynthesis of the unusual polyketide alkaloid coelimycin P1, a metabolic product of the *cpk* gene cluster of *Streptomyces coelicolor* M145. *Chem Sci* 3:2716–2720.
 7. Zhang L, Hashimoto T, Qin B, Hashimoto J, Kozone I, Kawahara T, Okada M, Awakawa T, Ito T, Asakawa Y, Ueki M, Takahashi S, Osada H, Wakimoto T, Ikeda H, Shin-ya K, Abe I. 2017. Characterization of Giant Modular PKSs Provides Insight into Genetic Mechanism for Structural Diversification of Aminopolyol Polyketides. *Angew Chemie* 129:1766–1771.
 8. Baltz RH. 2006. Molecular engineering approaches to peptide, polyketide and other antibiotics. *Nat Biotechnol* 24:1533–1540.
 9. Schwarzer D, Finking R, Marahiel MA. 2003. Nonribosomal peptides: From genes to products. *Nat Prod Rep* 20:275–287.
 10. Vasudevan R, Gale GAR, Schiavon AA, Puzorjov A, Malin J, Gillespie MD, Vavitsas K, Zulkower V, Wang B, Howe CJ, Lea-Smith DJ, McCormick AJ. 2019. CyanoGate: A Modular Cloning Suite for Engineering Cyanobacteria Based on the Plant MoClo Syntax. *Plant Physiol* 180:39–55.
 11. Klein CA, Emde L, Kuijpers A, Sobetzko P. 2019. MoCloFlex: A Modular Yet Flexible Cloning System. *Front Bioeng Biotechnol* 7:271.
 12. Hausner J, Jordan M, Otten C, Marillonnet S, Büttner D. 2019. Modular Cloning of the Type III Secretion Gene Cluster from the Plant-Pathogenic Bacterium *Xanthomonas euvesicatoria*. *ACS Synth Biol* 8:532–547.
 13. Jensen PR. 2016. Natural Products and the Gene Cluster Revolution. *Trends Microbiol* 24:968–977.
 14. Yurchenko T, Ševčíková T, Strnad H, Butenko A, Eliáš M. 2016. The plastid genome of some eustigmatophyte algae harbours a bacteria-derived six-gene cluster for biosynthesis of a novel secondary metabolite. *Open Biol* 6.
 15. Burlinson P, Studholme D, Cambray-Young J, Heavens D, Rathjen J, Hodgkin J, Preston GM. 2013. *Pseudomonas fluorescens* NZI7 repels grazing by *C. elegans*, a natural predator. *ISME J* 7:1126–1138.
 16. von Heijne G. 1990. The signal peptide. *J Membr Biol* 115:195–201.
 17. Klicki K, Ferreira D, Hamill D, Dirks B, Mitchell N, Garcia-Pichel F. 2018. The Widely Conserved *ebo* Cluster Is Involved in Precursor Transport to the Periplasm during Scytonemin Synthesis in *Nostoc punctiforme*. *mBio* 9: e02266-18.
 18. Blin K. 2020. ncbi-acc-download. 0.2.7.

19. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* 47:W81–W87.
20. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 37:420–423.
21. Medema MH, Takano E, Breitling R. 2013. Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol Biol Evol* 30:1218–23.
22. Da T, Agostini-Costa S, Vieira RF, Bizzo HR, Silveira D, Gimenes MA. 8 Secondary Metabolites.
23. Bertelli C, Laird MR, Williams KP, Lau BY, Hoad G, Winsor GL, Brinkman FSL. 2017. IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res* 45:W30–W35.
24. Juhas M. 2015. Type IV secretion systems and genomic islands-mediated horizontal gene transfer in *Pseudomonas* and *Haemophilus*. *Microbiol Res* 170:10–17.
25. Penn K, Jenkins C, Nett M, Udvary DW, Gontang EA, McGlinchey RP, Foster B, Lapidus A, Podell S, Allen EE, Moore BS, Jensen PR. 2009. Genomic islands link secondary metabolism to functional adaptation in marine Actinobacteria. *ISME J* 3:1193–1203.
26. Babakhani S, Oloomi M. 2018. Transposons: the agents of antibiotic resistance in bacteria. *J Basic Microbiol.* 58:905-917.
27. O’connor SE, Maresh JJ. 2005. Chemistry and biology of monoterpene indole alkaloid biosynthesis. *Nat Prod Rep* 23:523–547.
28. Smirnova GV, Oktyabrsky ON. 2005. Glutathione in Bacteria. *Biochemistry* 70:1459–1473.
29. Quistgaard EM, Löw C, Guettou F, Nordlund P. 2016. Understanding transport by the major facilitator superfamily (MFS): Structures pave the way. *Nat Rev Mol Cell Biol* 17:123–132.
30. Hollenstein K, Dawson RJ, Locher KP. 2007. Structure and mechanism of ABC transporter proteins. *Curr Opin Struct Biol* 17:412–418.
31. Kornberg A, Rao NN, Ault-Riché D. 1999. Inorganic Polyphosphate: A Molecule of Many Functions. *Annu Rev Biochem* 68:89–125.
32. Ii B, Freiburg D-. 1991. in *Pure Cultures of Sewage Bacteria* 25:9–13.

Tables and Figures

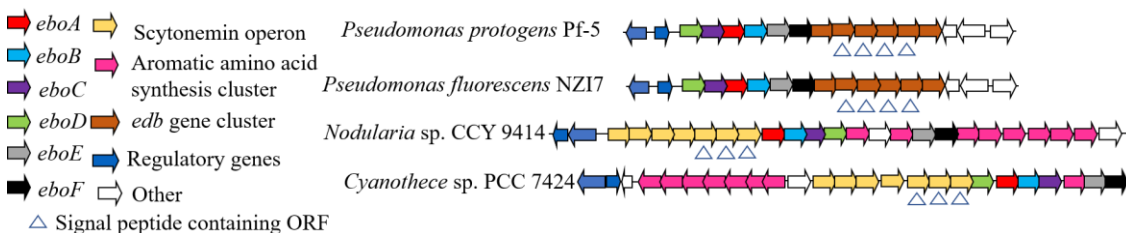


Fig. 1: Arrangement of the *ebo* gene cluster in known biosynthesis contexts. In the scytonemin producing cyanobacteria *Nodularia* sp. CCY 9414 and *Cyanothece* sp. PCC 7424 the *ebo* cluster is intercalated into the scytonemin synthesis operon and its associated aromatic amino acid synthesis cluster. In the soil gammaproteobacteria *Pseudomonas protogens* Pf-5 and *Pseudomonas fluorescens* NZI7, the *ebo* cluster is situated upstream of the *edb* cluster, the suspected biosynthetic cluster of a putative anti-nematodal secondary metabolite. Signal peptide containing ORFs are enriched around the scytonemin and *edb* gene clusters and are indicated with a white arrow.

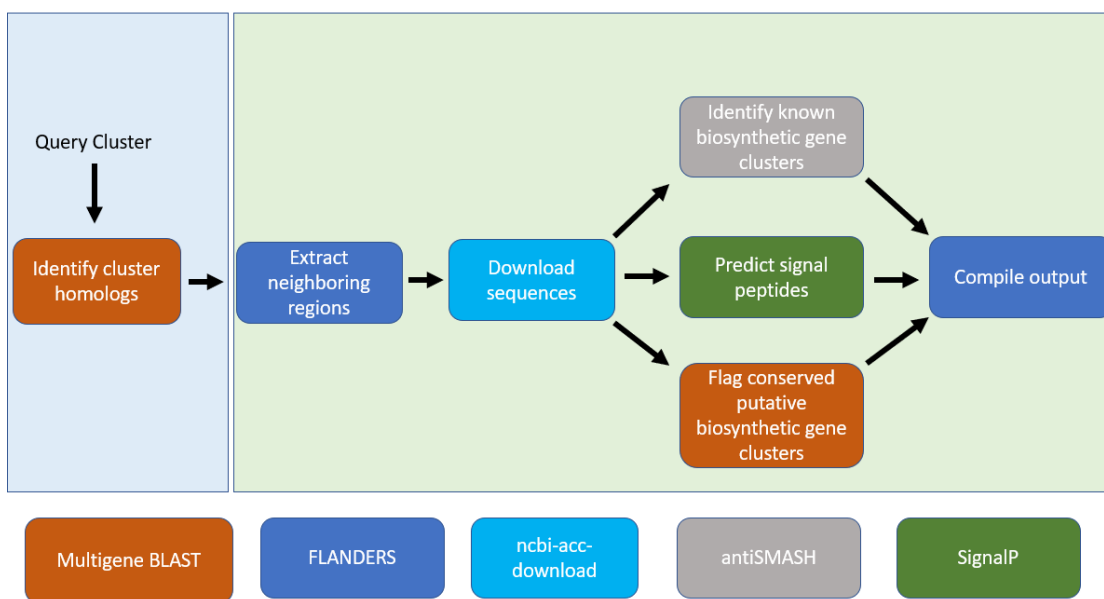


Fig. 2 Outline of the FLANDERS pipeline: A query gene cluster is searched against a user selected database for cluster homologs by multigene BLAST. Accession numbers for neighboring ORFs were generated and sequences downloaded from NCBI. These sequences were then passed to antiSMASH for detection of BGCs, SignalP to predict the presence of signal peptides, and multigene BLAST to assess conservation of neighboring clusters. The results for each analysis are then written to a convenient csv file for easy parsing. Blue background represents cluster homolog discovery stage prior to FLANDERS pipeline and green background represents stages of the FLANDERS pipeline. Steps are color coded to their respective dependency listed at bottom.

Table 1: Selected results of multigene BLAST search for *ebo* cluster homologs and FLANDERS assessment of neighboring gene clusters. Species names and accession numbers are listed for each genome in which an *ebo* cluster homolog was found. Total signal peptides correspond to the number of signal peptide containing genes out of the 4 upstream and 4 downstream genes. Total clusterblast hits correspond to the number of complete cluster homologs that were found for the upstream and downstream regions adjacent to the *ebo* cluster in each genome. *ebo* adjacent gene clusters that were previously described are listed in the known cluster column.

Species	Accession	Total signal peptides	Total clusterblast hits	Known cluster
<i>Lyngbya</i> sp. PCC 8106	AAVU01000005	5	21	Scytonemin (pseudocluster)
<i>Crocospaera chwakensis</i> CCY0110	AAXW01000042	2	2	
<i>Pedobacter</i> sp. BAL39	ABCM01000004	2	11	
<i>Plesiocystis pacifica</i> SIR-1	ABCS01000019	0	2	
<i>Pseudomonas fluorescens</i> WH6	AEAZ01000019	7	48	<i>edb</i>
<i>Tolypothrix</i> sp. PCC 7601 (Fremyella diplosiphon CCAP 1429/1)	AGCR01000110	0	6	Scytonemin
<i>Niabella soli</i> DSM 19437	AGSA01000026	5	3	
<i>Pseudomonas fluorescens</i> Q8r1-96	AHPO01000004	3	0	<i>edb</i>
<i>Cellvibrio</i> sp. BR	AICM01000008	5	5	
<i>Pseudomonas</i> sp. GM102	AKJB01000118	3	36	<i>edb</i>
<i>Pseudomonas</i> sp. GM50	AKJK01000048	3	34	<i>edb</i>
<i>Pseudomonas</i> sp. GM18	AKJT01000074	4	26	<i>edb</i>
<i>Leptospira meyeri</i> serovar Hardjo str. Went 5	AKXE01000002	2	0	
<i>Sorangium cellulosum</i> So ce56	AM746676	3	6	

Pseudomonas putida CSV86	AMWJ010001 64	1	3	<i>edb</i>
Mariniradius saccharolyticus AK6	AMZY020000 17	3	0	
Leptospira meyeri serovar Semaranga str. Veldrot Semarang 173	ANIL0100000 4	0	0	
Leptospira terpstrae serovar Hualin str. LT 11-33 = ATCC 700639	AOGW020000 18	2	0	
Leptospira yanagawae serovar Saopaulo str. Sao Paulo = ATCC 700523	AOGX020000 44	2	0	
Leptospira vanthielii serovar Holland str. Waz Holland = ATCC	AOGY020000 69	1	0	
Leptospira wolbachii serovar Codice str. CDC	AOGZ0200001 3	1	0	
Azospirillum sp. B510	AP010948	1	9	
Pseudomonas protegens Cab57	AP014522	4	29	<i>edb</i>
Winogradskyella sp. PG-2	AP014583	2	2	
Pseudomonas sp. Os17	AP014627	3	31	<i>edb</i>
Pseudomonas sp. St29	AP014628	3	31	<i>edb</i>
Lunatimonas lonarensis	AQHR0100011 0	1	1	
Roseivivax marinus	AQQW010000 12	2	3	
Catenovulum agarivorans DS-2	ARZY0100000 3	5	2	
Pseudomonas alcaligenes OT 69	ATCP0100019 1	2	27	<i>edb</i>
Pseudomonas sp. CFII68	ATLN0100004 0	3	26	<i>edb</i>
Cyclobacterium qasimii M12-11B	ATNM010001 34	1	3	
Pseudomonas mediterranea CFBP 5447	AUPB0100004 1	5	26	<i>edb</i>
Pseudomonas aestus	AVOY010000 99	3	3	
Pedobacter sp. V48	AWRU010000 32	3	2	

Rhodonellum psychrophilum GCM71 = DSM 17998	AWXR01000002	2	2	
Nostoc sp. PCC 7120 = FACHB-418 (Anabaena sp. PCC 7120)	BA000019	3	24	Scytonemin
Sporocytophaga myxococcoides	BBLT01000001	2	2	
Pseudomonas sp. SHC52	CBLV01000003	4	27	<i>edb</i>
Burkholderiales bacterium JOSHI_001	CM001438	2	51	
Pseudomonas protegens Pf-5	CP000076	4	29	<i>edb</i>
Leptospira biflexa serovar Patoc strain 'Patoc 1 (Ames)'	CP000777	2	13	
Leptospira biflexa serovar Patoc strain 'Patoc 1 (Paris)'	CP000786	1	16	
Crocospaera subtropica ATCC 51142	CP000806	2	3	
Nostoc punctiforme PCC 73102 (Nostoc punctiforme ATCC 29133)	CP001037	1	1	Scytonemin (Distal)
Gloeotheca citrififormis PCC 7424	CP001291	4	16	Scytonemin
Teredinibacter turnerae T7901	CP001614	2	0	
Dyadobacter fermentans DSM 18053	CP001619	1	3	
Chitinophaga pinensis DSM 2588	CP001699	1	3	
Robiginitalea biformata HTCC2501	CP001712	0	8	
Ketogulonicigenium vulgare WSH-001	CP002019	3	4	
Gloeotheca verrucosa PCC 7822	CP002198	2	3	
Ketogulonicigenium vulgare Y25	CP002225	3	4	
Leadbetterella byssophila DSM 17132	CP002305	3	2	
Sphingobacterium sp. 21	CP002584	4	2	

Pseudomonas brassicacearum subsp. brassicacearum NFM421	CP002585	4	29	<i>edb</i>
Runella slithyformis DSM 19594	CP002859	3	2	
Cyclobacterium marinum DSM 745	CP002955	2	17	
Muricauda ruestringensis DSM 13258	CP002999	0	25	
Pseudomonas protegens CHA0	CP003190	3	27	<i>edb</i>
Chroococcidiopsis thermalis PCC 7203	CP003597	2	20	
Gloeocapsa sp. PCC 7428	CP003646	1	17	Scytonemin (Distal?)
Sorangium cellulosum So0157-2	CP003969	2	2	
Pseudomonas sp. TKP	CP006852	3	9	<i>edb</i>
Pseudomonas cichorii JBC1	CP007039	4	76	<i>edb</i>
Sphingobacterium sp. ML3W	CP009278	4	3	
Pseudomonas chlororaphis subsp. aurantiaca	CP009290	1	51	<i>edb</i>
Pseudomonas soli	CP009365	6	29	<i>edb</i>
Pseudomonas parafulva	CP009747	2	1	
Hymenobacter sp. DG25B	CP010054	3	3	
Cyclobacterium amurskyense	CP012040	1	18	
Pseudomonas entomophila L48	CT573326	1	5	
Zobellia galactanivorans	FP476056	0	25	
Anditalea andensis	JMIH01000034	3	2	
Pedobacter antarcticus 4BY	JNFF01000045	4	2	
Pedobacter kyungheensis	JSYN01000004	4	16	
Scytonema millei VB511283	JTJC01000086	1	2	
Pseudomonas batumici	JXDG01000040	2	2	
Pseudomonas sp. FeS53a	JYFT01000015	4	27	<i>edb</i>
Aliterella atlantica CENA595	JYON01000001	1	20	

Muricauda eckloniae	LCTZ0100000 2	1	7	
Pseudomonas lini	LFQO0100003 8	3	36	<i>edb</i>
Algoriphagus marincola HL-49 (microbial mat metagenome)	LJXT01000039	3	2	
Pseudomonas sp. Root401	LMDO010000 01	3	28	<i>edb</i>
Pseudomonas sp. Leaf15	LMKI0100003 6	3	29	<i>edb</i>
Pedobacter sp. Leaf216	LMKM010000 20	3	1	
Pedobacter sp. Leaf176	LMPD0100000 4	3	5	
Dyadobacter sp. Leaf189	LMPS0100000 4	2	4	
Pedobacter sp. Leaf194	LMPU0100000 1	2	2	
Pedobacter ginsenosidimutans	LMZQ0100000 9	3	1	
Pseudomonas sp. CCOS 191	LN847264	6	29	<i>edb</i>

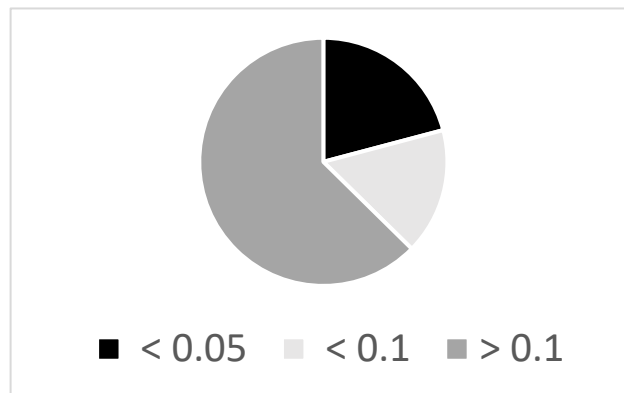
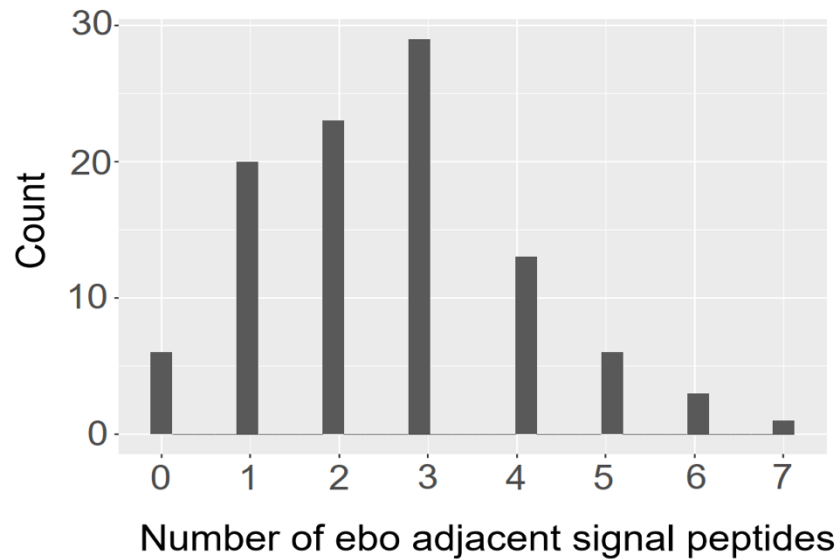


Fig. 3: Top: Number of signal peptide containing ORFs upstream and downstream of *ebo* cluster homologs surveyed. Four upstream and four downstream ORF sequences for each *ebo* cluster were scanned for signal peptide domains using signalP. The overwhelming majority of *ebo* adjacent clusters had at least one signal peptide containing ORF, while many had multiple. Bottom: proportion of genomes with a probability of finding the number of signal peptide containing ORFs found through FLANDERS in the 8 *ebo* adjacent by random given the total number of signal peptide containing ORFs in the genome.

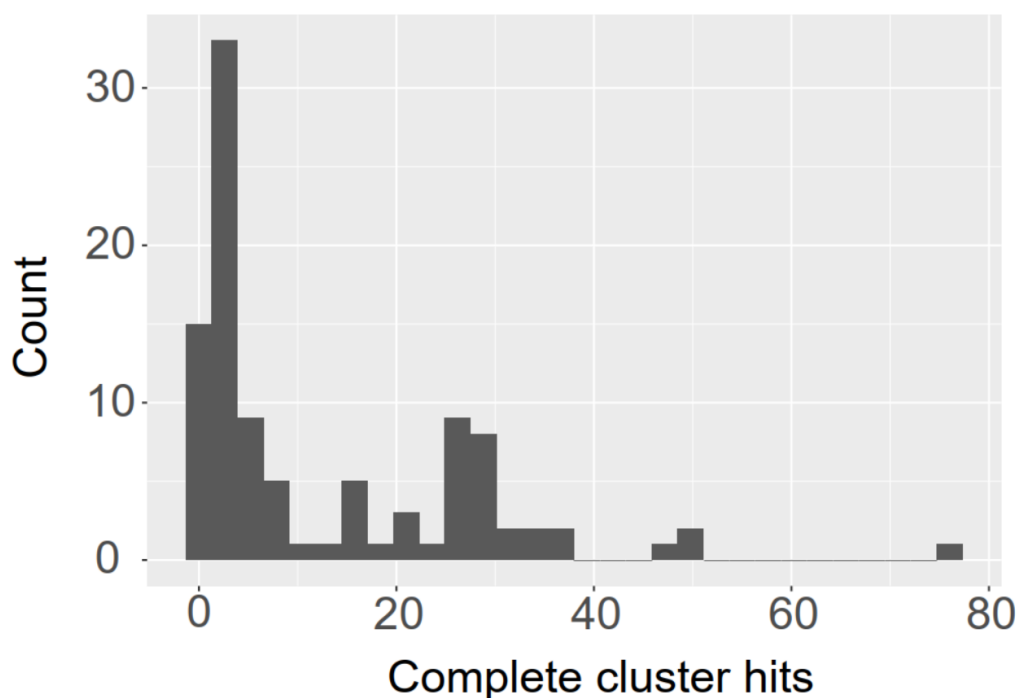


Fig. 4: Frequency distribution of complete clusterblast hits. *ebo* adjacent gene clusters that did not correspond to a known BGC were searched for cluster homologs using multigene BLAST to assess their conservation. Clusters which contained homologs of all four query genes were considered to be whole cluster homologs. A sizeable minority of query clusters had 20-40 homologs detected, the majority corresponded to *edb* cluster containing *Pseudomonads* and scytonemin producing cyanobacteria. Another large majority only had 1-2 homologs, suggesting that these clusters may be very specific to a given organism, or perhaps ancestral and non-functional.

4 - BET HEDGING SUNSCREEN PRODUCTION AND MOTILITY RESPONSES
AGAINST UV EXPOSURE IN A CYANOBACTERIUM

Submitted for publication to Current Biology, June 2021

Coauthors have acknowledged the use of this manuscript in my dissertation.

Kevin Klicki^{1,2}, Daniela Ferreira², Douglas Risser³, Ferran Garcia-Pichel^{1,2}

¹ Center for Fundamental and Applied Microbiomics, Biodesign Institute, Arizona State University, Tempe, AZ, 85281

² School of Life Sciences, Arizona State University, AZ 95281

³ Department of Biological Sciences, University of the Pacific, Stockton CA, 95211

Abstract

In the face of stressors, bacteria sometimes hedge their survival bets by activating heterogenous response circuits that lead to more than one phenotype in an isogenic population. We show that isogenic populations of the filamentous cyanobacterium *Nostoc punctiforme* respond to UV-A stress by concurrently producing the sunscreen scytonemin and developing into motile hormogonia, but segregating the responses at the filament level. We identify a four-gene regulatory system (*hcyA-D*) that mediates the cross-talk between the respective hormogonia and scytonemin canonical regulatory circuitry. On the one hand, the transcription of typical hormogonial genes and that of a partner-switching regulatory system (*hcyA-D*) is upregulated by UVA irradiation through the scytonemin two component regulator (scyTRC), as deduced from transcriptomic analyses of scyTRC knock-out mutants. Analyses of knockout mutants in the *hcyA-D* system show its direct involvement in mediating UVA signal transduction into the hormogonium response, and its role of modulating this response based on visible light cues. On the other hand, we show that the sigma factor cascade known to regulate developmental commitment to hormogonia differentiation can also activate transcription of *hcyA-D* system components and strongly suppress the synthesis of scytonemin through downregulation of the scyTRC itself. Through this complex bidirectional mechanism, *Nostoc* is apparently able to utilize two fundamentally different UV stress mitigation strategies, either hunker down or flee, in a single population.

