Published in final edited form as:

Nat Prod Rep. 2018 April 25; 35(4): 357–378. doi:10.1039/c7np00053g.

Parallel lives of symbionts and hosts: Chemical mutualism in marine animals

Maho Morita and Eric W. Schmidt

Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah, USA 84112

Abstract

Symbiotic microbes interact with animals, often by producing natural products (specialized metabolites; secondary metabolites) that exert a biological role. A major goal is to determine which microbes produce biologically important compounds, a deceptively challenging task that often rests on correlative results, rather than hypothesis testing. Here, we examine the challenges and successes from the perspective of marine animal-bacterial mutualisms. These animals have historically provided a useful model because of their technical accessibility. By comparing biological systems, we suggest a common framework for establishing chemical interactions between animals and microbes.

Graphical abstract

1. Introduction

Chemistry mediates the interactions between animals and microbes (largely, bacteria). While it is straightforward to obtain correlative data, it remains a great challenge to demonstrate the individual bacterial strains and chemical compounds that are relevant to animal biology. This is an important problem both because animal wellbeing is directly impacted by the microbiota and symbiosis,^{1–4} and because specialized metabolites resulting from these interactions are often good drug leads.⁵

The challenge is that it is difficult to show that a compound, found in an animal, is actually produced by a microbial strain living within that animal. Other alternatives include a dietary source, host-derived biosynthesis, multiple symbiotic microbes making the same compound, collaborative biosynthesis, and the host and microbes both making the same compounds, among others.⁶ Each of these alternatives has seen experimental evidence in specific systems. It is obvious that just because a microbe isolated from an animal produces a compound in culture, it is not necessarily the biologically relevant producer in nature. Further, it is a challenge to establish a biological function for individual compounds

produced in microbe-host interactions. For example, bacteria are often isolated from animals and shown to produce certain chemicals in culture. However, whether those same strains produce the same compounds within the animal, or whether those strains or chemicals even exist within the animal at all, are often unknown.

Correlative studies include high-throughput omics methods and cultivation efforts that lead to the initial identification of consistent co-associations between compounds, bacteria, phenotypes, gene sequences, and / or hosts. They are excellent at providing essential data for further testing. However, because of many confounding factors, further specific hypothesis-testing experiments tying a symbiont to a molecule or a phenotype need to be performed. One of the main limitations holding back the field is a lack of commonly accepted types of experimental data to demonstrate biologically relevant chemical interactions. In turn, this leads to widespread confusion as to what has or has not been established. Further complicating this is a lack of commonly accepted use of terms (see Section 6).

There are many excellent reviews that either focus on or touch upon symbiotic sources of natural products.⁶⁻⁹ Here, we focus on experimental evidence for chemical mutualism in marine animals. These organisms are rich with information because of the historical importance of marine natural products, and the commonly noted supply problem, which have led to numerous investigations into the source and biosynthesis of animal compounds. Following these examples, and based upon previous successes and challenges in the field, we propose a common framework for establishing the origins and roles of natural products in symbioses.

2. Experimentally defined defensive mutualisms

In some cases, good evidence demonstrates that compounds are 1) produced by bacteria and 2) defensive for the host animal. Since the bacteria derive nutrition from the host, these are experimentally defined defensive mutualisms, usually with extensive evidence obtained by multiple research groups.

2.1 Protection from UV light: The MAAs

In high-UV tropical environments, several marine organisms employ mycosporine-like amino acids (MAAs) as UV protectants. Among marine natural products, MAAs such as shinorine (**1**) and mycosporine glycine (**2**) have among the best evidence for the defensive role. Prior to an elucidation of the chemical structures,¹⁰ MAAs were observed as “UV-absorbing substances” in extracts of several corals.¹¹ Early work used correlative studies to understand the effects of MAAs. In corals for example, there appeared to be correlation between UV-exposure levels and MAAs concentrations.^{12, 13}

Beyond correlation, Jokiel & York compared levels of UV in dinoflagellates from corals that were raised in isolation or in the animals.¹³ Purified dinoflagellates were damaged by long-term UV exposure, while dinoflagellates grown in coral were not, implying that corals protect their symbionts from UV radiation using MAAs. This hypothesis was reinforced by another example, where cyanobacterial *Prochloron* sp. form photosynthetic symbioses with tunicates. Dionisio-sese showed that photosynthesis in purified *Prochloron* cells were

severely inhibited by UV irradiation, while *Prochloron* in the host tunicate was not.¹⁴ It is noteworthy that these early studies were carried out in the natural environment or in conditions close to nature, providing good evidence for the defensive role of the natural products.