Introduction

Organisms in nature are subject to a wide range of often concurrent environmental stressors. In response, adaptive strategies for sensing and responding to these environmental challenges have evolved at both the molecular and cellular levels. The classical paradigm of environmental stress response in bacteria consists of a molecular circuit that modulates cellular activity by exerting transcriptional, post-transcriptional, or post-translational control over processes relevant to the environmental stimulus (1, 2). This model implies that there is a one-to-one relationship between stress factor and stress response. In reality, even genetically identical populations of bacteria can exhibit phenotypic heterogeneity in response to a uniform environmental stimulus and yield differing responses from individual cells by a combination of genetically programmed and stochastic mechanisms (3, 4). One evolutionary strategy that exploits phenotypic heterogeneity to increase population fitness is bet-hedging, wherein phenotypic diversity in an isogenic population allows for the population as a whole to enhance its combined probability of survival to environmental fluctuations that may be either too rapid to be addressed by genetic regulation, or for which the maintenance of separate, dedicated signal transduction systems would be too costly (5, 6). This approach may be particularly useful to an organism subject to a simultaneous barrage, or a rapidly oscillating variety of environmental stressors (3, 7). A range of molecular mechanisms leading to phenotypic variation in an isogenic population have been identified that can be epigenetic, stochastic, and cell-cycle based in nature(8, 9). Importantly, as subpopulations commit to diverging

responses, they must regulate cellular functions in a manner consequent with that commitment.

Among microbes, cyanobacteria are particularly prone to radiative environmental insults in that their photosynthetic metabolism exposes them by necessity to solar radiation (10). Many cyanobacteria are dominant in habitats characterized by high light exposure (11–13), in which exposure to ultra-violet radiation (UVR) can bring about significant collateral damage, and have thus developed a suite of adaptations to mitigate its deleterious effects, including DNA and protein repair, behavioral avoidance of exposure by migration patterns, the synthesis of antioxidants, or the production of UV absorbing sunscreens (14, 15). While some of these seem to be universal among cyanobacteria, like DNA damage repair or specific replacements of the D1 protein in photosystem one (15), some are of restricted distribution. Only some cyanobacteria, for example, can produce the sunscreen scytonemin (16), a brownish-yellow lipid soluble sunscreen that accumulates in the extracellular polysaccharide layer (EPS) in response to UV-A (315-400 nm) exposure (17). This response is regarded as a passive adaptation particularly useful in cyanobacteria that endure periods of quiescence(18). By contrast, active behavioral responses based on photophobic responses to UV have only been described for filamentous, non-heterocystous cyanobacteria like *Oscillatoria spp.* (19) and *Microcoleus spp* (20). These two types of adaptations would seem to be mutually exclusive given this ecological niche separation, in other words, either the organisms hunker down and endure the insults, or they flee. *Nostoc punctiforme* has become the model organism to study scytonemin, as it is the only producer that is genetically

tractable. Although *N. punctiforme* is sessile (i.e., non-motile) in its vegetative life stage, it is also capable of motility responses after undergoing a developmental cycle to form hormogonia, which are specialized short filaments that are characteristically motile and buoyant and that serve as a means for colony dispersal. Fluctuating nutrient or salt concentrations, diminishing light intensity, or the presence of symbiotic partners (21) will all elicit in *N. punctiforme* the genetic expression patterns that lead to differentiation of motile hormogonia, which will ‘glide’ to an area more conducive to growth and promptly de-differentiate into sessile, colonial vegetative forms (21–23). In general, the genus *Nostoc* is widely distributed in the terrestrial environment, from Antarctic valleys to arid deserts (24), where it is often found as a symbiont to a variety of plants and fungi (24, 25). During the establishment of symbiosis, *N. punctiforme* will differentiate into hormogonia in response to nutrient conditions and/or plant secondary messengers (26, 27).

Gene expression for both types of adaptations in *Nostoc* is tightly controlled. In the case of scytonemin, a canonical (28) two-component regulatory system (henceforth the scytonemin TCR or scyTCR) directly controls its downstream polycistronic biosynthetic operon (29). The scyTCR is essential for scytonemin production (30), and its histidine kinase contains chromophore binding motifs; while the chromophore is unknown, it is likely responsible for sensing its environmental stimulus, UV-A radiation. Hormogonia differentiation requires several disparate cellular functions to work in concert, including a specialized type IV pilus system that provides the primary motive force to carry filaments along (31), gas vesicles to provide buoyancy in aqueous environments (32), the

production of specific polysaccharide, as well as genes involved in reductive cell division such as *ftsZ* and *ftsY* (33). Most of the hormogonia related genes are coregulated by a tripartite hierarchical sigma factor cascade (33), consisting of sigma factors J, F and C, that chaperone the RNA polymerase to their transcriptional targets (34).

Because UV-A radiation exposure in *Nostoc* can elicit a large number of presumably adaptive responses beyond scytonemin synthesis (35) for which the regulatory components are unknown, we began a survey of regulatory targets of the scytonemin TCR by RNA sequencing of scytonemin TCR deletion mutants under the hypothesis that its regulatory capacity might be more general than previously thought. Surprisingly, these analyses provided evidence for a regulatory cross talk between hormogonia differentiation and scytonemin synthesis, which we explore further in this contribution.

Materials and Methods

Strains and culture conditions

Mutant construction

In-frame deletion mutants Δ Reg, Δ RR, Δ sigJ, Δ hcyA and Δ hcyC were generated from their respective parent strain by homologous recombination as described in Cai and Wolk 1990 (36). PCR primers were designed to amplify DNA upstream and downstream of the desired deletion site (1.0 kb to 3.0 kb on each side) and containing overlapping sequences (36, 37) (Table S1). PCR products were checked and cloned into carrier plasmid pRL278. Plasmids were introduced into the wild type by conjugation, and selected by a

combination of antibiotic and sucrose counter selection markers, as explained in detail in Gonzalez et al 2019 (33).

Phenotypic characterization

Scytonemin induction assays. Scytonemin production was induced by growing the strains in 15x100mm glass petri dishes under white light at $50 \mu\text{mol photons m}^{-2}\text{s}^{-1}$ supplemented with a UV-A flux of $7.5 \mu\text{mol photons m}^{-2}\text{s}^{-1}$ supplied by black-light fluorescent bulbs (General Electric) for 5 days at 21°C . For some experiments, as noted where applicable, white light was omitted. Scytonemin production was detected by microscopic observation first, using a Carl Zeiss Axioscope.A1 at 100x magnification and by HPLC on a Waters e2695 equipped with a Supelco Discovery HS F5-5 column connected to a Waters 2998 PDA UV-Vis diode array detector as described in Klicki et al 2018 (38).

Standard Hormogonia induction assays. Hormogonia induction was carried out following Khayatan et al. 2015 (31): *Nostoc* strains were streaked onto agar-solidified medium supplemented with 5% sucrose and grown heterotrophically in the dark until punctate colonies formed. Colonies were then transferred to agar-solidified media lacking sucrose and exposed to $10 \mu\text{mol photons m}^{-2}\text{s}^{-1}$ provided at a 90-degree angle. Formation of hormogonia was detected visually under the dissection scope. Plate motility assays were performed as previously described (31). This assay yields a massive differentiation of filaments involving the large majority of existing filaments.

Hormogonia/ scytonemin co-induction was carried out on solid AA medium using aluminum foil wrapped glass petri plates provided with a side window, so that white light could enter at a 90-degree angle at an intensity of $10 \mu\text{mol photons m}^{-2}\text{s}^{-1}$. UV-A radiation was concurrently delivered from a top window, after passing through a 400 nm cut-off short-pass optical filter (Edmund Optics) to remove any visible light produced by the UV-A bulbs. This was necessary because we found in pilot experiments that the additional vertically supplied visible light inhibited standard phototactic hormogonia induction. For analyses, scytonemin/hormogonia co-induction cultures were transferred from agar plates to microscope slides using a scalpel and photographed on a Carl Zeiss Axioscope.A1 at 1000x magnification using an Axiocam 105 color camera. They were inspected for morphologically identifiable hormogonia, and for the typical extracellular coloration imparted by scytonemin. UV-A exclusive hormogonia induction was carried out by plating strains from liquid cultures on to solid AA medium and incubating under UV-A irradiation without supplemental PAR for 48-96 hours. Hormogonia spreading was quantified by measuring the area occupied by initial colony biomass and the area swept out by whisks of hormogonia using the Carl Zeiss Zen 3.1 software, subtracting the former from the latter, and dividing the difference by the initial biomass area.

RNA extraction, sequencing and RNAseq analyses

Total RNA was extracted from appropriately incubated cultures (UV-A-induced or uninduced) of each test strain (wild type, ΔRR and ΔReg) using previously described methods (39). Ribodepletion, cDNA synthesis and subsequent RNA sequencing were performed at the Arizona State University Genomics Facility as follows. Four μg of total

RNA was submitted to ribodepletion using the MICROBExpress Bacterial mRNA Enrichment Kit (Ambion). Directional cDNA libraries were synthesized using rRNA depleted samples. Three libraries per strain were multiplexed and treatments sequenced independently on an Illumina NextSeq 500 using a v2 2x75 pair end kit sequenced generating 76 basepair reads, obtaining on average 9,773,321 +/- 5,550,035 reads per sample, excluding 16s and 23s rRNA genes. Sequences were aligned and assembled to the *Nostoc punctiforme* PCC 73102 genome using Rockhopper (40) on default parameters independently for each replicate. Differential expression between mutants and UV treatments was calculated using the DEseq algorithm in Rockhopper (40). Given the roughly 6000 non rRNA genes in the *N. punctiforme* genome, only transcripts whose differential expression had a q-value (a false discovery rate corrected p-value) of less than 10^{-15} were considered significant for downstream analyses. Contiguous genes of known function that exhibited similar degrees of differential expression were considered to be co-regulated. Reads per kilobase million for each gene from the Rockhopper output were square root transformed to accentuate expressional differences over a broader dynamic range. Heatmaps were constructed using Pheatmap (41). Raw RNA sequencing data can be accessed through the NCBI Sequence Read Archive with accession number PRJNA728749.

Targeted analysis by RT-qPCR

Five-hundred ng of total RNA was used to synthesize cDNA using the Protoscript first-strand cDNA synthesis kit and random hexamer primers (New England Biolabs) according to manufacturer's protocols. 1 μ l of resulting cDNA was used as qPCR

template. Transcripts were amplified using primer sets targeting specific genes as described in Supplementary Table 1, using an ABI 7900 HT Real time PCR system and PerfeCTa SYBR Fastmix according to manufacturer's protocols. Transcript level was quantified in three technical replicates of each treatment using the $2^{-\Delta\Delta CT}$ method with expression normalized relative to *mnpB* as previously done by Gonzalez et al. 2019 (33), and results averaged per treatment.

Bacterial adenylate cyclase two-hybrid assays

The bacterial adenylate cyclase two-hybrid assay (BACTH) was employed to probe protein-protein interaction between HcyC (Npun_F1684) and HcyA (Npun_F1682) with the putative phosphorylated serine residues replaced with alanine, as well as sigma factors C, F, and J. To construct a plasmid encoding HcyC fused to the T25 fragment of *Bordetella pertussis* adenylate cyclase for bacterial adenylate cyclase two-hybrid (BACTH) analysis (42, 43), the coding region gene was amplified via PCR and cloned into pKT25 as an XbaI-KpnI fragment using restriction sites introduced on the primers. BTH101 (adenylate cyclase-deficient) *E. coli* strains transformed with appropriate plasmids were streaked onto Lysogeny Broth (LB) agar plates containing 100 µg/ml Ampicillin and 50 µg/ml Kanamycin, and incubated at 30°C for 48 h. For each strain, 1 ml of LB containing 100 µg/ml Ampicillin, 50 µg/ml Kanamycin and 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) was inoculated with several colonies from the plate, and cultures were then incubated overnight at 30°C with shaking at 230 rpm. Subsequently, 2 µl were taken from each overnight culture and spotted onto MacConkey

agar plates containing 100 $\mu\text{g/ml}$ Ampicillin and 50 $\mu\text{g/ml}$ Kanamycin. Plates were incubated at 30°C for 48 h prior to imaging.

Results

Patterns in RNAseq expression analysis

The original intent of the present study was to identify ancillary regulatory targets of the scytonemin TCR. To this end, RNA sequencing was conducted on *N. punctiforme* UCD 153 (wild type) and derived scytonemin TCR (scyTCR) system mutants Δ RR and Δ Reg, both under UV-A inductive conditions and in the absence of UV (non-induced). The expectations were that any genes under control of this TCR would show a relative downregulation in the mutants, and that the differences would be dependent on inductive (UV-A exposed) conditions. We found that out of the 6219 ORFs in the genome, 527 were statistically differentially expressed ($q < 10^{-15}$) between wild type and either one of the two regulatory mutants under scytonemin-inductive conditions (Fig. 1). Most of this response was seen in the double mutant rather than in the response regulator mutant, consistent with the crucial role of the histidine kinase in the photosensory response. Under non-inductive conditions the number of differentially expressed ORFs between wild type and mutants was reduced by more than half, but not completely abolished, indicating that the scytonemin TCR can also exert regulatory control based on non-UV-A sensory inputs. Unexpectedly, we found a large group of genes (207 under UV-A induction and 132 in its absence) that were more intensely expressed in Δ Reg than in the wild type. While most of these genes were annotated as hypothetical, one contiguous cluster (Npun_F0198 to Npun_F0237), included a heat shock protein HSP20, UV damage endonuclease, and manganese containing catalase among other general stress response genes (Supplementary Fig. 1). These results would imply that scyTCR can act

as an inhibitor, in addition to its known activator function, directly or indirectly, under conditions leading to scytonemin synthesis. Clearly the regulatory reach, mode of action, and sensory versatility of the scyTCR goes well beyond what was previously assumed. However, many of the presumptive target genes, while statistically significant, were either hypothetical proteins or annotated to an insufficient level to extract useful information.

As expected, expression of genes in the scytonemin operon (Npun_R1276-Npun_R1260) was virtually abolished in either one of the TCR deletion mutants under UV-A exposure (205-190 fold downregulation vs. the wild type), consistent with the results of Naurin et al 2016 (30). Further, no statistical difference in expression between wild type and TCR deletion mutants was detected in the absence of UV-A (Fig. 2), indicating the pure photosensory nature of its regulatory action on the scytonemin operon. Interestingly, the *ebo* gene cluster (Npun_R5231-Npun_F5236), which is essential for the synthesis of scytonemin as an intermediary export system to the periplasm (38), did not exhibit differential expression between wild type and scytonemin TCR mutants (data not shown), despite being demonstrably induced by UV-A exposure (22). Thus, the *ebo* operon in *Nostoc* is likely under the control of a separate photoregulatory system.

Three other functional gene clusters (Fig. 2) displayed conspicuous differential downregulation in the TCR mutants: one encoding gas vesicle proteins, another one, just downstream of it, coding for a variety of chemotaxis-related proteins (Npun_F2147 to Npun_F2164) and the *pilM*-Q cluster (Npun_F5005 to Npun_F5008) encoding type IV pilus system components (33). All of these clusters are typically and exclusively

expressed in *Nostoc* during the development of hormogonia. The level of wild type expression for these clusters under UV-A induction was 3-4 fold higher than that in the uninduced wild type cultures, showing that UV-A did play an inductive role in their expression. However, while the *scy* genes were downregulated very strongly (some 200-fold), these went down only by 5-6 fold with *scyTCR* deletion.

Finally, expression of a 4-gene cluster of putative, but unspecified, regulatory function (Npun_F1682-Npun_R1685) was also negatively affected by the deletion of the *scytonemin TCR* (2-3 fold; $q = 10^{-15}$) under conditions of UV-A induction. This effect was not detected in uninduced cultures ($q > 0.4$), indicating that the *scyTCR* has a UV-A dependent role in the promotion of their expression in the wild type, similar to the canonical effect on the *scytonemin* operon. Differential expression levels, however, were more in line with those seen in the hormogonia-related genes discussed above (Fig. 2). The cluster contains a gene predicted to be an anti-sigma factor antagonist (henceforth *hcyA*), followed by a RsbU-like phosphatase (henceforth *hcyB*), an anti-sigma factor (henceforth *hcyC*), and a multi-sensor hybrid histidine kinase (henceforth *hcyD*).

UV-A as an environmental cue for hormogonia development

The preceding results showing the positive regulation by *scyTRC* of expression of typical hormogonia genes suggest that UV-A may function as an environmental cue to their differentiation. But no such reports existed in the literature. To test this, we exposed cultures of the wild type to standard UV-A radiation fluxes in the dark (i.e., with no visible light). Abundant hormogonia developed within a few days (Fig. 3), whereas no

hormogonia developed in controls without UV. However, if as little as $5 \mu\text{mol m}^{-2} \text{s}^{-1}$ of visible light was included in the assays, no hormogonia developed.

Involvement of the *hcyA-D* cluster in UV-mediated hormogonia differentiation

The RNAseq results suggested that the scytonemin TCR may be responsible for a regulatory linkage between scytonemin production and hormogonia differentiation within the frame of a higher-order regulatory network, wherein the *hcyA-D* genes may be the key to this linkage. To probe for UV-A dependent hormogonia-related phenotypes associated with these genes, in frame deletion knockouts ($\Delta hcyA$ and $\Delta hcyC$) were constructed. Indeed, $\Delta hcyA$ displayed a decrease in hormogonia differentiation capacity relative to wild type *N. punctiforme* (Fig. 3) under UV induction assays, whereas $\Delta hcyC$ showed no such inhibition. When screened for scytonemin production in standard UV-A induction assays, both $\Delta hcyA$ and $\Delta hcyC$ were positive and showed no difference in accumulation compared to their wild type counterpart (Supplementary Fig. 4). These results are consistent with the notion that HcyC acts as an inhibitor of the hormogonium cascade so that its absence would not bring about a decrease in hormogonia differentiation, and that HcyA acts as an HcyC inhibitor whose absence would allow HcyC to inhibit the process more strongly. To test if these two putative regulatory elements could be part of a coordinated system, we conducted bacterial two-hybrid tests, which confirmed that HcyA and HcyC can physically interact *in vivo* (Supp Fig. 3), adding evidence beyond their genomic location that HcyC is a target of HcyA activity. As was seen in the wild type, addition of visible light at low intensity could suppress the UV-A induced development of hormogonia in the $\Delta hcyC$ mutant, but unlike in the wild

type, exposure to high fluxes of visible light restored the response. This restoration by high light was not detected in $\Delta hcyA$.

Surprisingly, two-hybrid interaction screens of HcyC against SigC, SigF, and SigJ (the sigma factors whose regulon controls hormogonia differentiation) did not yield any evidence of protein-protein interaction suggesting that the target of HcyC's activity is a different sigma factor or regulatory element all together (data not shown). Overall, the results implicate the hcyA-D regulatory system in the hormogonium differentiation regulatory cascade elicited by UV-A, rather than in the scytonemin regulatory system.

Suppression of scytonemin synthesis in hormogonia under UV-A

To examine regulatory interactions in the reverse direction (i.e., the hormogonium regulatory cascade acting on the scytonemin regulation) we conducted experiments with *N. punctiforme* ATCC 29133 and its derived sigma factor regulatory mutants in which hormogonia formation was induced with the standard phototactic assay, but with concurrent exposure to UV radiation. Cells were exposed to visible light laterally, and UV-A radiation from above. After five days, cultures were examined microscopically. We could determine that scytonemin had been produced in all wild type as well as $\Delta hcyA/C$ strains, and concurrently, hormogonia had formed in all strains (Fig.4). Deletion mutants did thus not interfere with scytonemin formation under phototactic induction, as they did not under UV induction (see above), and both responses proceeded concurrently. However, we noticed a clear populational segregation in the responses. In all cases, only a small part of the *Nostoc* filaments had differentiated into hormogonia, and those that did, did not exhibit the characteristic extracellular brownish coloration of scytonemin.

Conversely, scytonemin production occurred only in the subpopulation that remained in the vegetative state (Fig. 4). These findings confirm the notion that hormogonia differentiation and scytonemin production cannot proceed simultaneously in a single filament and indicate that a regulatory mechanism must be at play to suppress scytonemin production once a cell is committed to differentiation, even under concurrent UV-A irradiation. This mechanism cannot be reliant on the *hcy* system, given that the deletion mutants were unaffected in its ability to segregate the responses.

We re-examined the RNAseq time course conducted by Gonzalez et al 2019 (33) in which expression changes were followed during the process of phototactic hormogonium differentiation paying particular attention to changes in expression of scytonemin-related genes under standard phototactic hormogonia induction. Expectedly, since the cultures were not exposed to UV, neither scytonemin nor *ebo* operons displayed any significant changes in expression. However, the histidine sensor kinase of the scyTCR was down-regulated 7-fold during this 24-hour time course (Fig. 5), essentially blinding to UV-A any cells committed to develop into hormogonia. In a parallel analysis of the sigma factor J regulatory mutant, no such down-regulation of the scytonemin histidine kinase was observed over the same time course (Fig. 5) indicating that sigma factor J is, directly or indirectly, responsible for this suppression. The Gonzalez transcriptome also revealed expression of *hcyA* and *hcyC* to be dependent on sigma factor C (33). The expression of sigma factor C is itself directly dependent on sigma factor J (33), thus deletion of sigma factor J has the effect of abolishing phototactic induction of *hcyA-C* (Fig. 5)

Dynamics of UV-A mediated upregulation of scytonemin and hormogonia responses

Two reasons prompted us to inspect more closely the induction of the hormogonia response genes that we saw with the RNAseq data in Fig. 2 under conditions of scytonemin induction (UV-A in the presence of visible light). First scytonemin production in response to UV-A takes place over a period of 4-6 days (22) while the typical phototactic hormogonia differentiation takes only 24 hours (33). Second, that induction, although statistically significant, did demonstrably not lead to an actual, phenotypically detectable formation of hormogonia (Fig. 3). For this we conducted an RT-qPCR based time-course analysis of relevant genes. We chose hormogonia gas vesicle protein A (*gvpA*, Npun_F2159) and type IV pilus assembly protein (*pilM*, Npun_F5005) to represent the hormogonia response. The analyses were carried out with the wild type strains and the scyTCR and sigma factor regulatory mutants under standard scytonemin induction conditions over a period of 120 hours. The results (Fig. 6) corroborated the RNAseq data in that a clear induction by 5 days was seen in *pilM* and *gvpA* genes in the wild type, but not in Δ RR or Δ Reg. This induction was, however, dynamically much more sluggish and less intense than that seen under phototactic induction of hormogonia differentiation (compare with Fig. 5). The involvement of the *hcy* system in this induction was also patent in that Δ *hcyA* mutants failed to respond. Interestingly, Δ *hcyC* mutants had a short-lived peak of expression during the first day, while both wild type and all other mutants showed a relative expression decrease early on. These results are consistent with the presence of competing sensory signals and regulatory systems: one that can induce the hormogonia genes based on the UV-A

signals, and another one repressing these genes through other inputs (possibly visible light, see Fig. 3) in which HcyC participates, given that its absence allows their expression at least during the early response.

Discussion

RNAseq analyses of deletion mutants in the *scyTCR* of *Nostoc punctiforme* demonstrated that the suite of genes under its influence is much more diverse than previously considered, that it includes both up and downregulation of various genes and multigene loci, and that it likely also incorporates sensory inputs other than UV-A radiation. It apparently downregulates at least some genes involved in photooxidative stress responses alternative to scytonemin production. In its regulon, the most puzzling targets were sets of genes presumably involved in the development of hormogonia, the meaning of which we pursued in this contribution.

Bet-hedging in response to UV-A exposure

Our experiments demonstrate that the segregation between motility and sunscreen responses to UVA (i.e., the choice between hunkering down and passively enduring its stress vs. actively fleeing), both of which were known from different species of cyanobacteria (44), have an intra-population parallel in a single species genetically capable of both types of responses. Our work demonstrated that UV-A cues in themselves suffice to elicit hormogonia differentiation in *Nostoc*, concurrently with the well-known synthesis of scytonemin (22). However, the responses were mutually exclusive at the

cellular level, as *Nostoc* hormogonia failed to produce scytonemin even in the presence of UV-A radiation levels that elicited its synthesis in cells of the same original population that had remained in the vegetative state.