In terms of the origin of MAAs, tunicate-*Prochloron* associations provided some early data. Several studies analyzed MAA localization in photosymbiotic tunicates, mainly based on chemical analyses of dissected tunicate tissues and isolated *Prochloron*, or UV-light microscopy.^{14–17} In *Lissoclinum patella*, for example, the surface layer contained all of the *Prochloron* cells and a higher concentration of MAAs than the basal layer.¹⁴ While in *L. patella*, *Prochloron* contained less MAA than the host, in other tunicates MAAs were found only in *Prochloron*.¹⁷ Although MAA origin was still unknown, it was likely that MAAs are produced by the symbionts, transported to the host tunic, and/or concentrated on the tunic surface. The MAAs may be needed in tunicates because the animals are often exposed to high levels of UV irradiation at low tide, even sometimes being above the waterline.

Since MAAs were also isolated from free-living algae,^{18, 19} MAAs were believed to be biosynthesized through the shikimate pathway, which animals typically lack.²⁰ The first evidence that MAA biosynthesis required the shikimate pathway was provided by Shick *et al.*²¹ In a coral-dinoflagellate association, production of MAAs was severely inhibited by glyphosate, a specific inhibitor of the shikimate pathway. Biochemical evidence for bacterial MAA biosynthesis was provided in 2010.²² This study identified and heterologously expressed the shinorine biosynthetic gene cluster from the cyanobacterium *Anabaena variabilis*. The cluster encoded four proteins, including dehydroquinate synthase (DHQS), an *O*-methyltransferase, an ATP-grasp enzyme, and a nonribosomal peptide synthetase homolog that collectively synthesize **1** from sedoheptulose 7-phosphate (**3**) via 4-deoxygadusol (**4**) (Scheme 1). Thus, MAAs are not made via shikimate, but instead via a related DHQS enzyme that is part of the larger family of sedoheptulose 7-phosphate modifying enzymes including DHQS.²³ Subsequently, many bacterial MAA pathways were identified, including one found in *Prochloron* living within a tunicate,²⁴ further reinforcing Hirose's hypothesis about the symbiotic origin of MAAs in those animals.

Beyond bacterial and algal MAAs, genomic studies revealed the presence of MAA genes in invertebrates and vertebrates. Several corals and sea anemones contain MAAs, despite the absence of MAA-producing symbionts or diet.^{25–28} In 2008, sequence analyses strongly suggested the presence of a shikimate pathway in a sea anemone *Nematostella vectensis*.²⁹ The anemone had two genes encoding a DHQS-like enzyme and an *O*-methyltransferase, which might putatively be responsible for MAAs biosynthesis. The two gene sequences were highly similar to those of the dinoflagellates. This might suggest a recent horizontal gene transfer from a symbiotic dinoflagellate to a host cnidarian. At that time, it was still unclear that these genes were truly related to MAA biosynthesis, as MAA biosynthesis had not yet been elucidated.

More recently, genomic analysis of a coral *Acropora digitifera* indicated that the animal has the genes for producing MAAs *de novo*.³⁰ In the whole genome, there were homologues of all four enzymes of the cyanobacterial shinorine biosynthetic cluster.²² The coral seemed to

lack an enzyme essential for cysteine biosynthesis, implying that its symbiont might provide nutritional metabolites to the coral host in return for UV protection by MAAs. Overall, it is clear that in some cases, corals make their own MAAs, while in other cases MAAs may be made by symbiotic dinoflagellates.

In addition to invertebrates, vertebrates such as marine fishes also have been known to harbor MAA-related UV protective compounds, the gadusols.^{10, 31–33} Gadusols were initially considered to be accumulated from the diet or provided by symbiotic bacteria. Surprisingly, however, a recent study found that fish have the ability to biosynthesize gadusols independently.³⁴ Vertebrates including amphibians, reptiles, and birds also have homologous pathways, although humans and other mammals lack two essential genes for gadusol biosyntheses. Furthermore, the gadusol pathway from fish was experimentally characterized by heterologous expression in a yeast strain. The gadusol-containing cell lysate showed UV-protection effects on both UVB-sensitive and wild-type yeast strains.

Taken together, biosynthetic pathways of MAAs and related compounds are widespread among bacteria, fungi (not described here), algae, invertebrates, and vertebrates, suggesting the widespread importance of UV protection for organisms. The examples described above illustrate the complexity of chemical mutualism and the difficulty of making simplifying assumptions about who makes the molecules and who benefits from them. This is especially complicated in the MAA scenario, where both horizontal gene transfer and convergence play a role in the generation of UV protective substances.

2.2 Defense against infection: Crustaceans and their epibionts

The association between crustaceans and epibiotic bacteria is one of the earliest examples of studying marine defensive chemical symbioses. Since these studies took place prior to the genomic era, they relied upon electron microscopy, cultivation of isolated microbes, and chemical analysis. However, the totality of evidence is quite convincing because it meets the classical “Koch’s postulates” standards. In 1989, Gil-Turnes *et al.* observed that a shrimp *Palaemon macrodactylus* harbored an epibiotic bacterium *Alteromonas* species that produces 2,3-indolinedione (isatin, **5**).³⁵ This small molecule showed anti-infective activity against a fungus *Lagenidium callinectes*, a pathogen of many crustaceans. Isatin was also able to restore the unprotected embryos to a defended state. Similarly, in the lobster *Homarus americanus*, another epibiotic Gram-negative bacterium was found to protect the embryos from *L. callinectes*.³⁶ The defensive compound was identified as 4-hydroxyphenethyl alcohol (tyrosol, **6**), which was previously discovered as a metabolite of plant endophytic fungi that protects against plant pathogens.³⁷

2.3 Defense against predation: The bryostatins

Bryostatins (e.g., bryostatin 1, **7**) are a family of macrolides originally isolated from a marine bryozoan *Bugula neritina* because of their activity in cancer assays.^{38–41} Bryostatins are among the best known of the marine natural products because of their potential uses in a variety of diseases and the many clinical trials spanning decades.^{42–44} However, bryostatins were initially isolated in very small yields, spawning significant efforts both in chemical

synthesis and in understanding their biosynthetic origins for potential biotechnological production.⁴⁵

Based on 16S rRNA-targeting PCR and microscopic analyses, Haygood and Davidson proposed that the γ -proteobacterial symbiont “*Candidatus Endobugula sertula*” is the most likely producer of bryostatins.^{45, 46} Studies focused on bryozoan larvae, which had a simpler microbial community. In situ hybridization indicated that “*Ca. Endobugula sertula*” was found in a ring surrounding the pallial sinus of *B. neritina* larvae, as well as in other tissues in the adults. The presence of these bacteria specifically in bryozoans, in addition to their dominance in the microbiome and their vertical transmission in larvae (where bryostatin concentrations are highest) provided early indirect but compelling correlative evidence that the bacteria produce bryostatins.

Additional microbiological data was obtained, further strengthening this connection. Lim and Haygood identified a symbiont closely related to “*Ca. E. sertula*” from another *Bugula* species *B. simplex*.⁴⁷ Chemical analyses and bioassays using *B. simplex* extracts suggested the presence of bryostatin analogs (although the chemical structures were not determined in this study). The first evidence for the defensive effect of bryostatins was provided by Lopanik *et al.*, in which symbionts were cured from the larvae using antibiotic treatment.⁴⁸ Bryostatins appeared to protect the larvae *B. neritina* from predation, but did not protect the adults. The palatability of *B. neritina* was correlated with the presence and absence of “*Ca. E. sertula*”, supporting the proposed ability of this bacterium to biosynthesize bryostatins and also providing evidence that bryostatins serve a defensive purpose.⁴⁸