In terms of populational fitness it may well make sense for *Nostoc* to hedge its bets by deploying both types of responses concurrently. This may ensure a higher rate of survival under ultraviolet stress of unknown magnitude, as sunscreens may or may not provide sufficient protection under harsh regimes, and hormogonia are not guaranteed a safe journey to more benign settings. This strategy of risk spreading in the face of fluctuating environmental conditions is known as bet hedging (45). Maintaining genetically encoded phenotypic heterogeneity may be of particular utility in the face of unpredictably fluctuating environmental conditions (3). The crux is how to manage regulatory stress responses for both types of adaptations in a coordinated manner. In other words, how to obtain the desired heterogeneity in response first, and then how to silence the alternative pathway once a commitment is made at the cellular level. The first part of the problem is not very different from the choices faced by *Nostoc* populations under stressors other than UV-A, in that differentiation of hormogonia is apparently generally not universal, except under controlled laboratory conditions (46), and only portions of the natural population will commit to the developmental pathway (47). Many instances of phenotypic response heterogeneity have been observed across bacteria (3), suggesting that a diversity of stress responses in a clonal population can increase overall fitness in the face of fluctuating environmental stressors (48). The presence of double-negative, imbricated regulatory systems (4, 49, 50), as well as transcriptional noise and

stochasticity (59, 71) are thought to play a key role in the generation of phenotypic response heterogeneity. A conspicuous case of phenotypic heterogeneity occurs in cyanobacteria during nitrogen fixation, wherein a small subset of cells in an isogenic population modulate expression of nitrogenase components based on a circadian clock (51, 52). The basic mechanism that allows phenotypic heterogeneity to arise under UV-A may well be intrinsic to the core hormogonia regulatory cascade (Fig. 7), although the specifics are neither known nor easily derived from current knowledge. Our research provides some mechanistic insights into a means for coordination and cross-talk between the two canonical regulatory pathways involved.

How and why hormogonia suppress scytonemin synthesis

The regulatory basis of this phenomenon can be traced to the downregulation of the scyTCR expression by sigma J, as demonstrated by the Δ sigJ knockouts during phototactic hormogonium differentiation (Fig. 4). The extent of this downregulation in committed hormogonia is in fact not fully encompassed by the populational RNAseq data, in that a 6-7 fold downregulation signal stemming from a just a portion of cells (committed hormogonia under the present UV-A inductive conditions) that may make up at most some 10% of the population, implies that the single-cell downregulation effect in hormogonia themselves could be close to an order of magnitude higher, although the precise experimental conditions used may influence this proportion, as standard phototactic assays yield near 100% differentiation of hormogonia. In any event, this downregulation of the scyTCR likely denies hormogonia the ability to respond to UV-A cues that would otherwise strongly upregulate scytonemin genes. Although the specific

mechanism of this transcriptional repression cannot be directly discerned from our data, it is unlikely that sigma J acts directly as a repressor of the scyTCR transcription because sigma factors promote expression by recruiting RNA polymerase to their target transcriptional start site (53, 54). Sigma J may in turn upregulate the expression of a putative secondary transcriptional repressor that targets the scytonemin TCR locus. This regulatory separation makes sense in terms of energetics of investment, in that any scytonemin formed by hormogonia would be exported to the extracellular EPS, and thus be left behind as they glide away from it, negating any potential beneficial effects to the hormogonium. The investment in an ineffective sunscreen would add to the already burdened metabolism of hormogonia differentiation, which involves the cessation of vegetative metabolic activities in exchange for overproduction of hormogonia specific proteins (55).

UV-A can elicit hormogonia differentiation and this response is modulated by visible light

Our results point to the *hcyA-D* quadripartite regulatory system as a crucial link between the canonical scytonemin and hormogonia regulatory cascades (Fig. 7). On the one hand, as a unit, its expression was enhanced by the UV-A sensing scyTCR (Fig. 2), and mutants of one of its components, *hcyA*, failed to develop hormogonia under induction by UV-A (Fig. 3), both linking its expression to the scytonemin circuit. On the other hand, it is overexpressed during phototactic hormogonia differentiation in a process dependent on the master regulator of the hormogonia cascade, sigma J (Fig 6) and, additionally some of its expression of some of its genes is controlled by members of the tripartite

hierarchical sigma factor cascade (sigma factors J,F, and C) (33). Notably, we found no differential expression or phenotypic effects of the hcy deletion mutants on the scytonemin systems, indicating that its main ultimate regulatory target is hormogonia differentiation. These results clearly link the system to the canonical hormogonia regulatory circuitry. Interestingly, the hcyA-D cluster is conserved in sequence homology of its components and genomic architecture only among a small group of cyanobacterial genomes corresponding to genera that are capable of differentiating hormogonia and producing scytonemin (*Nostoc*, *Calothrix*, *Tolypothrix*), but not in other cyanobacteria, which speaks for adaptations to UV beyond *Nostoc* itself.

Further, this system is also at least partly responsible for the integration of visible light signals that can demonstrably counteract the induction of the hormogonium circuit by UVA. HcyC seems to play a role in this signal integration as its deletion mutant was unable to carry out fully the suppression of UV-A-induction mediated by (high intensity) visible light (Fig. 3) and displays a differential early burst of hormogonia gene expression (Fig. 6).

While the precise mechanisms of the hcyA-D mediated interaction cannot be directly determined from our data, bioinformatic examination may provide some insights. The system contains four genes, three of which are annotated as potential components of a partner switching regulatory unit, including a putative anti-sigma factor (*hcyC*), an anti-sigma factor antagonist (*hcyA*) and a phosphatase (*hcyB*), and we demonstrate (Supplementary Fig. 3) that *hcyC* and *hcyA* interact physically *in vivo*. The multi-sensor hybrid histidine kinase encoded by *hcyD* (Npun_R1685) contains two GAF and one PAS

sensory domains according to the NCBI Conserved Protein Domain Database (56–58). Both are known as chromophore binding domains (59) suggesting that it may play a photosensory role (60), and its histidine kinase domain may activate other components of the hcy system. By analogy to similar partner switching systems, a potential candidate is the phosphatase HcyB, with homology to RsbU (61), which is known in *Bacillus subtilis* to be activated by a sensor kinase, and in turn dephosphorylates an anti-sigma factor antagonist (61). By analogy, HcyB could dephosphorylate HcyA, relieving its inhibition of HcyC. Similarly, one might be inclined to postulate the hcyA-D system exerts its control over the expression of hormogonia development through post-translational interactions between HcyC and Sigma J. However, it is clear from two-hybrid analyses that HcyC does not directly interact with any of the three hormogonia related sigma factors *in vivo*. If HcyC were to actually act as an anti-sigma factor, the target would have to be found outside of the canonical tripartite cascade. However, other studies have found that similar partner-switching systems, such as HmpWUV in *N. punctiforme*, can have final output effects on non-sigma factor targets (62–64). This could well be the case with hcyA-D given that a lack of hcyA led to a complete loss of UV-A induced hormogonia formation as shown in Figure 4, suggesting that HcyC is inhibiting some regulatory element that positively affects hormogonia development. In the Δ hcyC mutant background, HcyA has no target to inhibit, and thus hormogonia differentiation proceeds unabated.

It is also possible that the apparent complexity in regulatory circuitry may contribute to the process of populational bifurcation in the phenotypic responses to UV-A, since

negative feedback loops similar to the regulatory mechanism that mediates it, directly lead to cases of phenotypic heterogeneity (4, 49, 50). In the present working model (Fig. 7), sigma factor J indirectly and negatively regulates the scytonemin TCR and indirectly and positively regulates expression of the *hcyA-D* system, which in turn negatively regulates, directly or indirectly, hormogonia responses.

Outlook

Clearly, further experimentation will be needed to characterize the components, precise interactions and phenotypic outcomes of this complex gene regulatory network to the point of gaining a sound mechanistic understanding. Our current interpretation was unexpectedly hampered by the discovery that strain UCD 153, which was used to derive the scyTCR mutants, was in fact deficient in hormogonia differentiation, which prevented us from further probing effects of its mutants on the phenotype directly. The interrogation of new mutant strains based on the ATCC 29133 strain, both targeting the scyTCR and other components shown by our study to be important in the *hcy* system, particularly *hcyD*, should prove useful in this respect. Additionally, further probing of environmental cues in various combinations may help clarify the precise conditions that bring about an increase in fitness to *Nostoc*, and perhaps also other similar genera of cyanobacteria, by fleeing UV-A exposure via developmental processes, in that currently this outcome occurs under exposure to UV-A alone (with no exposure to visible light), a situation that does not happen in nature. In this regard, it is useful to be reminded of the limitations of using model laboratory strains kept for decades under laboratory conditions to interpret situations in their original habitat, in that they are released from the selective pressures

that honed adaptations of fitness value there, as the case of strain UCD 153 most poignantly reminds us.

References

1. Gao R, Stock AM. 2009. Biological insights from structures of two-component proteins. *Annu Rev Microbiol* 63:133–154.
2. Magasanik B. 1961. Catabolite repression. *Cold Spring Harb Symp Quant Biol* 26:249–256.
3. Ackermann M. 2015. A functional perspective on phenotypic heterogeneity in microorganisms. *Nat Rev Microbiol* 13:497–508.
4. Smits WK, Kuipers OP, Veening JW. 2006. Phenotypic variation in bacteria: The role of feedback regulation. *Nat Rev Microbiol* 4:259–271.
5. Beaumont HJE, Gallie J, Kost C, Ferguson GC, Rainey PB. 2009. Experimental evolution of bet hedging. *Nature* 462:90–93.
6. Levy SF, Ziv N, Siegal ML. 2012. Bet Hedging in Yeast by Heterogeneous, Age-Correlated Expression of a Stress Protectant. *PLoS Biol* 10:1001325.
7. Holland SL, Reader T, Dyer PS, Avery S V. 2014. Phenotypic heterogeneity is a selected trait in natural yeast populations subject to environmental stress. *Environ Microbiol* 16:1729–1740.
8. Avery S V. 2006. Microbial cell individuality and the underlying sources of heterogeneity. *Nat Rev Microbiol*. Nature Publishing Group.
9. Raj A, van Oudenaarden A. 2008. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell* 135:216–226.
10. Whitton BA, Potts M. 2012. Introduction to the cyanobacteria. *Ecol Cyanobacteria II Their Divers Sp Time*. Springer Netherlands.
11. van Gemerden H. 1993. Microbial mats: A joint venture. *Mar Geol* 113:3–25.
12. Revsbech NP, Jørgensen BB. 1983. Photosynthesis of benthic microflora measured with high spatial resolution by the oxygen microprofile method: Capabilities and limitations of the method. *Limnol Oceanogr* 28:749–756.
13. Potts M. 1994. Desiccation tolerance of prokaryotes. *Microbiol Rev* 58:755–805.
14. Singh SP, Häder D-P, Sinha RP. Cyanobacteria and ultraviolet radiation (UVR) stress: Mitigation strategies. *Ageing Res Rev* 9:79–90.
15. Rastogi RP, Sinha RP, Moh SH, Lee TK, Kottuparambil S, Kim Y-J, Rhee J-S, Choi E-M, Brown MT, Häder D-P, Han T. 2014. Ultraviolet radiation and

- cyanobacteria. *J Photochem Photobiol B Biol* 154–169.
16. Garcia-Pichel F, Castenholz RW. 1991. Characterization and Biological Implications of Scytonemin, a Cyanobacterial Sheath Pigment. *J Phycol* 27:395–409.
 17. Gao Q, Garcia-Pichel F. 2011. Microbial ultraviolet sunscreens. *Nat Rev Microbiol* 9:791–802.
 18. Soule T, Stout V, Swingley WD, Meeks JC, Garcia-Pichel F. 2007. Molecular genetics and genomic analysis of scytonemin biosynthesis in *Nostoc punctiforme* ATCC 29133. *J Bacteriol* 189:4465–72.
 19. Nadeau T, Howard-Williams C, Castenholz R. 1999. Effects of solar UV and visible irradiance on photosynthesis and vertical migration of *Oscillatoria* sp. (Cyanobacteria) in an Antarctic microbial mat. *Aquat Microb Ecol* 20:231–243.
 20. Bebout BM, Garcia-Pichel F. 1995. UV B-Induced Vertical Migrations of Cyanobacteria in a Microbial Mat. *Appl Environ Microbiol* 61:4215–4222.
 21. Meeks J, Campbell E, Summers M, Wong F. 2002. Cellular differentiation in the cyanobacterium *Nostoc punctiforme*. *Arch Microbiol* 178:395–403.
 22. Soule T, Garcia-Pichel F, Stout V. 2009. Gene expression patterns associated with the biosynthesis of the sunscreen scytonemin in *Nostoc punctiforme* ATCC 29133 in response to UVA radiation. *J Bacteriol* 191:4639–4646.
 23. Risser DD, Chew WG, Meeks JC. 2014. Genetic characterization of the hmp locus, a chemotaxis-like gene cluster that regulates hormogonium development and motility in *Nostoc punctiforme*. *Mol Microbiol* 92:222–233.
 24. Dodds WK, Gudder DA, Mollenhauer D. 1995. The Ecology of *Nostoc*. *J Phycol* 31:2–18.
 25. Meeks JC, Elhai J, Thiel T, Potts M, Larimer F, Lamerdin J, Predki P, Atlas R. 2001. An overview of the Genome of *Nostoc punctiforme*, a multicellular, symbiotic Cyanobacterium. *Photosynth Res* 70:85–106.
 26. Campbell EL, Christman H, Meeks JC. 2008. DNA microarray comparisons of plant factor and nitrogen deprivation-induced hormogonia reveal decision-making transcriptional regulation patterns in *Nostoc punctiforme*. *J Bacteriol* 190:7382–7391.
 27. Ekman M, Picossi S, Campbell EL, Meeks JC, Flores E. A *Nostoc punctiforme* Sugar Transporter Necessary to Establish a Cyanobacterium-Plant Symbiosis. *Plant Physiol* 161:1984–1992.
 28. Rajeev L, Luning EG, Dehal PS, Price MN, Arkin AP, Mukhopadhyay A. 2011. Systematic mapping of two component response regulators to gene targets in a model sulfate reducing bacterium. *Genome Biol* 12:R99.

29. Soule T, Palmer K, Gao Q, Potrafka R, Stout V, Garcia-Pichel F. 2009. A comparative genomics approach to understanding the biosynthesis of the sunscreen scytonemin in cyanobacteria. *BMC Genomics* 10:336.
30. Naurin S, Bennett J, Videau P, Philmus B, Soule T. 2016. The response regulator Npun_F1278 is essential for scytonemin biosynthesis in the cyanobacterium *Nostoc punctiforme* ATCC 29133. *J Phycol* 52:564–571.
31. Khayatan B, Meeks JC, Risser DD. 2015. Evidence that a modified type IV pilus-like system powers gliding motility and polysaccharide secretion in filamentous cyanobacteria. *Mol Microbiol* 98:1021–1036.
32. Armstrong RE, Hayes PK, Walsby AE. 1983. Gas Vacuole Formation in Hormogonia of *Nostoc muscorum*. *Microbiology* 129:263–270.
33. Gonzalez A, Riley KW, Harwood T V, Zuniga EG, Risser DD, Ellermeier CD. 2019. A Tripartite, Hierarchical Sigma Factor Cascade Promotes Hormogonium Development in the Filamentous Cyanobacterium *Nostoc punctiforme* 4:e00231-19.
34. Davis MC, Kesthely CA, Franklin EA, MacLellan SR. 2017. The essential activities of the bacterial sigma factor. *Can J Microbiol* 63:89–99.
35. Soule T, Gao Q, Stout V, Garcia-pichel F. 2013. The Global Response of *Nostoc punctiforme* ATCC 29133 to UVA Stress , Assessed in a Temporal DNA Microarray Study. *Photochem Photobiol* 89:415–423.
36. Cai Y, Wolk CP. 1990. Use of a Conditionally Lethal Gene in *Anabaena* sp. Strain PCC 7120 To Select for Double Recombinants and To Entrap Insertion Sequences. *J Bacteriol* 172:3138–3145.
37. Cohen MF, Wallis JG, Campbell EL, Meeks JC. Transposon mutagenesis of *Nostoc* sp. strain ATCC 29133, a filamentous cyanobacterium with multiple cellular differentiation alternatives. *Microbiology* 140:3233–3240.
38. Klicki K, Ferreira D, Hamill D, Dirks B, Mitchell N, Garcia-Pichel F. 2018. The widely conserved ebo cluster is involved in precursor transport to the periplasm during scytonemin synthesis in *Nostoc punctiforme*. *MBio* 9:e02266-18.
39. Campbell EL, Summers ML, Christman H, Martin ME, Meeks JC. 2007. Global gene expression patterns of *Nostoc punctiforme* in steady-state dinitrogen-grown heterocyst-containing cultures and at single time points during the differentiation of akinetes and hormogonia. *J Bacteriol* 189:5247–56.
40. Tjaden B. 2015. De novo assembly of bacterial transcriptomes from RNA-seq data. *Genome Biol* 16:1.
41. Kolde R, Antoine L. 2015. pheatmap: Pretty Heatmaps. <https://github.com/raivokolde/pheatmap>.

42. Karimova G, Pidoux J, Ullmann A, Ladant D. 1998. A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci U S A* 95:5752–5756.
43. Battesti A, Bouveret E. 2012. The bacterial two-hybrid system based on adenylate cyclase reconstitution in *Escherichia coli*. *Methods* 58:325–334.
44. Castenholz RW, Garcia-Pichel F. 2007. Cyanobacterial Responses to UV-Radiation. *Ecol Cyanobacteria* 21:591–611.
45. De Jong IG, Haccou P, Kuipers OP. 2011. Bet hedging or not? A guide to proper classification of microbial survival strategies. *BioEssays* 33:215–223.
46. Splitt SD, Risser DD. 2016. The non-metabolizable sucrose analog sucralose is a potent inhibitor of hormogonium differentiation in the filamentous cyanobacterium *Nostoc punctiforme*. *Arch Microbiol* 198:137–147.
47. Campbell EL, Christman H, Meeks JC. 2008. DNA microarray comparisons of plant factor- and nitrogen deprivation-induced Hormogonia reveal decision-making transcriptional regulation patterns in *Nostoc punctiforme*. *J Bacteriol* 190:7382–91.
48. Holland SL, Reader T, Dyer PS, Avery S V. 2014. Phenotypic heterogeneity is a selected trait in natural yeast populations subject to environmental stress. *Environ Microbiol* 16:1729–1740.
49. Schurr MJ, Martin DW, Mudd MH, Hibler NS, Boucher JC, Deretic V. 1993. The *algD* promoter: Regulation of alginate production by *Pseudomonas aeruginosa* in cystic fibrosis. *Cell Mol Biol Res* 39:371–376.
50. McCaw ML, Lykken GL, Singh PK, Yahr TL. 2002. ExsD is a negative regulator of the *Pseudomonas aeruginosa* type III secretion regulon. *Mol Microbiol* 46:1123–1133.
51. Mohr W, Vagner T, Kuypers MMM, Ackermann M, LaRoche J. 2013. Resolution of Conflicting Signals at the Single-Cell Level in the Regulation of Cyanobacterial Photosynthesis and Nitrogen Fixation. *PLoS One* 8:1–7.
52. Adams DG. 2000. Heterocyst formation in cyanobacteria. *Curr Opin Microbiol* 3:618–624.
53. Wösten MMSM. 1998. Eubacterial sigma-factors. *FEMS Microbiol Rev* 22:127–150.
54. Helmann JD, Chamberlin MJ. 1988. Structure and Function of Bacterial Sigma Factors. *Annu Rev Biochem* 57:839–872.
55. Marsac NT. 1994. Differentiation of Hormogonia and Relationships with Other Biological Processes, p. 825–842. *In The Molecular Biology of Cyanobacteria*. Springer Netherlands.

56. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko M V., Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH. 2011. CDD: A Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* 39:D225–D229.
57. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Res* 43:D222–D226.
58. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Geer LY, Bryant SH. 2017. CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 45:D200–D203.
59. Wagner JR, Brunzelle JS, Forest KT, Vierstra RD. 2005. A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* 438:325–331.
60. Wiltbank LB, Kehoe DM. 2019. Diverse light responses of cyanobacteria mediated by phytochrome superfamily photoreceptors. *Nat Rev Microbiol* 17:37–50.
61. Delumeau O, Dutta S, Brigulla M, Kuhnke G, Hardwick SW, Völker U, Yudkin MD, Lewis RJ. 2004. Functional and structural characterization of RsbU, a stress signaling protein phosphatase 2C. *J Biol Chem* 279:40927–40937.
62. Riley KW, Gonzalez A, Risser DD. 2018. A partner-switching regulatory system controls hormogonium development in the filamentous cyanobacterium *Nostoc punctiforme*. *Mol Microbiol* 109:555–569.
63. Morris AR, Visick KL. 2013. The response regulator SypE controls biofilm formation and colonization through phosphorylation of the *syp*-encoded regulator SypA in *Vibrio fischeri*. *Mol Microbiol* 87:509–525.
64. Mercer RG, Lang AS. 2014. Identification of a predicted partner-switching system that affects production of the gene transfer agent RcGTA and stationary phase viability in *Rhodobacter capsulatus*. *BMC Microbiol* 14:1–16.
65. Campbell EL, Summers ML, Christman H, Martin ME, Meeks JC. 2007. Global gene expression patterns of *Nostoc punctiforme* in steady-state dinitrogen-grown heterocyst-containing cultures and at single time points during the differentiation of akinetes and hormogonia. *J Bacteriol* 189:5247–56.
66. Ferreira D, Garcia-Pichel F. 2016. Mutational Studies of Putative Biosynthetic Genes for the Cyanobacterial Sunscreen Scytonemin in *Nostoc punctiforme* ATCC

29133. *Front Microbiol* 7:1–10.

67. Bertani G. 1959. Sensitivities of Different Bacteriophage Species to Ionizing Radiation. *J Bacteriol* 79:387–393.
68. Riley KW, Gonzalez A, Risser DD. 2018. A partner-switching regulatory system controls hormogonium development in the filamentous cyanobacterium *Nostoc punctiforme*. *Mol Microbiol* 109:555–569.

Tables and Figures

Table 1. *Nostoc punctiforme* strains utilized in this work.

Strain	Description	Mutation Locus Tag	Function	Mutant Phenotype
ATCC 29133	Wild Type/ WT	-	-	-
UCD153	Strain derived from ATCC 29133	-	-	High rate of conjugal transfer, no hormogonia differentiation
Δ RR	Deletion of scytonemin TCR response regulator	Npun_F1278	DNA-binding transcription factor initiating transcription of scytonemin operon	Total abolition of scytonemin production
Δ Reg	Deletion of scytonemin response regulator and histidine kinase	Npun_F1278, Npun_F1277	Two component regulatory system controlling expression of the scytonemin operon	Total abolition of scytonemin production
Δ hcyA	Deletion of anti-sigma factor antagonist	Npun_F1682	Inhibitor of anti-sigma factor	Mild reduction in hormogonia differentiation under phototactic induction
Δ hcyC	Deletion of anti-sigma factor	Npun_F1684	Inhibitor of sigma factor	No reduction in hormogonia differentiation under phototactic induction
Δ σ J	Deletion of sigma factor J	Npun_R1337	Regulator of phototactic hormogonia response	Total abolition of hormogonia response

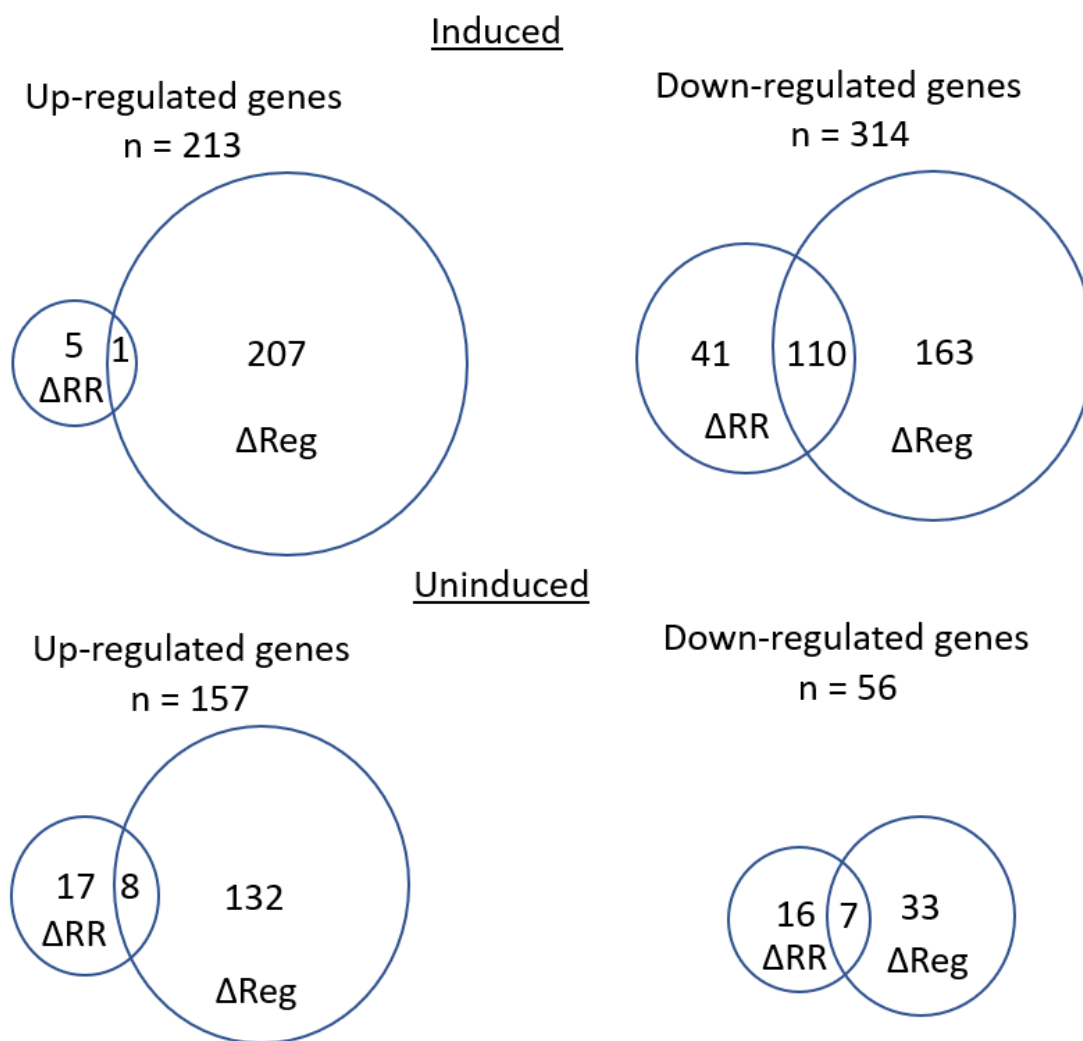


Fig. 1: Numerical overview of differentially regulated genes between *N. punctiforme* UCD 153 (wild type) and scytonemin two-component regulatory system deletion mutants (response regulator, ΔRR ; response regulator and histidine kinase double mutants, Δreg) either under five days of white light alone incubation (non-induced) or white light and UV-A irradiation (induced). Venn diagrams not to scale.

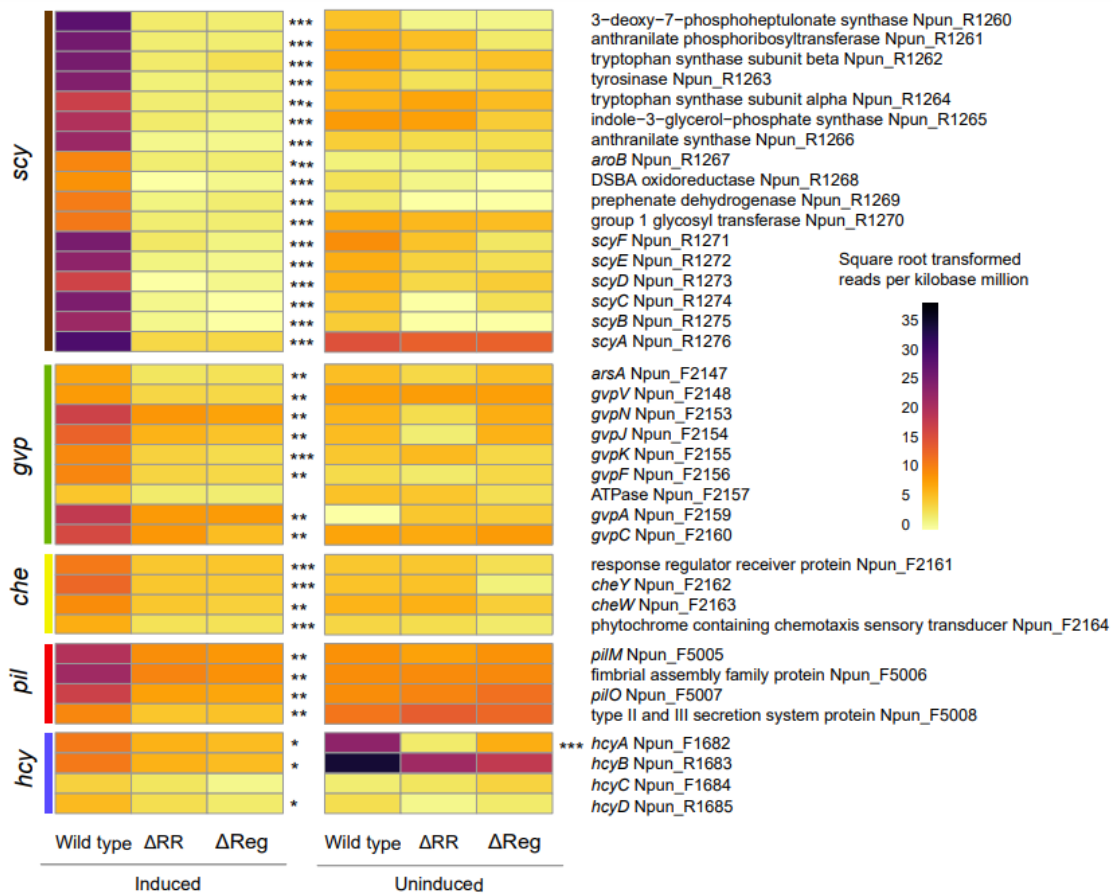


Fig. 2: Subset of differentially expressed genes from RNA sequencing of *N. punctiforme* wild type (UCD 153), ΔRR, and ΔReg under UV-A and standard white light conditions. Square root transformed RPKM for each gene was used to produce heatmap. The scytonemin operon (Npun_R1276-Npun_R1260) was strongly downregulated in mutants lacking the scytonemin TCR. A cluster of genes from Npun_F2147 to Npun_F2164 implicated in gas vesicle synthesis and other motility related processes were also downregulated in scytonemin TCR mutants under inductive conditions, as were genes Npun_F5005-Npun_F5008, involved in the synthesis of type-IV pilus components. Several components of the anti-sigma factor/anti-sigma factor partner switching system Npun_1682- Npun_1685 (*hcyA-D*) were also differentially expressed between wild type and scytonemin TCR mutants. Significance thresholds between wild type and mutant strains: * ($q \leq 10^{-15}$), ** ($q \leq 10^{-20}$), *** ($q \leq 10^{-50}$).

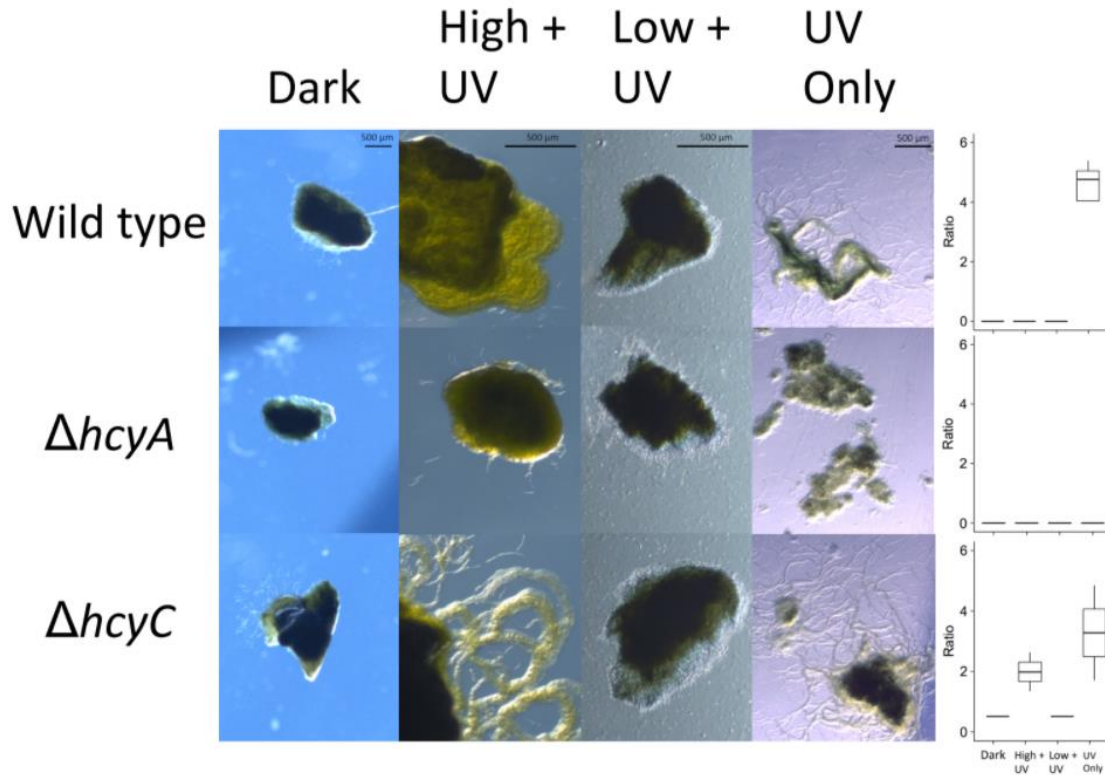


Fig. 3: Phenotypic responses of wild type, $\Delta hcyA$, $\Delta hcyC$ after 5 days in dark, high light ($100 \mu\text{mol m}^{-2} \text{s}^{-1}$) + UV, low light ($10 \mu\text{mol m}^{-2} \text{s}^{-1}$) + UV, and dark + UV conditions. Hormogonia can be seen spreading from the central clump of vegetative cells in wild type and $\Delta hcyC$ in UV-only conditions, but not in $\Delta hcyA$, while they can be seen emanating only from $\Delta hcyC$ under high light + UV conditions. Box plots represent hormogonium spreading quantification for each mutant and condition (see Materials and Methods) of 3 independent colonies. The experiment was repeated three times independently with qualitatively identical results.

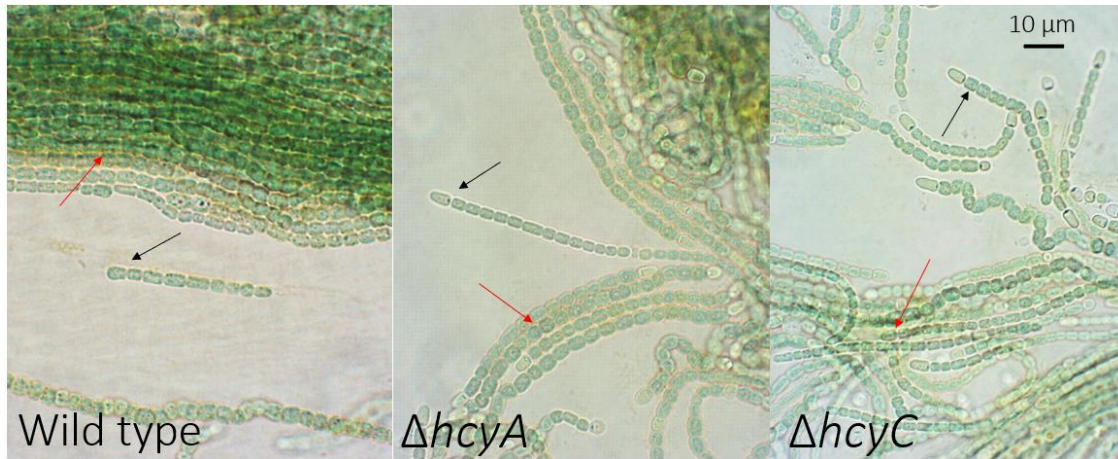


Fig. 4: Co-induction of scytonemin production and hormogonia differentiation in wild type, $\Delta hcyA$, and $\Delta hcyC$ mutants. Scytonemin can be seen as a brownish yellow sheath accumulating around vegetative cells, but not around hormogonia in all three strains. Black arrows point to hormogonium filaments and red arrows point to extracellular scytonemin accumulation in each frame. Quantitative assessment of scytonemin production is in Supplementary Fig 4

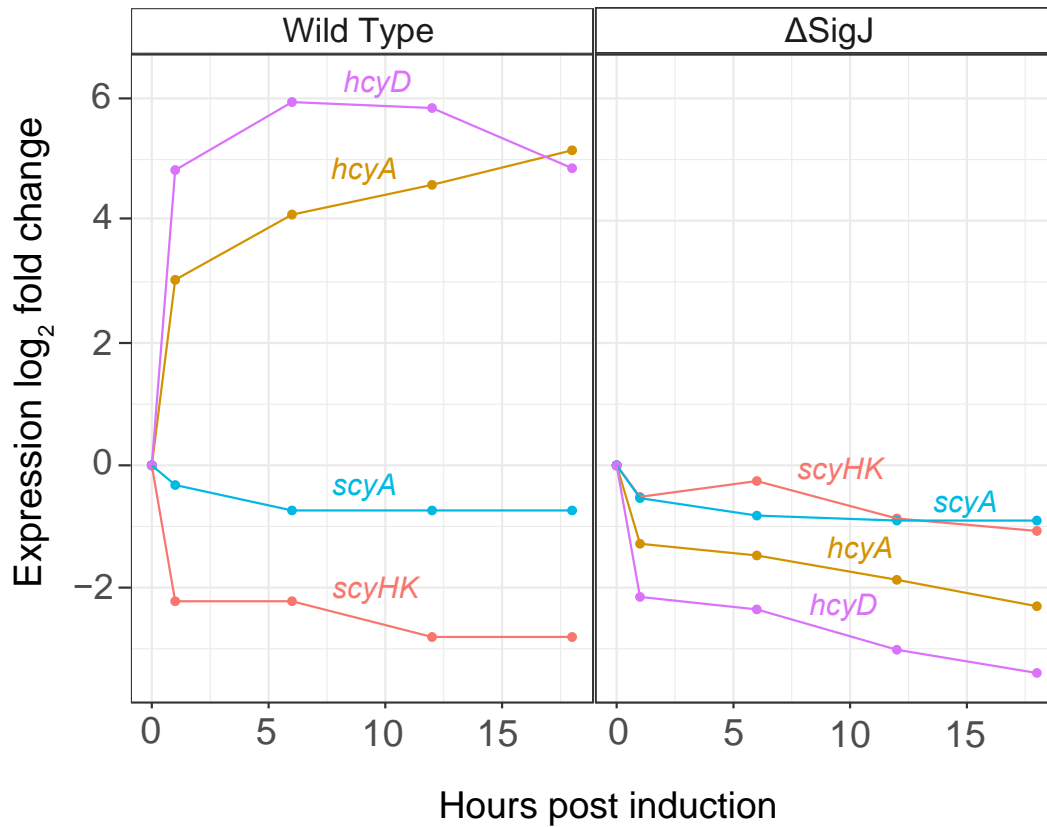


Fig. 5: Expression of scytonemin biosynthetic (*scyA*) and regulatory (*scyHK*) genes, as well as *hcyA* and *hcyD*, over a phototactic hormogonium induction time course (raw data from Gonzalez et al 2019) for the wild type and $\Delta sigJ$ deletion mutant. Expression levels are relative to those in wild type at time 0. The scyt TCR histidine kinase is downregulated during phototactic hormogonium induction in the wild type, but not in $\Delta sigJ$. Genes in the *hcyA-D* system show the opposite effect.

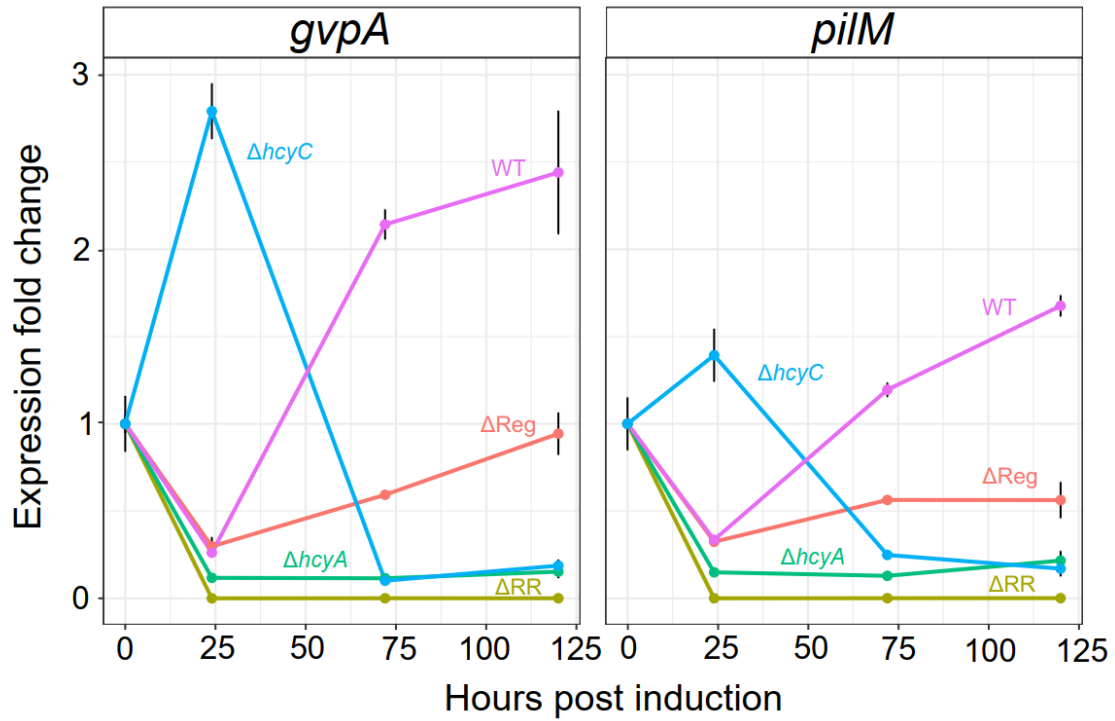


Fig. 6: Expression changes relative to time 0 of hormogonia-associated genes *gvpA* and *pilM* in wild type (WT), ΔRR , ΔReg , $\Delta hcyA$, and $\Delta hcyC$ mutants under standard scytonemin inductive conditions (UV-A plus white light) over 120 hours. Both genes showed increased expression over the time course in the wild type, but no such induction in ΔRR , ΔReg , or $\Delta hcyA$. $\Delta hcyC$ showed a burst of early up-regulation but decreased to low levels after the first day.

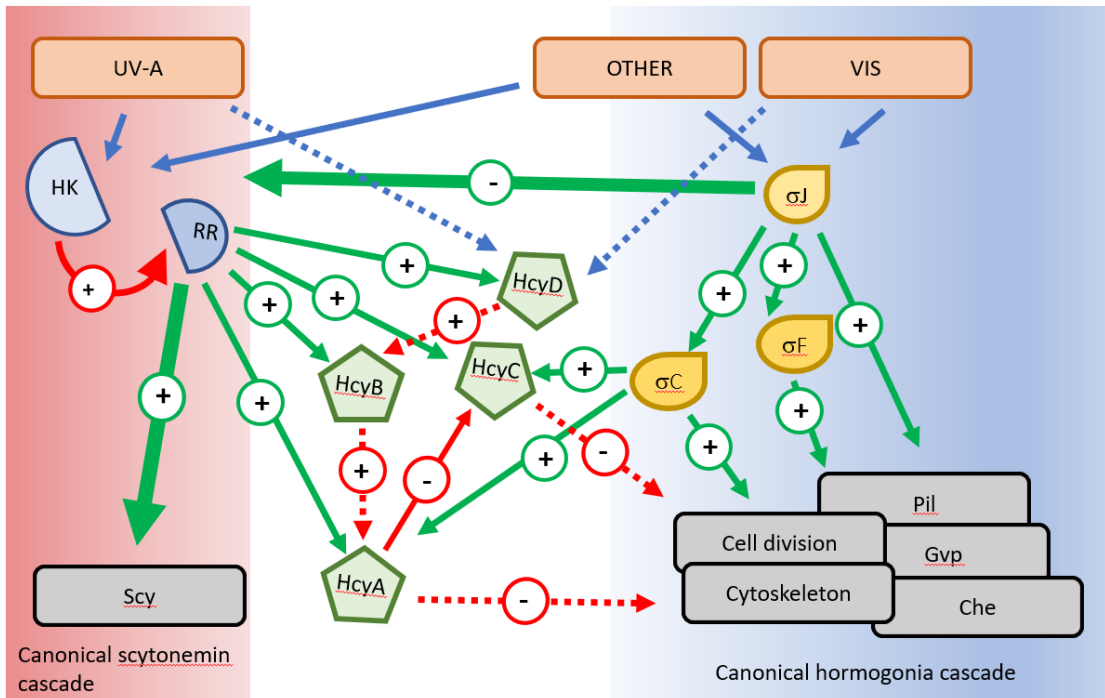


Fig. 7: Working model of the scytonemin-hormogonium gene regulatory network. The extent of the canonical networks for each is shown over colored background. Cross-talk of the canonical systems is enabled by the one-way repression by sigma factor J over the scytonemin TCR on the one hand, and by the hcyA-D partner switching system on the other, which receives transcriptional controls from both canonical networks but exerts control over the hormogonia genes only. Orange rectangles and blue arrows represent input signals. Grey boxes represent functional outputs. Green arrows denote transcriptional control, and red lines represent post-translational control. Solid lines represent demonstrated interactions, and dotted lines represent suspected or deduced interactions. Positive and negative signs refer to the respective type of interaction.

5 - Conclusions

This dissertation investigates the latter stages of the biosynthesis, as well as the genetic regulation of the cyanobacterial sunscreen scytonemin, and in doing so represents the first description of a biological role of the widely conserved *ebo* gene cluster and the induction of hormogonia differentiation by UVR. These findings prompted the development of the FLANDERS pipeline in the service of gleaning further insights into the distribution and broader role of the *ebo* gene cluster.

In chapter 2, UV irradiated knock-out mutants of the scytonemin synthesis genes *scyD*, *scyE*, and *scyF* as well as *ebo* gene cluster knock-out mutants were analyzed by high pressure liquid chromatography, revealing the presence of a novel compound in mutant strains that was absent in the wild-type. This compound was analyzed by spectrofluorometry and mass spectroscopy and found to be consistent with the scytonemin monomer (1). Because the scytonemin monomer was found to be autofluorescent, the intracellular localization of the scytonemin monomer was assessed by fluorescence microscopy. The localization differed between the *scyE* mutant and all 5 *ebo* mutants: in the absence of any one of the *ebo* genes, the scytonemin monomer remained relegated to the cytoplasm, while the *scyE* mutant saw the scytonemin monomer accumulate in the periplasm. These results, along with presence of a sec-signal peptide sequence in *scyE* are consistent with periplasmic localization of the final dimerization during scytonemin synthesis (2). This transport-like role for the *ebo* gene cluster is the first description of its role in any organism, and begs the question if this role is common in the hundreds of *ebo* harboring bacteria (3). In the case of many scytonemin

producing cyanobacteria (excluding *N. punctiforme* ATCC 29133) as well as the *edb* cluster of select soil *Pseudomonas* (4), the *ebo* gene cluster is contiguous with the biosynthetic cluster which it interacts with. The arrangement of the *ebo* gene in these cases suggested that colocalization with their cognate biosynthetic gene cluster may be a common occurrence, and thus examining the *ebo* adjacent regions of *ebo* containing genomes may reveal the biosynthetic gene clusters with which they are associated. In order to assess the veracity of this hypothesis, the fourth chapter examined *ebo* cluster neighboring regions of 92 genomes were searched for known biosynthetic clusters and other sequence based attributes that may suggest the presence of unknown biosynthetic clusters through development and implementation of the Functional Landscape and Neighbor Determining gEnomic Region Search (FLANDERS) pipeline. While no known biosynthetic gene clusters were found, many *ebo* adjacent genes contained signal peptides, much like *scyE* and several *edb* genes. Not only did these findings tacitly corroborate the compartmentalized biosynthesis of scytonemin outlined in chapter 2, they also demonstrated the broader applicability of the FLANDERS pipeline to answering specific question about contiguous gene clusters and their relationships.

Along with elucidating the latter stages of scytonemin synthesis, the present dissertation also seeks to characterize the regulatory network controlled by the scytonemin two-component regulatory system in the third chapter. Towards this end, RNA sequencing of a knock-out mutant of the scytonemin response regulator and a knock-out mutant of the scytonemin response regulator and histidine kinase was conducted under standard and scytonemin inducing conditions. In addition to downregulation of the scytonemin

operon, scytonemin two component regulatory system mutants also showed a downregulation of a number of hormogonia related genes. These results prompted experiments that demonstrated that UV exposure could elicit hormogonia differentiation in a subset of cells in a population of *Nostoc*. Among the downregulated hormogonia genes under UV conditions was a putative partner switching regulatory system composed of an anti-sigma factor, and an anti-sigma factor antagonist coined *hcyC* and *hcyA* respectively. Knock-out mutants of these genes exhibited differential induction of hormogonia differentiation under UV irradiation. These experiments revealed that scytonemin synthesis is inhibited in actively differentiating hormogonia, and through this partner switching mechanism, cells that are actively producing scytonemin cannot differentiate into hormogonia. This regulatory linkage illustrates that *Nostoc* has a varied response to UV, which in turn allows for its perseverance in environments with high light exposure.

Together the findings of the present dissertation elaborate on the previous studies of scytonemin synthesis in *N. punctiforme* and represent the first description of a biological role for the *ebo* gene cluster. The third chapter also represents the first description of hormogonia differentiation induced by UV irradiation. Overall, this thesis contributes to the field of cyanobacterial physiology by broadening knowledge about how *N. punctiforme* responds to UV radiation, including how scytonemin is synthesized and exported from the cell. *N. punctiforme* is of particular interest in the field of biotechnology, as its photosynthetic metabolism and genetic tractability make it an attractive strain for development as a platform for drug and chemical production.

Additionally, the development of the FLANDERS pipeline represents a contribution to the field of data mining bioinformatics that allows for the high throughput analysis of genomic neighborhoods.

References

1. Malla S, Sommer MOA. 2014. A sustainable route to produce the scytonemin precursor using *Escherichia coli* †. *Green Chem* 16:3255–3265.
2. Klicki K, Ferreira D, Hamill D, Dirks B, Mitchell N, Garcia-Pichel F. 2018. The widely conserved ebo cluster is involved in precursor transport to the periplasm during scytonemin synthesis in *Nostoc punctiforme*. *MBio* 9:e02266-18.
3. Yurchenko T, Ševčíková T, Strnad H, Butenko A, Eliáš M. 2016. The plastid genome of some eustigmatophyte algae harbours a bacteria-derived six-gene cluster for biosynthesis of a novel secondary metabolite. *Open Biol* 6.
4. Burlinson P, Studholme D, Cambray-Young J, Heavens D, Rathjen J, Hodgkin J, Preston GM. 2013. *Pseudomonas fluorescens* NZI7 repels grazing by *C. elegans*, a natural predator. *ISME J* 7:1126–1138.

REFERENCES

1. Whitton BA, Potts M. 2012. Introduction to the cyanobacteria. *Ecol Cyanobacteria II Their Divers Sp Time*. Springer Netherlands.
2. Rossi F, Potrafka RM, Pichel G, De Philippis R. 2012. The role of the exopolysaccharides in enhancing hydraulic conductivity of biological soil crusts. *Soil Biol Biochem* 46:33–40.
3. Quesada A, Vincent WF. 1997. Strategies of adaptation by antarctic cyanobacteria to ultraviolet radiation. *Eur J Phycol* 32:335–342.
4. Castenholz RW. 1992. Species Usage, Concept, and Evolution in the Cyanobacteria (Blue-Green Algae). *J Phycol* 28:737–745.
5. Bryant DA. 1994. *The Molecular Biology of Cyanobacteria*. Springer Netherlands.
6. Saini DK, Pabbi S, Shukla P. 2018. Cyanobacterial pigments: Perspectives and biotechnological approaches. *Food Chem Toxicol* 120:616–624.
7. Niedzwiedzki DM, Swainsbury DJK, Canniffe DP, Neil Hunter C, Hitchcock A. 2020. A photosynthetic antenna complex foregoes unity carotenoid-to-bacteriochlorophyll energy transfer efficiency to ensure photoprotection. *Proc Natl Acad Sci U S A* 117:6502–6508.
8. Zhang H, Liu H, Niedzwiedzki DM, Prado M, Jiang J, Gross ML, Blankenship RE. 2014. Molecular mechanism of photoactivation and structural location of the cyanobacterial orange carotenoid protein. *Biochemistry* 53:13–19.
9. Hayashi Y, Ito T, Yoshimura T, Hemmi H. 2018. Utilization of an intermediate of the methylerythritol phosphate pathway, (*E*)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate, as the prenyl donor substrate for various prenyltransferases. *Biosci Biotechnol Biochem* 82:993–1002.
10. Garcia-Pichel F, Castenholz RW. 1991. Characterization And Biological Implications of Scytonemin, A Cyanobacterial Sheath Pigment. *J Phycol* 27:395–409.
11. Potts M. 2006. Nostoc, p. 465–504. *In The Ecology of Cyanobacteria*. Kluwer Academic Publishers.
12. Dodds WK, Gudder DA, Mollenhauer D. 1995. The Ecology of Nostoc. *J Phycol* 31:2–18.
13. Wolk CP, Ernst A, Elhai J. 1994. Heterocyst Metabolism and Development, p. 769–823. *In The Molecular Biology of Cyanobacteria*. Springer Netherlands.
14. Chailakhyan LM, Glagolev AN, Glagoleva TN, Murvanidze G V., Potapova T V., Skulachev VP. 1982. Intercellular power transmission along trichomes of

- cyanobacteria. *BBA - Bioenerg* 679:60–67.
15. Marsac NT. 1994. Differentiation of Hormogonia and Relationships with Other Biological Processes, p. 825–842. *In The Molecular Biology of Cyanobacteria*. Springer Netherlands.
 16. Adams DG, Duggan PS. 1999. Heterocyst and akinete differentiation in cyanobacteria. *New Phytol*. Cambridge University Press.
 17. Mazor G, Kidron GJ, Vonshak A, Abeliovich A. 2006. The role of cyanobacterial exopolysaccharides in structuring desert microbial crusts. *FEMS Microbiol Ecol* 21:121–130.
 18. Lüttge U. 1997. Cyanobacterial Tintenstrich Communities and their Ecology. 526 *Naturwissenschaften* 84:526–534.
 19. Meeks JC, Elhai J, Thiel T, Potts M, Larimer F, Lamerdin J, Predki P, Atlas R. 2001. An overview of the Genome of *Nostoc punctiforme*, a multicellular, symbiotic Cyanobacterium. *Photosynth Res* 70:85–106.
 20. Bentley R. 1999. Secondary metabolite biosynthesis: The first century. *Crit Rev Biotechnol* 19:1–40.
 21. Pichersky E, Gang DR. 2000. Genetics and biochemistry of secondary metabolites in plants: An evolutionary perspective. *Trends Plant Sci* 5:439–445.
 22. Thapa SS, Grove A. 2019. Do Global Regulators Hold the Key to Production of Bacterial Secondary Metabolites? *Antibiotics* 8:160.
 23. Rafailidis PI, Ioannidou EN, Falagas ME. 2007. Ampicillin/sulbactam: Current status in severe bacterial infections. *Drugs* 67:1829–1849.
 24. Vining LC. 1990. Functions of Secondary Metabolites. *Annu Rev Microbiol* 44:395–427.
 25. Tyc O, Song C, Dickschat JS, Vos M, Garbeva P. 2017. The Ecological Role of Volatile and Soluble Secondary Metabolites Produced by Soil Bacteria. *Trends Microbiol* 25:280–292.
 26. Jiménez-Bremont JF, Marina M, Guerrero-González M de la L, Rossi FR, Sánchez-Rangel D, Rodríguez-Kessler M, Ruiz OA, Gárriz A. 2014. Physiological and molecular implications of plant polyamine metabolism during biotic interactions. *Front Plant Sci* 5:1–14.
 27. Smercina DN, Evans SE, Friesen ML, Tiemann LK. 2019. To fix or not to fix: Controls on free-living nitrogen fixation in the rhizosphere. *Appl Environ Microbiol* 85:e02546-18.
 28. Guan N, Li J, Shin H dong, Du G, Chen J, Liu L. 2017. Microbial response to environmental stresses: from fundamental mechanisms to practical applications.

- Appl Microbiol Biotechnol 101:3991–4008.
29. Hider RC, Kong X. 2010. Chemistry and biology of siderophores † ‡. *Nat Prod Rep* 27:637–657.
 30. Gao Q, Garcia-Pichel F. 2011. Microbial ultraviolet sunscreens. *Nat Rev Microbiol* 9:791–802.
 31. Barnard AML, Bowden SD, Burr T, Coulthurst SJ, Monson RE, Salmond GPC. 2007. Quorum sensing, virulence and secondary metabolite production in plant soft-rotting bacteria. *Philos Trans R Soc B Biol Sci* 362:1165–1183.
 32. Stone MJ, Williams DH. 1992. On the evolution of functional secondary metabolites (natural products). *Mol Microbiol* 6:29–34.
 33. Challis GL, Naismith JH. 2004. Structural aspects of non-ribosomal peptide biosynthesis. *Curr Opin Struct Biol* 14:748–756.
 34. Stachelhaus T, Marahiel MA. 1995. Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis. *FEMS Microbiol Lett* 125:3–14.
 35. Zhang L, Hashimoto T, Qin B, Hashimoto J, Kozono I, Kawahara T, Okada M, Awakawa T, Ito T, Asakawa Y, Ueki M, Takahashi S, Osada H, Wakimoto T, Ikeda H, Shin-ya K, Abe I. 2017. Characterization of Giant Modular PKSs Provides Insight into Genetic Mechanism for Structural Diversification of Aminopolyol Polyketides. *Angew Chemie* 129:1766–1771.
 36. Chen H, Du L. 2016. Iterative polyketide biosynthesis by modular polyketide synthases in bacteria. *Appl Microbiol Biotechnol* 100:541–557.
 37. Baltz RH. 2006. Molecular engineering approaches to peptide, polyketide and other antibiotics. *Nat Biotechnol* 24:1533–1540.
 38. Forsythe P, Paterson S. 2014. Ciclosporin 10 years on: Indications and efficacy. *Vet Rec* 174:13–21.
 39. Smela ME, Currier SS, Bailey EA, Essigmann JM. 2001. The chemistry and biology of aflatoxin B1: From mutational spectrometry to carcinogenesis. *Carcinogenesis* 22:535–545.
 40. Mavrodi D V., Peever TL, Mavrodi O V., Parejko JA, Raaijmakers JM, Lemanceau P, Mazurier S, Heide L, Blankenfeldt W, Weller DM, Thomashow LS. 2010. Diversity and evolution of the Phenazine Biosynthesis Pathways. *Appl Environ Microbiol* 76:866–879.
 41. Blankenfeldt W, Parsons JF. 2014. The structural biology of phenazine biosynthesis. *Curr Opin Struct Biol* 29:26–33.
 42. Saleh O, Gust B, Boll B, Fiedler HP, Heide L. 2009. Aromatic prenylation in

- phenazine biosynthesis: Dihydrophenazine-1-carboxylate dimethylallyltransferase from streptomyces anulatus. *J Biol Chem* 284:14439–14447.
43. Narsing Rao MP, Xiao M, Li WJ. 2017. Fungal and bacterial pigments: Secondary metabolites with wide applications. *Front Microbiol* 8:1113.
 44. Melander RJ, Minvielle MJ, Melander C. 2014. Controlling bacterial behavior with indole-containing natural products and derivatives. *Tetrahedron* 70:6363–6372.
 45. Gershenzon J, Dudareva N. 2007. The function of terpene natural products in the natural world. *Nat Chem Biol* 3:408–414.
 46. Dickschat JS. 2016. Bacterial terpene cyclases. *Nat Prod Rep* 33:87–110.
 47. Rathbone DA, Bruce NC. 2002. Microbial transformation of alkaloids. *Curr Opin Microbiol* 5:274–281.
 48. Proteau PJ, Gerwick WH, Garcia-Pichel F, Castenholz R. 1993. The structure of scytonemin, an ultraviolet sunscreen pigment from the sheaths of cyanobacteria. *Experientia* 49:825–829.
 49. Walsh CT, Fischbach MA. 2010. Natural Products Version 2.0: Connecting Genes to Molecules Introduction: Why Do Natural Products Still Matter? *J Am Chem Soc* 132:2469–2493.
 50. Gokulan K, Khare S, Cerniglia C. 2014. METABOLIC PATHWAYS | Production of Secondary Metabolites of Bacteria, p. 561–569. *In Encyclopedia of Food Microbiology*. Elsevier.
 51. Sidebottom AM, Carlson EE. 2015. A reinvigorated era of bacterial secondary metabolite discovery. *Curr Opin Chem Biol* 24:104–111.
 52. Corbell N, Loper JE. 1995. A global regulator of secondary metabolite production in *Pseudomonas fluorescens* Pf-5. *J Bacteriol* 177:6230–6236.
 53. Gallo M, Katz E. 1972. Regulation of Secondary Metabolite Biosynthesis: Catabolite Repression of Phenoxazinone Synthase and Actinomycin Formation by Glucose. *J Bacteriol* 109:659–667.
 54. Rutherford ST, Bassler BL. 2012. Bacterial quorum sensing: Its role in virulence and possibilities for its control. *Cold Spring Harb Perspect Med* 2:a012427.
 55. Blatchley ER, Dumoutier N, Halaby TN, Levi Y, Laîné JM. 2001. Bacterial responses to ultraviolet irradiation, p. 179–186. *In Water Science and Technology*. IWA Publishing.
 56. Rastogi RP, Sinha RP, Moh SH, Lee TK, Kottuparambil S, Kim Y-J, Rhee J-S, Choi E-M, Brown MT, Häder D-P, Han T. 2014. Ultraviolet radiation and cyanobacteria. *J Photochem Photobiol B Biol* 154–169.

57. Rastogi RP, Richa, Kumar A, Tyagi MB, Sinha RP. 2010. Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. *J Nucleic Acids* 2010:592980.
58. Nadeau T, Howard-Williams C, Castenholz R. 1999. Effects of solar UV and visible irradiance on photosynthesis and vertical migration of *Oscillatoria* sp. (Cyanobacteria) in an Antarctic microbial mat. *Aquat Microb Ecol* 20:231–243.
59. Singh SP, Hä Der D-P, Sinha RP. Cyanobacteria and ultraviolet radiation (UVR) stress: Mitigation strategies. *Ageing Res Rev* 9:79–90.
60. Oren A, Gunde-Cimerman N. 2007. Mycosporines and mycosporine-like amino acids: UV protectants or multipurpose secondary metabolites? *FEMS Microbiol Lett* 269:1–10.
61. Britton G, Armit GM, Lau SYM, Patel AK, Shone CC. 1982. Carotenoproteins, p. 237–251. *In* Carotenoid Chemistry and Biochemistry. Elsevier.
62. Latifi A, Ruiz M, Zhang CC. 2009. Oxidative stress in cyanobacteria. *FEMS Microbiol Rev* 33:258–278.
63. McCord JM, Fridovich I. 1969. Superoxide Dismutase. *J Biol Chem* 244:6049–6055.
64. Fita I, Rossmann MG. 1985. The active center of catalase. *J Mol Biol* 185:21–37.
65. Karsten U, Garcia-Pichel F. 1996. Carotenoids and Mycosporine-like Amino Acid Compounds in Members of the Genus *Microcoleus* (Cyanobacteria): A Chemosystematic Study. *Syst Appl Microbiol* 19:285–294.
66. Garcia-Pichel F, Wingard CE, Castenholz RW. 1993. Evidence regarding the ultraviolet sunscreen role of the Mycosporin-like compounds in the cyanobacterium *Gloeocapsa* sp. *Appl Environ Microbiol* 59:170–176.
67. Favre-Bonvin J, Bernillon J, Salin N, Arpin N. 1987. Biosynthesis of mycosporines: Mycosporine glutaminol in *Trichothecium roseum*. *Phytochemistry* 26:2509–2514.
68. Takanoa S, Uemurab D, Hiratab Y. 1977. Isolation and Structure of A Mycosporine from the Zoanthid *Palythoa Tuberculosa*. *Tetrahedron Lett* 28:2429–2430.
69. Bohm G, Pflleiderer W, Schere S. Structure of a Novel Oligosaccharide-Mycosporine-Amino Acid Ultraviolet A/B Sunscreen Pigment from the Terrestrial Cyanobacterium *Nostoc commune*. *J Biol Chem* 270:8536–8539.
70. Ehling-Schulz M, Scherer S. 1999. UV protection in cyanobacteria. *Eur J Phycol* 34:329–338.
71. Plonka PM, Grabačka M. 2006. Melanin synthesis in microorganisms-

- biotechnological and medical aspects *. *Acta Biochim Pol* 53:429–443.
72. Romero-Martinez R, Wheeler M, Guerrero-Plata A, Rico G, Torres-Guerrero H. 2000. Biosynthesis and functions of melanin in *Sporothrix schenckii*. *Infect Immun* 68:3696–3703.
 73. Funa N, Funabashi M, Ohnishi Y, Horinouchi S. 2005. Biosynthesis of hexahydroxyperylenequinone melanin via oxidative aryl coupling by cytochrome P-450 in *Streptomyces griseus*. *J Bacteriol* 187:8149–8155.
 74. Garcia-Pichel F, Castenholz RW. 1991. Characterization and Biological Implications of Scytonemin, a Cyanobacterial Sheath Pigment. *J Phycol* 27:395–409.
 75. Soule T, Stout V, Swingley WD, Meeks JC, Garcia-Pichel F. 2007. Molecular genetics and genomic analysis of scytonemin biosynthesis in *Nostoc punctiforme* ATCC 29133. *J Bacteriol* 189:4465–72.
 76. Soule T, Garcia-Pichel F, Stout V. 2009. Gene expression patterns associated with the biosynthesis of the sunscreen scytonemin in *Nostoc punctiforme* ATCC 29133 in response to UVA radiation. *J Bacteriol* 191:4639–4646.
 77. Balskus EP, Walsh CT. 2009. An Enzymatic Cyclopentyl[b]indole Formation Involved in Scytonemin Biosynthesis. *J Am Chem Soc* 131:14648–14649.
 78. Soule T, Palmer K, Gao Q, Potrafka R, Stout V, Garcia-Pichel F. 2009. A comparative genomics approach to understanding the biosynthesis of the sunscreen scytonemin in cyanobacteria. *BMC Genomics* 10:336.
 79. Naurin S, Bennett J, Videau P, Philmus B, Soule T. 2016. The response regulator Npun_F1278 is essential for scytonemin biosynthesis in the cyanobacterium *Nostoc punctiforme* ATCC 29133. *J Phycol* 52:564–571.
 80. Ferreira D, Garcia-Pichel F. 2016. Mutational Studies of Putative Biosynthetic Genes for the Cyanobacterial Sunscreen Scytonemin in *Nostoc punctiforme* ATCC 29133. *Front Microbiol* 7:1–10.
 81. Malla S, Sommer MOA. 2014. A sustainable route to produce the scytonemin precursor using *Escherichia coli* †. *Green Chem* 16:3255–3265.
 82. Yurchenko T, Ševčíková T, Strnad H, Butenko A, Eliáš M. 2016. The plastid genome of some eustigmatophyte algae harbours a bacteria-derived six-gene cluster for biosynthesis of a novel secondary metabolite. *Open Biol* 6.
 83. Burlinson P, Studholme D, Cambray-Young J, Heavens D, Rathjen J, Hodgkin J, Preston GM. 2013. *Pseudomonas fluorescens* NZI7 repels grazing by *C. elegans*, a natural predator. *ISME J* 7:1126–1138.
 84. Artsimovitch I. 2018. *Rebuilding the bridge between transcription and translation*. *Mol Microbiol*. Blackwell Publishing Ltd.

85. Stent GS. 1964. The Operon: On Its Third Anniversary New Series.
86. Fondi M, Emiliani G, Fani R. 2009. Origin and evolution of operons and metabolic pathways. *Res Microbiol* 160:502–512.
87. Wlodek A, Kendrew SG, Coates NJ, Hold A, Pogwizd J, Rudder S, Sheehan LS, Higginbotham SJ, Stanley-Smith AE, Warneck T, Nur-E-Alam M, Radzom M, Martin CJ, Overvoorde L, Samborsky M, Alt S, Heine D, Carter GT, Graziani EI, Koehn FE, McDonald L, Alanine A, Rodríguez Sarmiento RM, Chao SK, Ratni H, Steward L, Norville IH, Sarkar-Tyson M, Moss SJ, Leadlay PF, Wilkinson B, Gregory MA. 2017. Diversity oriented biosynthesis via accelerated evolution of modular gene clusters. *Nat Commun* 8:1206.
88. Gomez-Escribano JP, Song L, Fox DJ, Yeo V, Bibb MJ, Challis GL. 2012. Structure and biosynthesis of the unusual polyketide alkaloid coelimycin P1, a metabolic product of the cpk gene cluster of *Streptomyces coelicolor* M145. *Chem Sci* 3:2716–2720.
89. Schwarzer D, Finking R, Marahiel MA. 2003. Nonribosomal peptides: From genes to products. *Nat Prod Rep* 20:275–287.
90. Vasudevan R, Gale GAR, Schiavon AA, Puzorjov A, Malin J, Gillespie MD, Vavitsas K, Zulkower V, Wang B, Howe CJ, Lea-Smith DJ, McCormick AJ. 2019. CyanoGate: A Modular Cloning Suite for Engineering Cyanobacteria Based on the Plant MoClo Syntax. *Plant Physiol* 180:39–55.
91. Klein CA, Emde L, Kuijpers A, Sobetzko P. 2019. MoCloFlex: A Modular Yet Flexible Cloning System. *Front Bioeng Biotechnol* 7:271.
92. Hausner J, Jordan M, Otten C, Marillonnet S, Büttner D. 2019. Modular Cloning of the Type III Secretion Gene Cluster from the Plant-Pathogenic Bacterium *Xanthomonas euvesicatoria*. *ACS Synth Biol* 8:532–547.
93. Jensen PR. 2016. Natural Products and the Gene Cluster Revolution. *Trends Microbiol* 24:968–977.
94. von Heijne G. 1990. The signal peptide. *J Membr Biol* 115:195–201.
95. Klicki K, Ferreira D, Hamill D, Dirks B, Mitchell N, Garcia-Pichel F. 2018. The Widely Conserved ebo Cluster Is Involved in Precursor Transport to the Periplasm during Scytonemin Synthesis in *Nostoc punctiforme* 9.
96. Blin K. 2020. ncbi-acc-download. 0.2.7.
97. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* 47:W81–W87.
98. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide

- predictions using deep neural networks. *Nat Biotechnol* 37:420–423.
99. Medema MH, Takano E, Breitling R. 2013. Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol Biol Evol* 30:1218–23.
 100. Da T, Agostini-Costa S, Vieira RF, Bizzo HR, Silveira D, Gimenes MA. 8 Secondary Metabolites.
 101. Bertelli C, Laird MR, Williams KP, Lau BY, Hoad G, Winsor GL, Brinkman FSL. 2017. IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res* 45:W30–W35.
 102. Juhas M. 2015. Type IV secretion systems and genomic islands-mediated horizontal gene transfer in *Pseudomonas* and *Haemophilus*. *Microbiol Res* 170:10–17.
 103. Penn K, Jenkins C, Nett M, Udvary DW, Gontang EA, McGlinchey RP, Foster B, Lapidus A, Podell S, Allen EE, Moore BS, Jensen PR. 2009. Genomic islands link secondary metabolism to functional adaptation in marine Actinobacteria. *ISME J* 3:1193–1203.
 104. Babakhani S, Oloomi M. 2018. Transposons: the agents of antibiotic resistance in bacteria. *J Basic Microbiol* 58:905–917.
 105. O’connor SE, Maresh JJ. 2005. Chemistry and biology of monoterpene indole alkaloid biosynthesis. *Nat Prod Rep* 23:523–547.
 106. Smirnova GV, Oktyabrsky ON. 2005. Glutathione in Bacteria. *Biochemistry* 70:1459–1473.
 107. Quistgaard EM, Löw C, Guettou F, Nordlund P. 2016. Understanding transport by the major facilitator superfamily (MFS): Structures pave the way. *Nat Rev Mol Cell Biol* 17:123–132.
 108. Hollenstein K, Dawson RJ, Locher KP. 2007. Structure and mechanism of ABC transporter proteins. *Curr Opin Struct Biol* 17:412–418.
 109. Kornberg A, Rao NN, Ault-Riché D. 1999. Inorganic Polyphosphate: A Molecule of Many Functions. *Annu Rev Biochem* 68:89–125.
 110. Ii B, Freiburg D-. 1991. in *Pure Cultures of Sewage Bacteria* 25:9–13.
 111. Gao R, Stock AM. 2009. Biological insights from structures of two-component proteins. *Annu Rev Microbiol* 63:133–154.
 112. Magasanik B. 1961. Catabolite repression. *Cold Spring Harb Symp Quant Biol* 26:249–256.
 113. Ackermann M. 2015. A functional perspective on phenotypic heterogeneity in microorganisms. *Nat Rev Microbiol* 13:497–508.

114. Smits WK, Kuipers OP, Veening JW. 2006. Phenotypic variation in bacteria: The role of feedback regulation. *Nat Rev Microbiol* 4:259–271.
115. Beaumont HJE, Gallie J, Kost C, Ferguson GC, Rainey PB. 2009. Experimental evolution of bet hedging. *Nature* 462:90–93.
116. Levy SF, Ziv N, Siegal ML. 2012. Bet Hedging in Yeast by Heterogeneous, Age-Correlated Expression of a Stress Protectant. *PLoS Biol* 10:1001325.
117. Holland SL, Reader T, Dyer PS, Avery S V. 2014. Phenotypic heterogeneity is a selected trait in natural yeast populations subject to environmental stress. *Environ Microbiol* 16:1729–1740.
118. Avery S V. 2006. Microbial cell individuality and the underlying sources of heterogeneity. *Nat Rev Microbiol*. Nature Publishing Group.
119. Raj A, van Oudenaarden A. 2008. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell* 135:216–226.
120. van Gemerden H. 1993. Microbial mats: A joint venture. *Mar Geol* 113:3–25.
121. Revsbech NP, Jorgensen BB. 1983. Photosynthesis of benthic microflora measured with high spatial resolution by the oxygen microprofile method: Capabilities and limitations of the method. *Limnol Oceanogr* 28:749–756.
122. Potts M. 1994. Desiccation tolerance of prokaryotes. *Microbiol Rev* 58:755–805.
123. Soule T, Stout V, Swingley WD, Meeks JC, Garcia-Pichel F. 2007. Molecular genetics and genomic analysis of scytonemin biosynthesis in *Nostoc punctiforme* ATCC 29133. *J Bacteriol* 189:4465–72.
124. Bebout BM, Garcia-Pichel F. 1995. UV B-Induced Vertical Migrations of Cyanobacteria in a Microbial Mat. *Appl Environ Microbiol* 61:4215–4222.
125. Meeks J, Campbell E, Summers M, Wong F. 2002. Cellular differentiation in the cyanobacterium *Nostoc punctiforme*. *Arch Microbiol* 178:395–403.
126. Risser DD, Chew WG, Meeks JC. 2014. Genetic characterization of the hmp locus, a chemotaxis-like gene cluster that regulates hormogonium development and motility in *Nostoc punctiforme*. *Mol Microbiol* 92:222–233.
127. Dodds WK, Gudder DA, Mollenhauer D. 1995. The Ecology of *Nostoc*. *J Phycol* 31:2–18.
128. Campbell EL, Christman H, Meeks JC. 2008. DNA microarray comparisons of plant factor and nitrogen deprivation-induced hormogonia reveal decision-making transcriptional regulation patterns in *Nostoc punctiforme*. *J Bacteriol* 190:7382–7391.
129. Ekman M, Picossi S, Campbell EL, Meeks JC, Flores E. A *Nostoc punctiforme* Sugar Transporter Necessary to Establish a Cyanobacterium-Plant Symbiosis.

Plant Physiol 161:1984–1992.

130. Rajeev L, Luning EG, Dehal PS, Price MN, Arkin AP, Mukhopadhyay A. 2011. Systematic mapping of two component response regulators to gene targets in a model sulfate reducing bacterium. *Genome Biol* 12:R99.
131. Khayatan B, Meeks JC, Risser DD. 2015. Evidence that a modified type IV pilus-like system powers gliding motility and polysaccharide secretion in filamentous cyanobacteria. *Mol Microbiol* 98:1021–1036.
132. Armstrong RE, Hayes PK, Walsby AE. 1983. Gas Vacuole Formation in Hormogonia of *Nostoc muscorum*. *Microbiology* 129:263–270.
133. Gonzalez A, Riley KW, Harwood T V, Zuniga EG, Risser DD, Ellermeier CD. 2019. A Tripartite, Hierarchical Sigma Factor Cascade Promotes Hormogonium Development in the Filamentous Cyanobacterium *Nostoc punctiforme* 4:e00231-19.
134. Davis MC, Kesthely CA, Franklin EA, MacLellan SR. 2017. The essential activities of the bacterial sigma factor. *Can J Microbiol* 63:89–99.
135. Soule T, Gao Q, Stout V, Garcia-pichel F. 2013. The Global Response of *Nostoc punctiforme* ATCC 29133 to UVA Stress , Assessed in a Temporal DNA Microarray Study. *Photochem Photobiol* 89:415–423.
136. Cai Y, Wolk CP. 1990. Use of a Conditionally Lethal Gene in *Anabaena* sp. Strain PCC 7120 To Select for Double Recombinants and To Entrap Insertion Sequences. *J Bacteriol* 172:3138–3145.
137. Cohen MF, Wallis JG, Campbell EL, Meeks JC. Transposon mutagenesis of *Nostoc* sp. strain ATCC 29133, a filamentous cyanobacterium with multiple cellular differentiation alternatives. *Microbiology* 140:3233–3240.
138. Klicki K, Ferreira D, Hamill D, Dirks B, Mitchell N, Garcia-Pichel F. 2018. The widely conserved ebo cluster is involved in precursor transport to the periplasm during scytonemin synthesis in *Nostoc punctiforme*. *MBio* 9:e02266-18.
139. Campbell EL, Summers ML, Christman H, Martin ME, Meeks JC. 2007. Global gene expression patterns of *Nostoc punctiforme* in steady-state dinitrogen-grown heterocyst-containing cultures and at single time points during the differentiation of akinetes and hormogonia. *J Bacteriol* 189:5247–56.
140. Tjaden B. 2015. De novo assembly of bacterial transcriptomes from RNA-seq data. *Genome Biol* 16:1.
141. Kolde R, Antoine L. 2015. pheatmap: Pretty Heatmaps. <https://github.com/raivokolde/pheatmap>.
142. Karimova G, Pidoux J, Ullmann A, Ladant D. 1998. A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci U S A*

- 95:5752–5756.
143. Battesti A, Bouveret E. 2012. The bacterial two-hybrid system based on adenylate cyclase reconstitution in *Escherichia coli*. *Methods* 58:325–334.
 144. Castenholz RW, Garcia-Pichel F. 2007. Cyanobacterial Responses to UV-Radiation. *Ecol Cyanobacteria* 21:591–611.
 145. De Jong IG, Haccou P, Kuipers OP. 2011. Bet hedging or not? A guide to proper classification of microbial survival strategies. *BioEssays* 33:215–223.
 146. Splitt SD, Risser DD. 2016. The non-metabolizable sucrose analog sucralose is a potent inhibitor of hormogonium differentiation in the filamentous cyanobacterium *Nostoc punctiforme*. *Arch Microbiol* 198:137–147.
 147. Campbell EL, Christman H, Meeks JC. 2008. DNA microarray comparisons of plant factor- and nitrogen deprivation-induced Hormogonia reveal decision-making transcriptional regulation patterns in *Nostoc punctiforme*. *J Bacteriol* 190:7382–91.
 148. Holland SL, Reader T, Dyer PS, Avery S V. 2014. Phenotypic heterogeneity is a selected trait in natural yeast populations subject to environmental stress. *Environ Microbiol* 16:1729–1740.
 149. Schurr MJ, Martin DW, Mudd MH, Hibler NS, Boucher JC, Deretic V. 1993. The algD promoter: Regulation of alginate production by *Pseudomonas aeruginosa* in cystic fibrosis. *Cell Mol Biol Res* 39:371–376.
 150. McCaw ML, Lykken GL, Singh PK, Yahr TL. 2002. ExsD is a negative regulator of the *Pseudomonas aeruginosa* type III secretion regulon. *Mol Microbiol* 46:1123–1133.
 151. Mohr W, Vagner T, Kuypers MMM, Ackermann M, LaRoche J. 2013. Resolution of Conflicting Signals at the Single-Cell Level in the Regulation of Cyanobacterial Photosynthesis and Nitrogen Fixation. *PLoS One* 8:1–7.
 152. Adams DG. 2000. Heterocyst formation in cyanobacteria. *Curr Opin Microbiol* 3:618–624.
 153. Wösten MMSM. 1998. Eubacterial sigma-factors. *FEMS Microbiol Rev* 22:127–150.
 154. Helmann JD, Chamberlin MJ. 1988. Structure and Function of Bacterial Sigma Factors. *Annu Rev Biochem* 57:839–872.
 155. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko M V., Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH. 2011. CDD: A Conserved Domain Database for the functional

- annotation of proteins. *Nucleic Acids Res* 39:D225–D229.
156. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Res* 43:D222–D226.
 157. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Geer LY, Bryant SH. 2017. CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 45:D200–D203.
 158. Wagner JR, Brunzelle JS, Forest KT, Vierstra RD. 2005. A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* 438:325–331.
 159. Wiltbank LB, Kehoe DM. 2019. Diverse light responses of cyanobacteria mediated by phytochrome superfamily photoreceptors. *Nat Rev Microbiol* 17:37–50.
 160. Delumeau O, Dutta S, Brigulla M, Kuhnke G, Hardwick SW, Völker U, Yudkin MD, Lewis RJ. 2004. Functional and structural characterization of RsbU, a stress signaling protein phosphatase 2C. *J Biol Chem* 279:40927–40937.
 161. Riley KW, Gonzalez A, Risser DD. 2018. A partner-switching regulatory system controls hormogonium development in the filamentous cyanobacterium *Nostoc punctiforme*. *Mol Microbiol* 109:555–569.
 162. Morris AR, Visick KL. 2013. The response regulator SypE controls biofilm formation and colonization through phosphorylation of the *syp*-encoded regulator SypA in *Vibrio fischeri*. *Mol Microbiol* 87:509–525.
 163. Mercer RG, Lang AS. 2014. Identification of a predicted partner-switching system that affects production of the gene transfer agent RcGTA and stationary phase viability in *Rhodobacter capsulatus*. *BMC Microbiol* 14:1–16.
 164. Riley KW, Gonzalez A, Risser DD. 2018. A partner-switching regulatory system controls hormogonium development in the filamentous cyanobacterium *Nostoc punctiforme*. *Mol Microbiol* 109:555–569.
 165. Bauernfeind A, Burrows JR. 1978. Suggested Procedure Allowing Use of Plastic Petri Dishes in Bacteriocin Typing. *Appl Environ Microbiol* 35:970.
 166. Gao Q, Garcia-Pichel F. 2011. An ATP-Grasp Ligase Involved in the Last Biosynthetic Step of the Iminomycosporine Shinorine in *Nostoc punctiforme* ATCC 29133. *J Bacteriol* 193:5923–5928.
 167. Jöhler S, Stephan R, Althaus D, Ehling-Schulz M, Grunert T. 2016. High-resolution subtyping of *Staphylococcus aureus* strains by means of Fourier-

transform infrared spectroscopy. *Syst Appl Microbiol* 39:189–194.

APPENDIX A

CELLULAR AND MOLECULAR CHARACTERIZATION OF *EBO* GENE CLUSTER

KNOCKOUT MUTANTS IN THE SOIL BACTERIUM *PSEUDOMONAS*

FLUORESCENS NZ17

Abstract

To investigate the anti-nematodal capacities of *Pseudomonas fluorescens* NZ17, Burlinson et al 2013 created knock-out mutants that demonstrated the involvement of the *ebo* gene cluster in the production of a putative anti-nematodal compound. Mutants which had interruptions in any one of the *ebo* genes or the 6 preceding genes were edible by *Caenorhabditis elegans*, and the full 12 gene cluster was thus named the ‘edible’ or *edb* cluster. The findings of Klicki et al 2018 implicating the *ebo* gene cluster in scytonemin synthesis prompted interest in the *edb* cluster mutants, and thus the following experiments were performed to characterize the putative anti-nematodal agent: its diffusive properties were assessed through wild-type/*edb* mutant mixed culture incubations, its passive defense capabilities were investigated using killed wild-type and *edb* mutant cultures, and its physical properties were assessed through an array of analytical chemistry techniques. The putative anti-nematodal compound was found to be very poorly diffusive but can defend inactive cells from predation. The analytical techniques implemented were broadly inconclusive. Further experimentation will be needed to fully elucidate the structure and biochemical attributes of the putative *edb* encoded anti-nematodal compound.

Introduction

The discovery of the *ebo* gene cluster (1) and its implication in scytonemin synthesis (2) prompted a literature search for further examples of the cluster, if by another name than *ebo*. The annotations for *eboA-F* (hypothetical protein, TatD hydrolase, UbiA prenyltransferase, 5-dehydroquinase synthase, TIM-barrel containing xylose isomerase,

and type I phosphodiesterase) were searched for cooccurrence in google scholar, yielding one surprising result. While investigating the anti-nematodal properties of the mushroom pathogen *Pseudomonas fluorescens* NZI7, Burlinson et al 2013 (3) constructed a transposon mutant library from which several strains exhibited a mutant phenotype lacking resistance to predation by the soil nematode *Caenorhabditis elegans*. Sequencing of these mutants revealed a contiguous 12 gene cluster that when interrupted abolished nematode repellence. They designated this gene cluster as the *edb* gene cluster for edible, as knock-out mutants became edible by *C. elegans*. Surprisingly, the annotations of the first 6 genes in the *edb* cluster matched the annotations of *eboA-F* exactly, suggesting that like the scytonemin operon in scytonemin producing cyanobacteria besides *N. punctiforme* ATCC 29133, the *ebo* clusters is genomically colocalized with the synthesis operon of a hitherto unidentified nematode repellent, and furthermore is necessary for its proper synthesis and/or cellular function. The *ebo* knock-out mutants corresponding to interruptions in *eboA*, *eboC*, *eboD*, and *eboE*, as well as knock-outs of two downstream genes henceforth *edbA* and *edbB* were obtained from the Preston lab and subjected to cellular and molecular characterization to address the following questions:

- I. Is nematode repellence titratable from a population of producing cells to a population of non-producing cells?

Burlinson et al. 2013 (3) conducted a series of experiments to determine colonies of wild-type cells in close proximity to colonies of *edb* mutants could rescue their nematode repellence via presumptive diffusion of the putative anti-nematodal compound through

the media. Their results demonstrated that the putative anti-nematodal compound does not actively diffuse over distances greater than 5 mm. These results raised the question as to if this resistance is diffusible over distances on the order of micrometers or in cases of direct cell-cell contact between wild type and *edb* mutants.

In order to address these questions, colonies of wild-type, $\Delta eboC$, and $\Delta edbA$ were picked and grown overnight in Luria-Bertani broth at 28°C. Optical density of overnight cultures was measured, and aliquots were combined to produce mixtures containing 0%, 25%, 50%, 75% and 100% wild-type: $\Delta eboC$, wild-type: $\Delta edbA$, and $\Delta eboC$: $\Delta edbA$ respectively based on OD. 25 μ l of each mixture were spotted onto nematode growth media (NGM) and incubated at 28°C overnight. After incubation, 10 L4/adult nematodes were placed in a central location and allowed to explore the plate for 120 hours. Plates (n=3) were photographed every 48 hours.

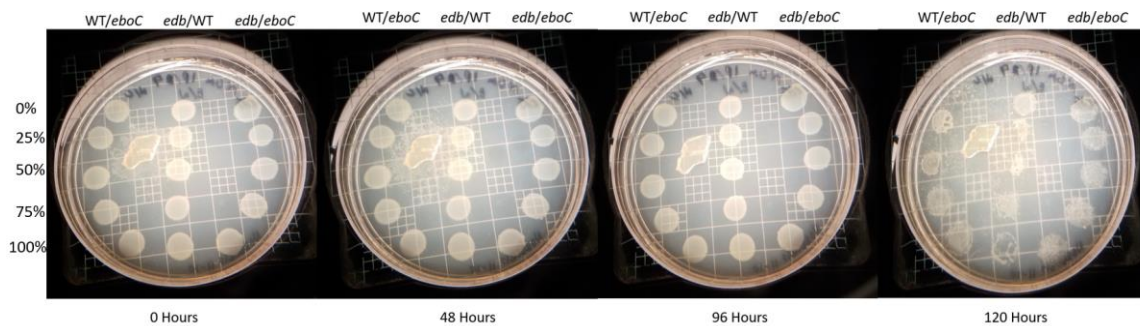


Fig. 1: *C. elegans* predation on mixed cultures of wild-type (WT) *P. fluorescens* NZI7, $\Delta eboC$ (*eboC*), and $\Delta edbA$ (*edb*). Percentages are relative to the first strain listed for each column of spots. After 120 hours, the spots containing no wild-type cells were most severely predated upon, while those with greater percentages of wild-type cells were traversed by *C. elegans*, but not as heavily predated upon.

After *C. elegans* had been on the plates for 120 hours, it was clear that spots which contained no wild-type cells had been devoured to a much higher degree than spots which contained at least some wild-type cells. These results indicate that while the putative anti-nematodal compound produced by the *edb* cluster cannot diffuse at distances greater than 5 mm, it may either be more actively diffusive at micro-scale distances, or perhaps *C. elegans* makes a decision whether or not to feed based on an aggregate assessment of the overall concentration of the putative anti-nematodal compound.

II. Is cell activity required for predation evasion/ does the putative anti-nematodal compound provide passive protection?

All experiments performed by Burlinson et al 2013 (3) were conducted under conditions wherein wild-type and *edb* mutant *P. fluorescens* strains were metabolically active. If the putative *edb* produced anti-nematodal compound can provide protection passively, killed cells should maintain resistance to predation. In order to assess the passive protection capacity of the putative anti-nematodal compound, wild-type *P. fluorescens* NZI7 were chemically killed leaving whole cells intact and the subjected to the standard *C. elegans* feeding assay (3).

Colonies of wild-type, Δ *eboA*, Δ *eboC*, and Δ *edbA* were picked and grown overnight in Luria-Bertani broth at 28°C. 25µl of each overnight culture was spotted onto NGM in separate quadrants and incubated at 28°C overnight. Plated cultures were killed by inverting the plate over a glass plate containing chloroform and incubating for 10 minutes

in the fume hood according to Bauernfeind and Burrows 1978 (4). 10 L4/adult nematodes were placed in a central location on both killed and live plates and allowed to explore the plates for 120 hours at which point plates were photographed.

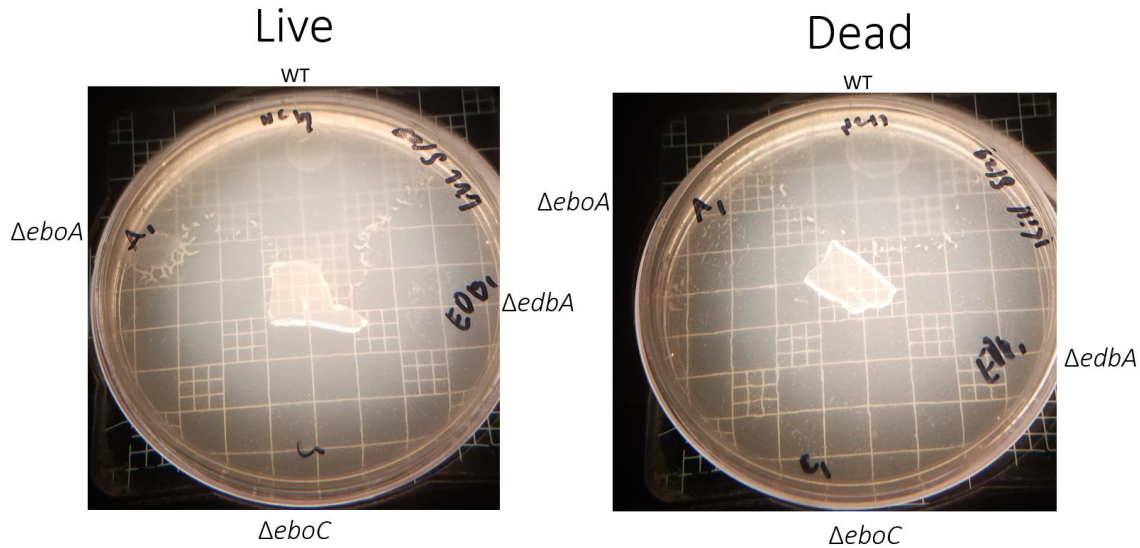


Fig. 2: *C. elegans* predation on killed wild-type (WT), $\Delta eboA$, $\Delta eboC$, and $\Delta edbA$ cells. The patterns of predation between the live and dead plates were identical: *C. elegans* fully consumed all mutants, but left wild-type virtually untouched, even when wild-type cells were not metabolically active.

There was no apparent difference between the live and killed cultures in terms of predation avoidance. Wild-type cells were able to repel predation whether metabolically active or killed by chloroform, while *edb* mutants continued to be edible. These results indicate that active cells are not required to maintain the nematode resistant phenotype, and further suggest that the putative anti-nematodal compound is produced constitutively and impervious to inactivation by cell death. This attribute is strikingly similar to the

case of scytonemin, which can provide passive protection from UV-A during prolonged periods of desiccation, wherein cells are not metabolically active.

III. Is the putative anti-nematodal *edb* product a small molecule, a protein, or something completely different?

The findings of Burlinson et al 2013 (3) along with the previous two experiments strongly suggested that the putative anti-nematodal product of the *edb* gene cluster is a poorly diffusible small molecule secondary metabolite. In attempts to isolate and identify the putative anti-nematodal compound, several small molecule analytical techniques were employed utilizing a range of extraction methods, lending to the mysterious nature of the compound's structure. To identify potential candidate compounds, analysis of wild-type and *ebo* and *edb* mutants were subjected to a series of analytical techniques, and the results compared for compounds present in the wild-type but absent from mutants.

For all chemical analyses, colonies of wild-type and *ebo* and *edb* mutants were inoculated into 15 ml LB and grown at 28°C overnight. 4ml of overnight cultures were plated onto NGM plates and incubated at 28°C overnight. Cells were collected from plates by applying 4 ml sterile PBS to plate surface and resuspending using a sterile cell spreader. Resuspended cells were decanted into 1.5 ml tubes, and residual PBS was removed by vacuum centrifugation. Resulting cell pellets were frozen at -20°C until extraction.

The poor diffusibility of the putative anti-nematodal compound suggested that it may be extractable using a non-polar organic solvent, and therefor acetone, ethyl acetate, and

DMSO respectively were used to extract the lipid soluble compounds from frozen cell pellets. 1 ml of organic solvent and .5 grams of 5 mm zirconium/silicate beads were added to the frozen cell pellet. Pellets were homogenized by bead beating at full force for 160 seconds. Cell lysates were left to extract at 4°C overnight in the dark. Lysates were then spun down, and supernatant transferred to new 1.5 ml tubes for vacuum centrifugation. High pressure liquid chromatography (HPLC) analysis was performed as in Klicki et al 2018 (2) by resuspending dried extracts in 50 µl methanol and running analytes on a Waters 2998 a Supelco HF-5 Discovery (5 µm membrane, 3.0 by 150 mm) normal phase column at 30°C and a flow rate of 1 ml/min with the following solvent composition: 1-5 minutes 20% methanol 80% acetonitrile, 5-15 minutes 100% acetonitrile, 15-20 minutes 20% methanol 80% acetonitrile. Column eluent absorbance was monitored continuously between 200 and 800 nm using a Waters 2998e PDA detector.

To determine if a water soluble compound was responsible for the observed anti-nematodal phenotype, the same extraction protocol was performed using methanol as the extraction solvent. After resuspension in 50 µl methanol, analytes were run as in Gao and Garcia-Pichel 2011 (5) on a Phenomenex Synergi Hydro-RP C18 analytical column (4-µm membrane; 3.0 by 150 mm) at 30°C and a flow rate of 1 ml/min with the following solvent composition: 0-3 min 100% 0.1 M triethylammonium acetate in water, 1.5-15 minutes 100% 0.1 M triethylammonium acetate in water to 50:50 with acetonitrile until 18 minutes, 18-23 minutes return to original composition. Column eluent

absorbance was monitored continuously between 200 and 800 nm using a Waters 2998e PDA detector.

In order to obtain an in-vivo assessment of the carbonaceous compounds present in wild-type and mutant cells, fourier transform infrared spectroscopy was employed. The frozen cell pellet was transferred to a zinc selenite (ZnSe) optical plate. Subsequently, FTIR spectra were recorded as in Johler et al 2016 (6) in transmission mode with an HTS-XT microplate adapter coupled to a Tensor 27 FTIR spectrometer (Bruker Optics GmbH, Ettlingen, Germany) using the following parameters: 4000–500 cm^{-1} spectral range, 6 cm^{-1} spectral resolution, zero-filling factor 4, Blackmann-Harris 3-term apodization, and 64 interferograms were averaged with background subtraction for each spectrum.

All spectra and chromatograms were visualized and analyzed using chemospec.

To gain a broader scope of the chemical differences between wild-type and *edb* mutants, cultures were prepared as previously described and subjected to LC-MS analyses at the West Coast Metabolomics Center. 100 μl of cell extracts were run on a thermo scientific QEHF in both positive and negative ion mode, and data were normalized to sums of internal standards. In total 8482 mass frag were identified across all strains, raw data can be found in supplementary table 1 and can be accessed at <https://drive.google.com/drive/folders/1n9gC5HmX6k5CMOoccbFL78s9StTopI7k?usp=sharing>. cursory examination of the differential abundance of each compound between wild-type and *edb* mutants found no compound was clearly absent from *edb* mutants but present in wild-type.

Given the lack of conclusive results in searching for a difference in small molecules produced between wild-type and mutants, the proteome of wild-type and mutants were compared by SDS-PAGE. Cells were lysed by adding .5 grams of 0.5 mm zirconium/silicate beads and 1 ml of ice cold PBS to frozen cell pellets and bead beating for 160 seconds as previously described. Cell lysates were then spun down and 50 μ l of supernatant were added to 50 μ l of Laemmli loading buffer, and incubated in a water bath at 95°C for 5 minutes prior to loading onto 12% agarose gel placed in Tris-glycine-EDTA buffer. Proteins were run at 120 V for one hour. Gels were stained in coomassie blue for 3 hours, then destained in ddH₂O overnight. Gels were visualized on a Biorad Gel-doc.

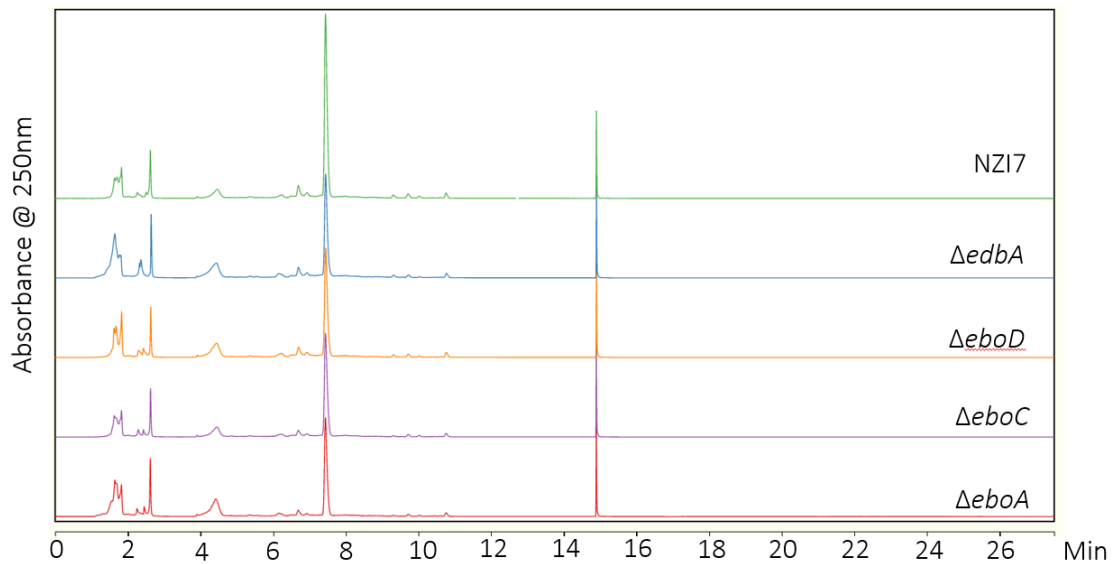


Fig. 3: Chromatograms of HPLC separation of organic solvent-based extractions of wild-type (NZI7), *edbA*, *eboD*, *eboC*, and *eboA* mutants. No significant differences could be discerned between mutants and wild-type for any organic solvent used (Acetone extract chromatograms shown)

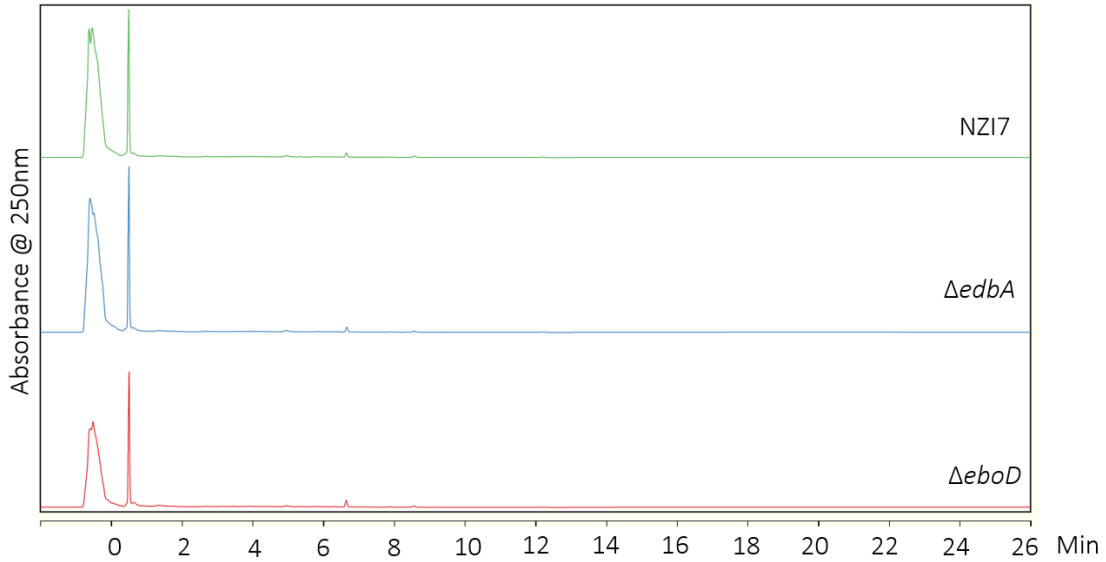


Fig. 4: Fig. 3: Chromatograms of HPLC separation of methanol-based extractions of wild-type (NZI7) and *edbA* and *eboD* mutants. No significant differences could be discerned between mutants and wild-type.

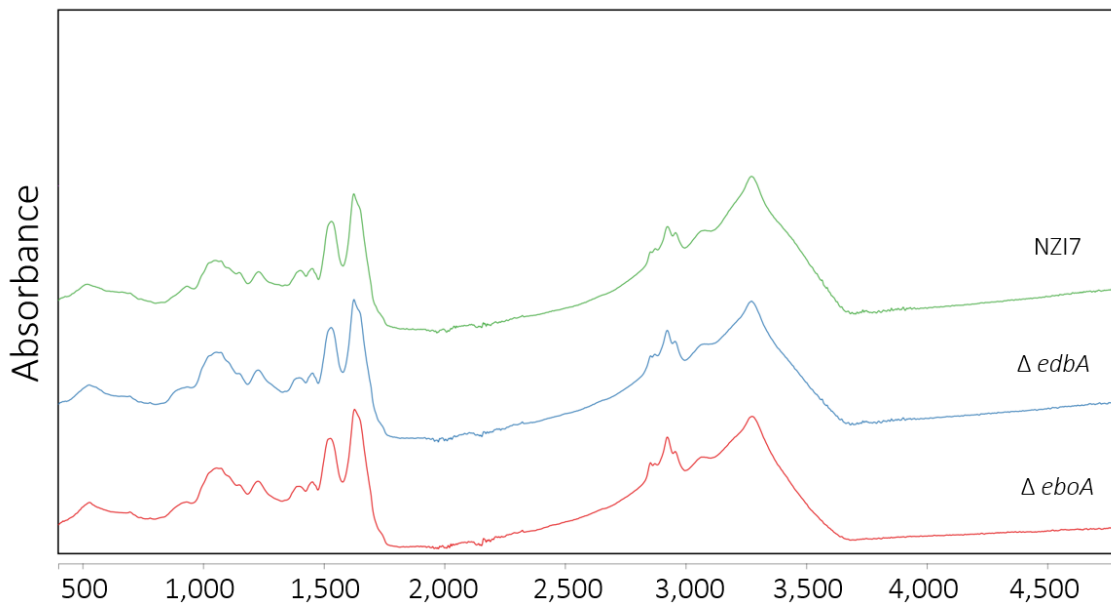


Fig. 5: Infrared spectra of wild-type (NZI7), *edbA* and *eboA* mutant dried cell pellets. No significant differences could be discerned between mutants and wild-type.

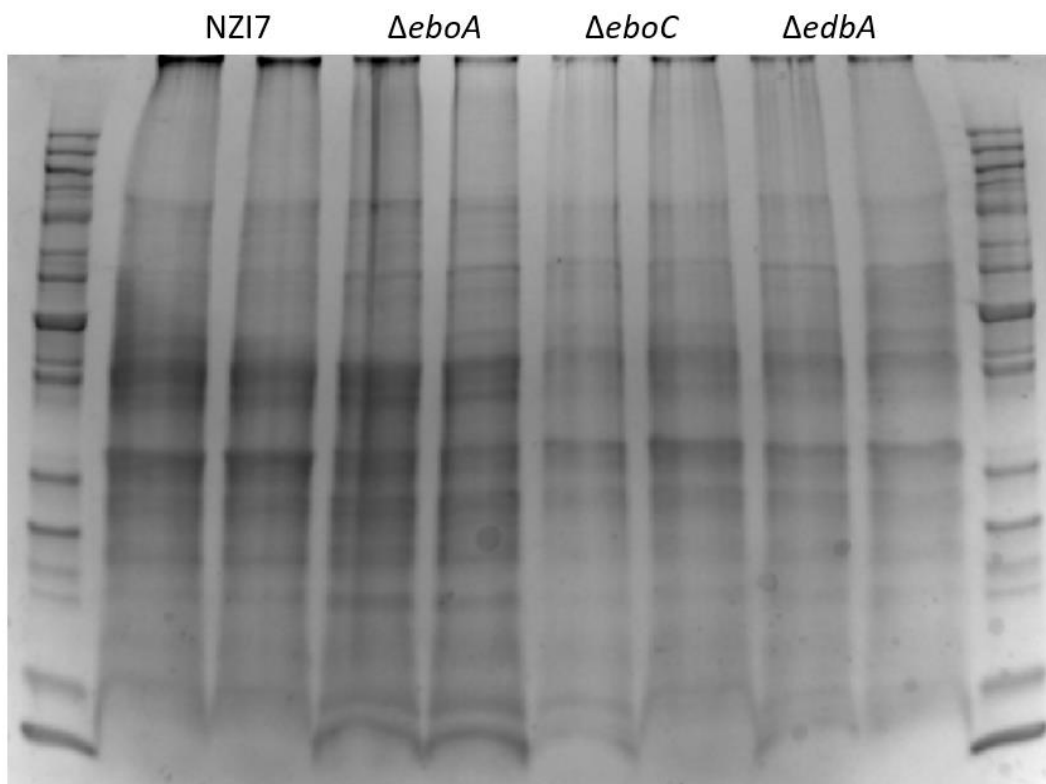


Figure 6: SDS PAGE gel of wild-type (NZI7), *edbA*, *eboC*, and *eboA* mutants. Each strain was run in duplicate. No discernible unique peptide bands were evident between mutants and wild-type.

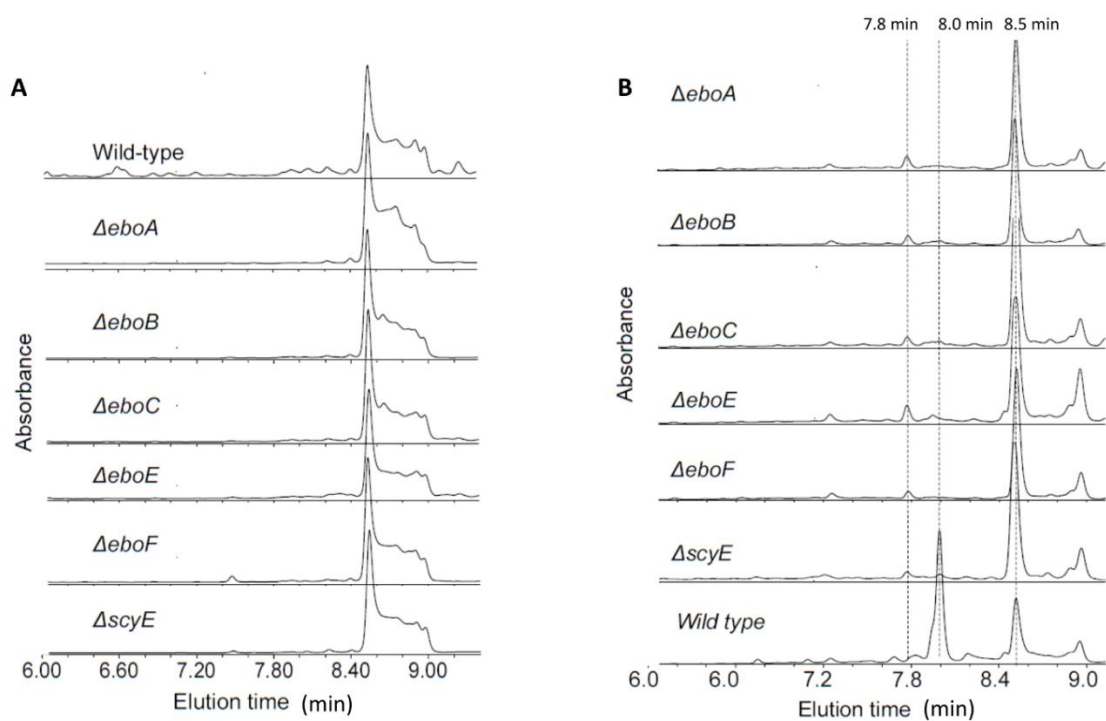
Collectively the analytical methods employed to potentially isolate and identify the putative anti-nematodal compound synthesized by the *edb* gene cluster yielded no clear candidate compound to investigate. These inconclusive results may be due to either the putative compound having characteristics that would prevent its detection by the methods utilized, or perhaps that it is present in such a low abundance compared to the bulk of cellular material that more sensitive instrumentation may be required. Further extraction techniques and broader analytical methods may be required to isolate and characterize the putative product of the *edb* gene cluster.

References

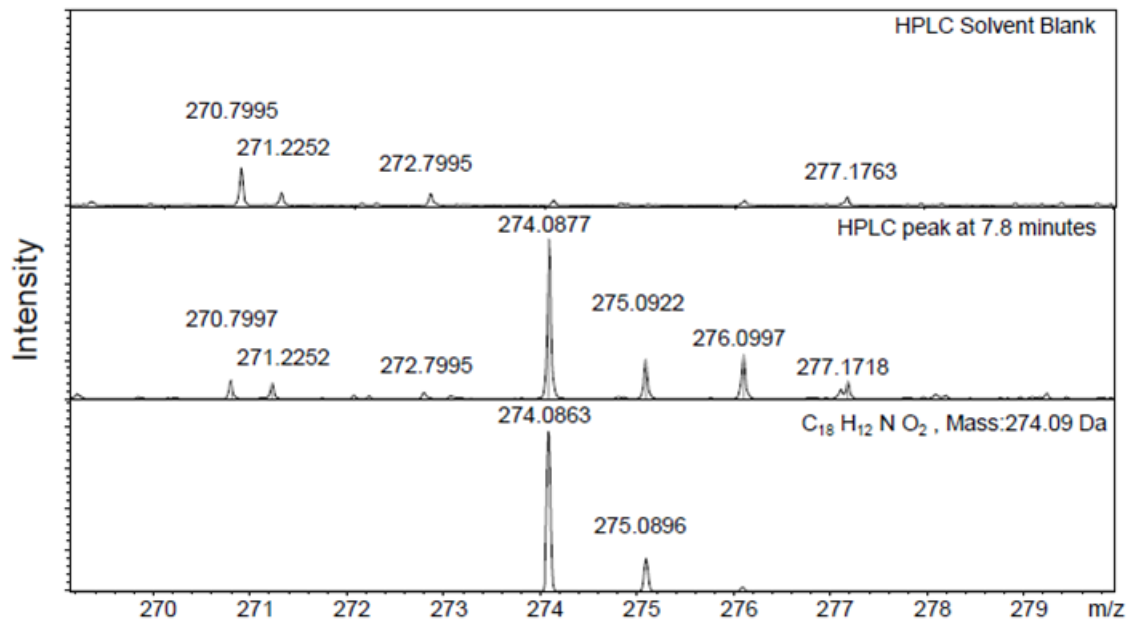
1. Yurchenko T, Ševčíková T, Strnad H, Butenko A, Eliáš M. 2016. The plastid genome of some eustigmatophyte algae harbours a bacteria-derived six-gene cluster for biosynthesis of a novel secondary metabolite. *Open Biol* 6.
2. Klicki K, Ferreira D, Hamill D, Dirks B, Mitchell N, Garcia-Pichel F. 2018. The widely conserved ebo cluster is involved in precursor transport to the periplasm during scytonemin synthesis in *Nostoc punctiforme*. *MBio* 9:e02266-18.
3. Burlinson P, Studholme D, Cambray-Young J, Heavens D, Rathjen J, Hodgkin J, Preston GM. 2013. *Pseudomonas fluorescens* NZI7 repels grazing by *C. elegans*, a natural predator. *ISME J* 7:1126–1138.
4. Bauernfeind A, Burrows JR. 1978. Suggested Procedure Allowing Use of Plastic Petri Dishes in Bacteriocin Typing. *Appl Environ Microbiol* 35:970.
5. Gao Q, Garcia-Pichel F. 2011. An ATP-Grasp Ligase Involved in the Last Biosynthetic Step of the Iminomycosporine Shinorine in *Nostoc punctiforme* ATCC 29133. *J Bacteriol* 193:5923–5928.
6. Jöhler S, Stephan R, Althaus D, Ehling-Schulz M, Grunert T. 2016. High-resolution subtyping of *Staphylococcus aureus* strains by means of Fourier-transform infrared spectroscopy. *Syst Appl Microbiol* 39:189–194.

APPENDIX B

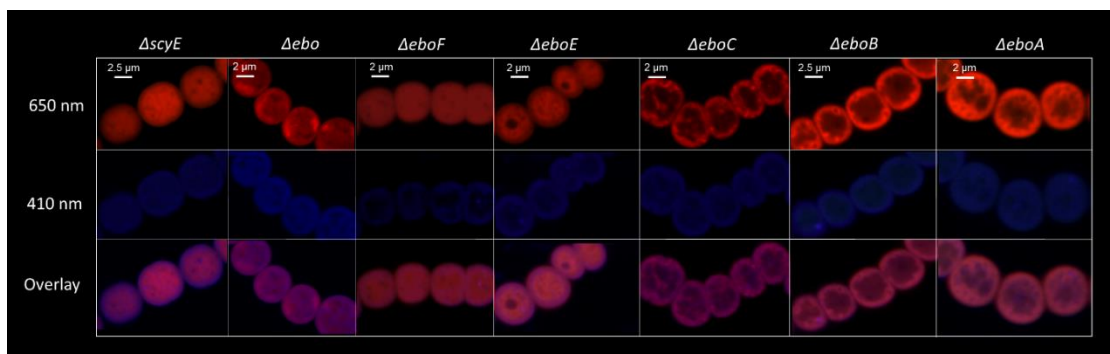
SUPPLEMENTARY MATERIALS FOR THE WIDELY CONSERVED *EBO*
CLUSTER IS INVOLVED IN PRECURSOR TRANSPORT TO THE PERIPLASM
DURING SCYTONEMIN SYNTHESIS IN *NOSTOC PUNCTIFORME*



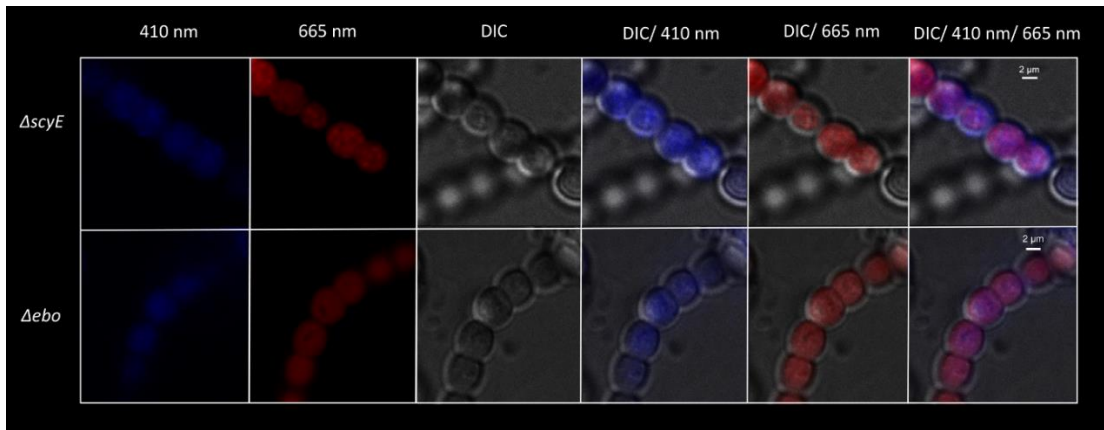
Supplementary Fig. S1: (A) HPLC chromatograms of *ebo* and *scyE* deletion mutant acetone extracts, after UVA induction, exhibiting novel compound at 7.8 minutes. (B) HPLC chromatograms of acetone extracts from uninduced cells of *ebo* and *scyE* deletion mutants.



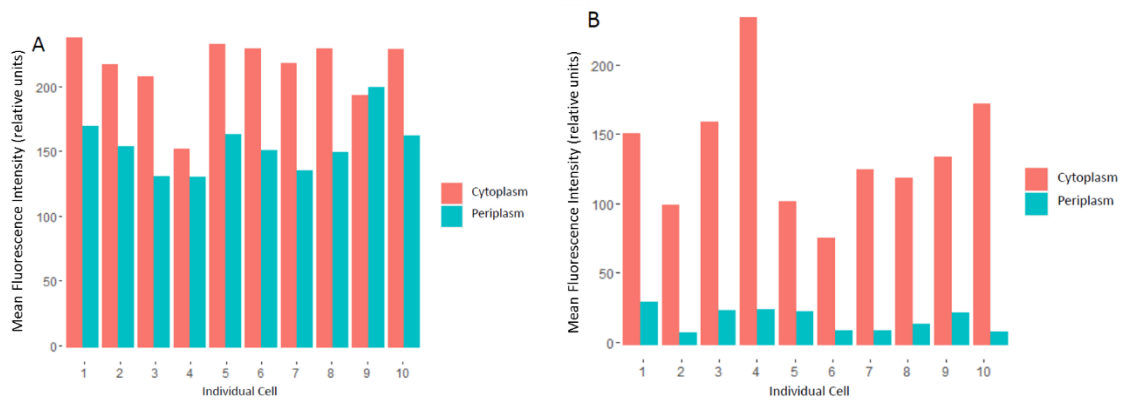
Supplementary Fig. S2: Electrospray ionization time-of-flight mass spectrometry analysis in negative ion mode of blank HPLC solvent (top), Δ eboC compound eluting at 7.8 minutes in HPLC, after eluent collection (middle). Prediction of mass spectral signal, given chemical formula of the scytonemin monomer negative ion fragment (bottom).



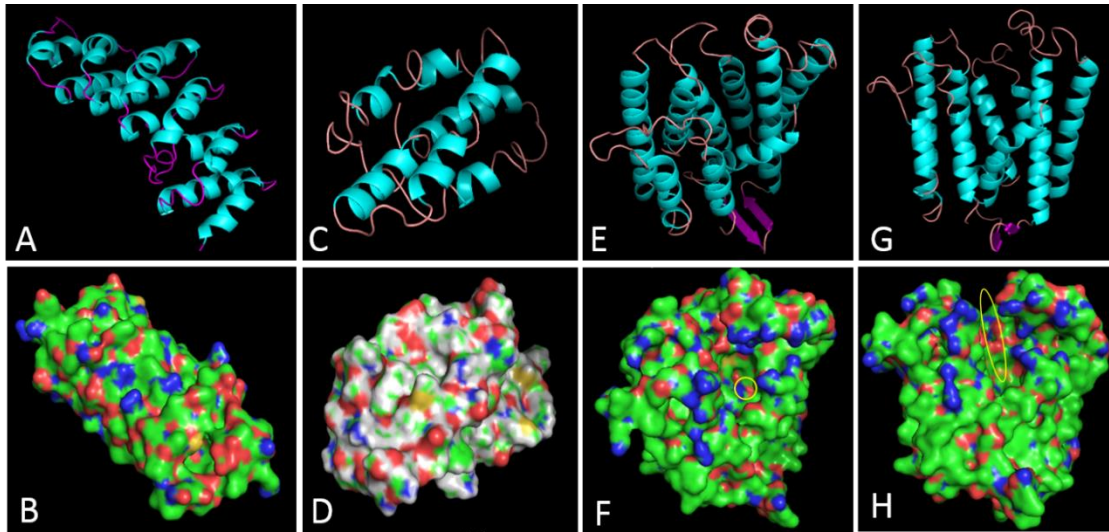
Supplementary Fig. S3: Intracellular localization of the scytonemin monomer in UVA induced *ebo* mutants by overlay of 665 nm emission over the 410 nm emission images. Accumulation was intracellular in all cases.



Supplementary Fig. S4: Overlay images of 665 nm and 410 nm emission channels and DIC channel. In all mutants fluorescence in the 410 nm channel was restricted to the total cellular diameter indicating that it remained intracellular in all cases.



Supplementary Fig. S5: (A) Fluorescence intensity of the scytonemin monomer in cytoplasm and periplasm of *ΔscyE* cells. Periplasmic levels of accumulation of scytonemin monomer are comparable to those reached in the cytoplasm. (B) Fluorescence intensity of scytonemin monomer in cytoplasm and periplasm in *Δebo* cells. There is no significant accumulation of the scytonemin monomer in the periplasm. Fluorescence scattering or leakage may account for some low level signal found outside of the cytoplasm.



Supplementary Fig. S6: I-TASSER (J Yang *et. al.* 2015) predicted structures of EboA and EboC. (A) Cartoon representation of EboA showing putative repeating alpha-helix motif consistent with the structure of solved TPR domain containing peptides. (B) Surface rendering of EboA. (C) Cartoon representation of EboG showing putative repeating alpha-helix motif, similar in structure to EboA. (D) Surface rendering of EboG. (E) Cartoon representation of EboC showing putative transmembrane domains formed by six alpha helices in a cylindrical organization. (F) Surface rendering of EboC showing the phenolic substrate accepting pocket containing basic residues, highlighted with the yellow circle. (G) Predicted structure of EboC shown in cartoon format angled towards the prenyl donor substrate accepting pocket (H) Surface rendering of EboC predicted structure, with prenyl acceptor pocket highlighted with yellow oval.

Supplementary Table S1: Primers used in this work for mutant construction and validation.

Table S1 A. *N. punctiforme* strains and oligonucleotides used in this work.

<i>N. punctiforme</i> strains	Relevant characteristics and details of mutants construction	Oligonucleotides used to confirm homozygosity of mutants
WT	UCD 153: ATCC 29133 wild-type derivate, scytonemin producer (Campbell <i>et al.</i> , 2007)	
Δ scyE	scytoneminless mutant (Ferreira and Garcia-Pichel, 2016)	
Δ ebo	Npun_F5232 to Npun_F5236 knockout. Primers used to create a 5364 bp deletion: sat-XhoI: ATCATCCTCGAGCGACAAATCCACCTGGTGGACTTC satI1: ACGCGGTACAGCACTAAATGCC satI2: GGCATTTAGTGCTGTACCGGTGGCTCTCACGGTAATATCACAC satR: CCCGACTTCTCAAGAAATCGGG	satF: AAGTTGGGCAACTTCTCCTG satR2: GCTGGATAAGGGTTACGTAG satF2: GGGTAAACAAGCCTACCATC F5234R satF3: GCCAAACTGGGCAGATGAAAGC F5232R
Δ eboA	Npun_F5232 knockout. Primers used to create a 606 bp deletion: sat-XhoI satI1 F5232I2: GGCATTTAGTGCTGTACCGGTGGTCCGAGTCTCCTTTACC F5232R: GCCGATGGAATCCACCATAAGG	satF F5233I1 satR2 F5232F: CTGTAGGGCGGTTTCCGGATAG
Δ eboB	Npun_F5233 knockout. Primers used to create a 720 bp deletion: sat-XhoI F5233I1: CGTAAGTAGTACGGGAACAC F5233I2: GTGTTCCCGTACTACTACGGCCAAACTGGCAGATGAAAGC F5233R: ACGCGCAAAACACAGATTGTTG	F5232F F5234I1 F5233R2: AAGCGTCCCAGCCTACAAGAC F5235L-XhoI
Δ eboC	Npun_F5234 knockout. Primers used to create a 657 bp deletion: F5234L-XhoI: ATCGCGCTCGAGAAGTTGGCAACTTCTCCTG F5234I1: AGCCGGAAGCAGCGAAACCTAC F5234I2: GTAGGTTTCGTGCTTCCGGCTTAGCCGAAATATCCGAAATGCG F5234R: CAGCCTCTACTGCCTTTC	satF3 F5232R F5234R2: CCGCCGCCGTATAAACCACTTG F5234F: CGATCGCAGCACTACCATTTCG
Δ eboE	Npun_F5235 knockout. Primers used to create a 768 bp deletion: F5235L-XhoI: ATCGCGCTCGAGGCTAGTCAGTTTGGCATCCAACAC F5232R F5235I2: CCTTATGGTGGATTCCATCGGCCTACCCAAAGATGACATCGCTACTG F5235R: CTGAATACACATCCGCCACTTAGG	F5234F F5236I1 F5233R F5235F: ACGTCGCACTGATGGTACACTG
Δ eboF	Npun_F5236 knockout. Primers used to create a 1131 bp deletion: F5236L-XhoI: ATCGCGCTCGAGGGCGGATATTCTTGTAGGTTTTCG F5236I1: ATTGACGCTACCTGTCTTTCG F5236I2: CGAAAGGACAGGTAGCGTCAATTGGCTCTCACGGTAATATCAC satR	F5235F F5234R F5236R2: GCGACAATCCCATGTTTCATC satF2

Underlined letters indicate introduced XhoI restriction endonuclease cleavage site.

Table S1 B. *E. coli* strains and oligonucleotides used in this work

<i>E. coli</i> strains	Relevant characteristics and details of mutant construction	Oligonucleotides used to confirm mutants
<i>E. coli</i> harboring pRSFDuet:scyABC	Forward primer containing NcoI restriction site: 5'-GCCGCCCATATGACTATACTGGTTCC-3' Reverse primer containing SacI restriction site: 5'-TTAGTTGGGAAGTCTAGGGATTCTTG-3'	5'- GCCGCCCATATGACTATACTGGTTCC- 3'
<i>E. coli</i> harboring pRSFDuet:scyABC-scyEF	Forward primer containing NdeI restriction site 5'-TGGGATTCGTTGTCTTTAAACCC-3' Reverse primer containing FseI restriction site 5'-TCAGCATTTGCTTTTGCAGTTCTTTC-3'	5'- TGGGATTCGTTGTCTTTAAACCC- 3'

Supplementary Table S2: Primers used in this work for RT-PCR assessment of polar transcriptional effects.

Supplementary Table S2

<i>ebo</i> cDNA target	Forward primer	Reverse primer
<i>eboB</i>	GCTAGTCAGTTTGGCATCCAACAC	GCTTCTGTCATGTCGTCATAGCC
<i>eboC</i>	CCGCCGCCGTATAAACCACTTG	ATCGCGCTCGAGGCTTGTAGGTTTCG
<i>eboE</i>	CCATATTGATTTAGAACCTGAGCCTGATGG	ACGCGCAAACACAGATTGTGG
<i>eboF</i>	GGGTAACAAGCCTACCATC	GGAAATTTTATTTGCGGATCGAGG

Supplementary Materials and Methods

Construction of mutants. PCR products were mixed and allowed to anneal via overlapping sequences. Subsequent PCR yielded a product containing the deletion, which was then cloned into the *sacB*-containing plasmid pRL278 (1) as a XhoI fragment using restriction sites contained within the primers. The resulting recombinant plasmids were sequenced to ensure fidelity. Gene deletion suicide-plasmids were introduced into wild-type *N. punctiforme* via triparental conjugation using *E. coli* strains as carriers of recombinant plasmids according to (2), with two variations from the original protocol: no sonication of *N. punctiforme* prior to conjugation, and no supplementation of conjugative plates with CO₂-enriched air. Single recombinant strains, in which the suicide-plasmid was integrated into the genome, were selected and maintained in medium containing neomycin. Strains in which the suicide-plasmid and wild-type gene were eliminated via a second recombination event, were selected by the presence of 5% (w/v) sucrose. PCRs performed as previously described (3) confirmed elimination of the wild-type gene and in-frame replacement with the respective 3' and 5' ends of the inactivated gene

(see supplementary Table S1 for oligonucleotide details). In addition to new mutants obtained in this work, we also analyzed deletion mutants $\Delta scyD$, $\Delta scyE$, and $\Delta scyF$, described previously (3).

Biochemical characterization of mutant strains. Cells from *N. punctiforme* wild-type and derived deletion mutants were tested for their ability to produce scytonemin upon induction by UVA radiation. Cultures grown for five days, with a chlorophyll *a* content of approximately 1–2 $\mu\text{g ml}^{-1}$, were concentrated to 50 μg of chlorophyll *a ml}^{-1} and spread over polycarbonate membrane filters (0.4 μm pore size), which were placed floating on liquid Allen and Arnon medium filled glass petri dishes, as previously described (4). The cells were exposed to white fluorescent light (7 Wm^{-2}) supplemented with UVA radiation (10 Wm^{-2}) as described previously (3), continuously for five days. Following UVA exposure, the cells were harvested and the lipid-soluble pigments were extracted in equal volumes of 100% acetone. Extracts were initially analyzed spectrophotometrically between 330 nm to 730 nm, a strong absorption peak at 384 nm indicating that scytonemin had accumulated in the cells (4). Following UVA exposure, water-soluble compounds were also extracted from whole cells in equal volumes of 25% aqueous methanol. 50 μl of concentrated acetone or methanol extracts from cells exposed to UVA radiation were also analyzed by high pressure liquid chromatography (HPLC on a Waters e2695 equipped with a Supelco Discovery HS F5-5 column connected to a Waters 2998 PDA UV-Vis diode array detector using previously described conditions (5,6) with the following revisions (6): flow rate was 1.0 mL /min-1 with linear gradients of 10 mM ammonium formate buffer (pH 3 adjusted with hydrochloric acid) and acetonitrile, (0-1.5 minutes 0-25% acetonitrile, 1.5-7.5 minutes 25-100% acetonitrile, 7.5-12.5 minutes 100-100% acetonitrile, 12.5-27.5 minutes 100-25% acetonitrile). Carotenoids, chlorophyll *a*, and scytonemin were monitored in the chromatograms at 384 nm, but spectra were recorded continuously between 200-800 nm using the*

PDA detector. Individual compounds were identified by their characteristic absorption maxima and appropriate retention time.

To obtain an authentic standard of the scytonemin monomer (Compound 1 in Figure 1) the results of Malla and Sommer (6) were repeated by cloning genes *scyA*, *scyB*, and *scyC* into the *NcoI/SacI* site of the pRSF-Duet-1 vector to produce the pRSF-ScyABC recombinant plasmid. Genes *scyE* and *scyF* were then cloned into the *NdeI/FseI* site of pRSF-ScyABC to produce the pRSF-ScyABC-ScyEF recombinant plasmid. Recombinant strains were cultured, induced, and extracted as published (6).

Supplementary References

1. Cai YP, Wolk CP. 1990. Use of a conditionally lethal gene in *Anabaena* sp. strain PCC 7120 to select for double recombinants and to entrap insertion sequences. *J. Bacteriol* 172:3138–3145.
2. Cohen MF, Wallis JG, Campbell EL, Meeks JC. 1994. Transposon mutagenesis of *Nostoc* sp. strain ATCC 29133, a filamentous cyanobacterium with multiple cellular differentiation alternatives. *Microbiology* 140:3233–3240.
3. Ferreira D, Garcia-Pichel F. 2016. Mutational Studies of Putative Biosynthetic Genes for the Cyanobacterial Sunscreen Scytonemin in *Nostoc punctiforme* ATCC 29133. *Front Microbiol* 7(May):1–10.
4. Garcia-Pichel F, Castenholz RW. 1991. Characterization and biological implications of scytonemin, a cyanobacterial sheath pigment. *J Phycol* 27:395-409.
5. Karsten U, Garcia-Pichel F. 1996. Carotenoids and Mycosporine-like Amino Acid Compounds in Members of the Genus *Microcoleus* (Cyanobacteria): A Chemosystematic Study. *Syst Appl Microbiol* 19(3):285–294.
6. Malla S, Sommer MOA. 2014. A sustainable route to produce the scytonemin precursor using *Escherichia coli*. *Green Chem* 16:3255–3265.

APPENDIX C

SUPPLEMENTARY MATERIALS FOR THE FUNCTIONAL LANDSCAPE AND
NEIGHBOR DETERMINING GENOMIC REGION SEARCH (FLANDERS)
BIOINFORMATIC DATA MINING PACKAGE/PIPELINE AND ITS USE TO PROBE
FUNCTIONS OF THE *EBO* GENE CLUSTER AMONG BACTERIA

Supplementary table 1: Complete results of the present work. In addition to the contents of table 1, the following columns are listed: Annotations of upstream and downstream *ebo* adjacent genes are listed in columns 6 and 7 respectively. Column 8 contains the total number of ORFs in each genome. Column 9 contains the total number of signal peptide containing ORFs in each genome, and column 10 contains the ratio of signal peptide containing ORFs to total ORFs in each genome. From these ratios the probability of finding the exact number of signal peptides found in the 8 *ebo* neighboring ORFs (column 3) in 8 randomly selected ORFs from each genome are listed in column 12. Column 13 contains the total number of transport related genes, and Column 14 contains the ratio of transport related genes to total genes in each genome. From these ratios the probability of finding the exact number of transporters found in the 8 *ebo* neighboring ORFs (column 15) in 8 randomly selected ORFs from each genome are listed in column 16. Supplementary table 1 can be accessed at <https://drive.google.com/drive/folders/1n9gC5HmX6k5CMOoccbFL78s9StTopI7k?usp=sharing>.

APPENDIX D

SUPPLEMENTARY MATERIALS FOR BET HEDGING SUNSCREEN
PRODUCTION AND MOTILITY RESPONSES AGAINST UV EXPOSURE IN A
CYANOBACTERIUM

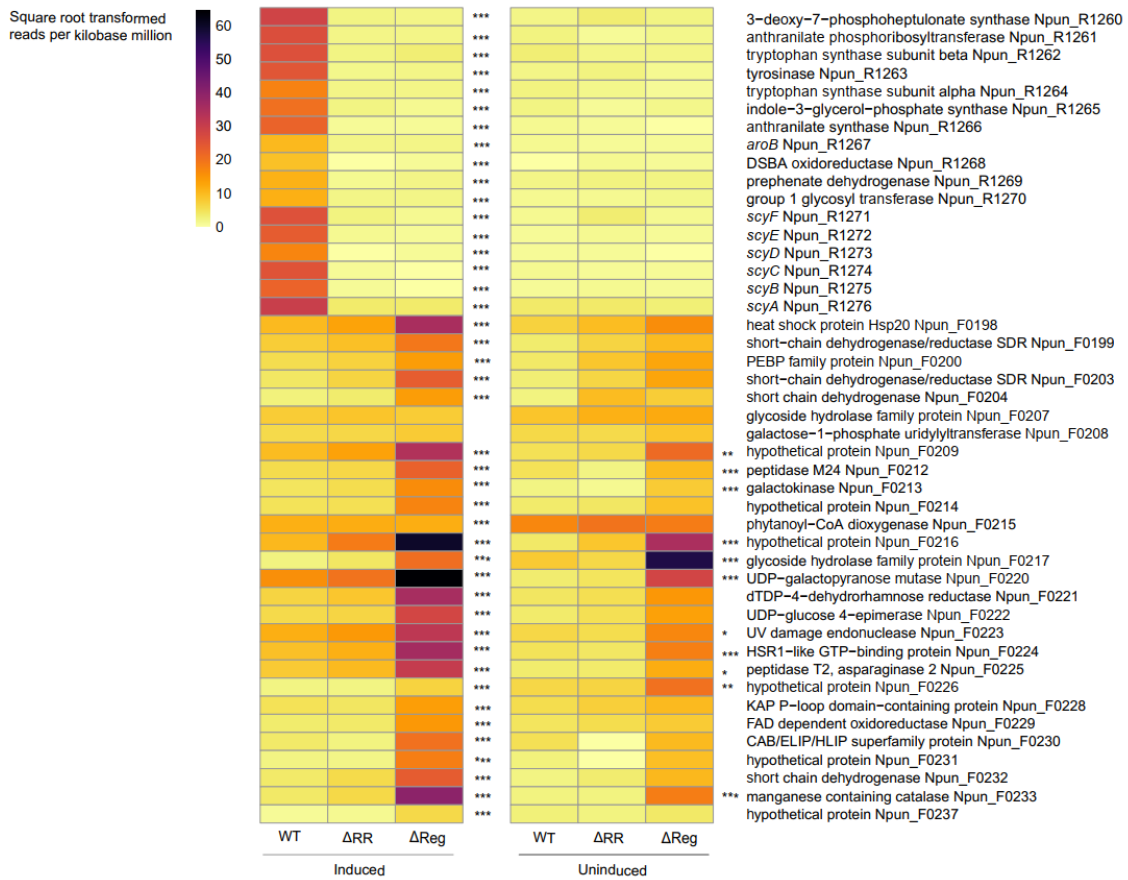
Supplementary Table 1: PCR primers utilized in this work

Gene target	Forward primer sequence	Reverse primer sequence
<i>rnpB</i>	TAAGAGCGCACCAGCAGTAT	CATTGAGCGGAACTGGTAAA
<i>gvpA</i>	TGACCGCATCTTGGACAAAGG	TGTTAAACCAACCGCCTCAGC
<i>pilM</i>	CTAACGGATGAACTGCGCCG	TGGCAAGCTCAATCGTTGGG
Npun_ F1278- Npun_ F1277- 3' (Scy TCR)	GTGGACGGATCCCAAAGGCTGG CCATTGC	GGTCAGGAGCTCATCCCTACT TGCTCC
Npun_ F1277- 3' (Scy RR)	GTGGAGGATCCGTCTATTTTTTTG TATTCCTC	GGTCAGGGAGCTCCATAACT TGTAGCAG
Npun_ F1278- 1277 and Npun_ F1277- 5' (Scy TCR)	TGTGGATCCACCGTGAAACTTTG GTAAAAGTCG	CTTGAGAGCTCCCGGATTGCT ATAGTAG
Npun_ F1684- Two Hybrid (<i>hcyC</i>)	ATATATCTAGATGTGAAAAATA AAATTTATCTAAAAGTCAAC	ATATAGGTACCCTTAGATAAG AGGATTAATAGTATTTAC
Npun_ F1682- 5' (<i>hcyA</i>)	ATATAGGATCCCCCTACAAATG ATTCATATTC	CTTACTGTGTAATGCTCATGG TTTTATCCTATTG
Npun_ F1682- 3' (<i>hcyA</i>)	CCATGAGCATTACACAGTAAGTT CCTAGACTAG	ATATAGAGCTCTGAAAGCGCG GGTAAGAGC
Npun_ F1684- 5' (<i>hcyC</i>)	ATATAGGATCCTCCTCGTCGATC CAATAATG	GAGGATTAATAGTAAAGGTCT GTGTTGACTTTTAG

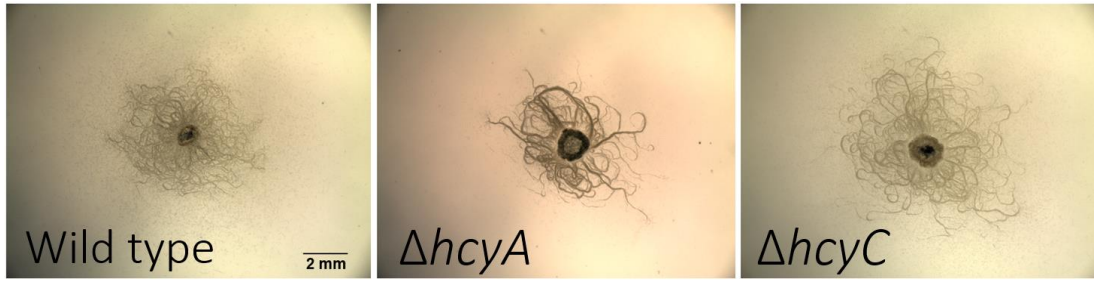
Npun_ F1684- 3'(hcyC)	CACAGACCTTTACTATTAATCCT CTTATCTAAGTTTTG	ATATAGAGCTCGCGGGAATG GACGACTATC
---------------------------------	--	------------------------------------

Supplementary Table 2: Plasmids utilized in this work

Plasmid	Description	Source
pDDR415	pRL278:: Δ NpF1682 (BamHI-SacI)	This study
pDDR417	pRL278:: Δ NpF1684 (BamHI-SacI)	This study
pDDR437	pKT25::NpF1684 (XbaI- KpnI)	This study
pDDR435	pUT18c::NpF1682 S49A (XbaI-KpnI)	Riley et al., Mol Micro 2018 (68)
pDDR415	pRL278:: Δ NpF1682 (BamHI-SacI)	Riley et al., Mol Micro 2018 (68)
pDPF1278	pRL278:: Δ NpF1278-1277 (BamHI-SacI)	This study
pDPF1277	pRL278:: Δ NpF1278-1277 (BamHI-SacI)	This study



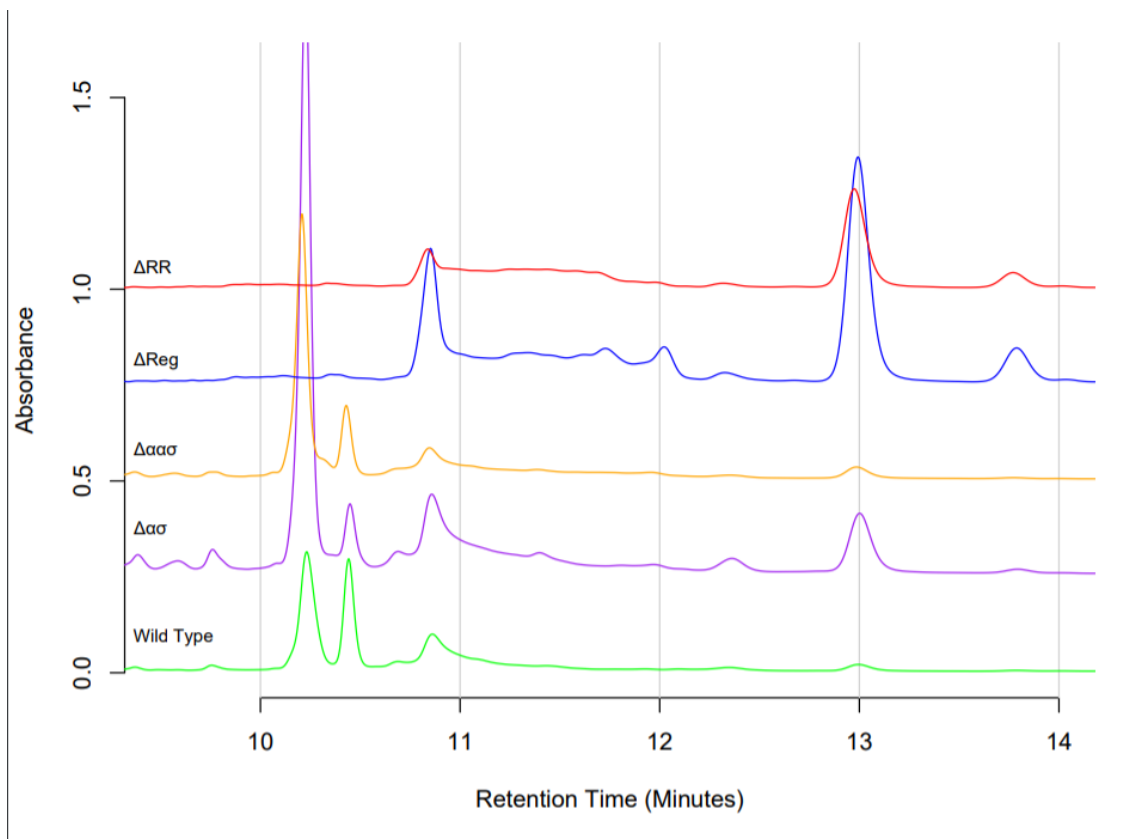
Supplementary Fig. 1: Genes upregulated in Δ Reg compared to Wild type (WT) and Δ RR under scytonemin inductive (UV-A irradiation) and non-inductive (while light only). Results for the scytonemin operon included for comparison at the top. The gene cluster from Npun_F0198 to Npun_F0237 contains several general stress response genes and seems to be upregulated in the absence of the sensor histidine kinase component of the scytonemin TCR.



Supplementary Figure 2: Plate spreading motility assay of Wild type, $\Delta hcyA$, and $\Delta hcyC$. No significant difference in motility was observed under standard phototactic conditions between Wild type and mutant strains.



Supplementary Figure 3: Bacterial Two-Hybrid interaction assay between HcyA and HcyC indicating protein-protein interaction between HcyA and HcyC. Positive and negative controls were conducted as in Karimova and Ullman, 1998 (46).



Supplementary Fig. 4: HPLC chromatograms at 384 nm to assess scytonemin production capability of mutant strains used in this study. Wild type, $\Delta hcyC$, and $\Delta hcyA$ showed a clear peak at 10.2 minutes corresponding to scytonemin production while ΔReg and ΔRR did not.

Materials availability

Mutant strains, natural variants and plasmids used in this work are freely available directly from the lead contact upon request. The wild-type strain is also available from the American Type Culture Collection. There are no newly developed reagents associated with this work.

Data and code availability

Sequence datasets generated in this study are available at NCBI's SRA database with accession number PRJNA728749. Other phenotypic or analytical data are available in the manuscript in graphical or photographic form, and source data, when applicable, is available on request from the corresponding author.