The bryostatins and their putative producer “*Ca. E. sertula*” were localized in the host *B. neritina* through its life cycle.⁴⁹ Since protein kinase C (PKC) is the target of bryostatins,^{50, 51} the compounds were detected using fluorescently labeled PKC, while fluorescent *in situ* hybridization (FISH) was used to follow bacteria. The results indicated that bryostatins were found distributed on the exterior of *B. neritina* larvae, not just in “*Ca. E. sertula*” cells. High levels of both bryostatins and symbionts were detected during embryogenesis and metamorphosis, prior to formation of protective chitinous shell. This was proposed to provide a chemical defense while the physical defense was developing.⁴⁹

The biosynthetic gene cluster for bryostatins was discovered by screening a cosmid library.^{52, 53} Bryostatins are formed through the action of a large, multi-protein polyketide synthase (PKS) with an acyl transferase (AT) acting in trans, rather than in cis, during chain elongation. Since the cluster was initially sequenced prior to the discovery of the so-called trans-AT PKSs in cultivated and experimentally tractable bacteria,^{54–56} this slowed down the analysis. The trans-AT PKSs initially did not match the colinearity rules learned from the cis-AT PKSs, and other strange features were apparent, such as the presence of specific pyran synthesizing dehydratase domains. This meant that, initially, the predicted function of the bryostatin PKS could not be convincingly correlated to the structure.

The first characterized cluster consisted of five modular trans-AT PKS genes. Analysis of the bryostatin clusters from two different *Bugula* samples indicated that in one, the cluster was contiguous, while in another it was split into two different loci.⁵⁷ With a better knowledge of

trans-AT PKS clusters, it became obvious that the structures of bryostatins matched closely what would be predicted from the *bry* cluster. Ultimately, experimental functional evidence for the assignment of the gene cluster was obtained using small proteins from the pathway, including trans-AT BryP⁵⁸ and β -branching enzyme BryR.⁵⁹ As these reactions were shown to be promiscuous, and as these types of enzymes are generic to trans-AT PKSs, there remains a missing link between sequence and function.

B. neritina probably acquires “*Ca. E. sertula*” through a vertical mode of transmission, based on the presence of bryostatins in reproductive tissue.⁴⁵ Often, vertical transmission leads to cospeciation of symbiont and host.^{60, 61} Lim-Fong *et al.* sequenced host and symbiont 16S rRNA genes.⁶² Contrary to expectation, phylogenetic trees of host *Bugula* species and “*Ca. Endobugula*” symbionts were different enough to suggest that the host and symbiont have more complex history involving at least some horizontal acquisition.⁶²

Three types of *B. neritina* (Type D, S, and N) were recognized as sibling species. Among them, only Type S and D were thought to harbor the symbiont and bryostatins, but Type N was not.^{63, 64} However, Linneman *et al.* reported the presence of symbiotic Type N and aposymbiotic Type S *B. neritina* hosts, based on PCR targeting and FISH analyses.⁶⁵ This observation left open the possibility that “*Ca. E. sertula*” could transiently survive outside of its host and transmit to aposymbiotic hosts.

More recently, the complete genome of uncultivated “*Ca. E. sertula*” was assembled and analyzed.⁵⁷ Often, the genomes of Gram-negative symbionts undergo a process of genome reduction, where non-essential genes are lost, leaving a very small chance that bacteria can survive outside of their hosts.⁶⁶ “*Ca. E. sertula*” did not have remarkable signs of genome streamlining, such as a small size genome, low GC content, and low coding density.⁵⁷ The genome size (3.3 Mbp) was larger than the 2 Mbp expected based on flow cytometry.⁶⁴ While there were no clear signs of genome reduction, several pieces of genomic evidence hinted at a dependence upon the host. “*Ca. E. sertula*” appeared to lack genes for biosyntheses of several amino acids, which might be host-supplied. Additionally, metatranscriptomics indicated that a gene that might be associated with symbiosis was highly expressed in “*Ca. Endobugula sertula*”.⁵⁷

Overall, a large amount of high-quality evidence indicates that bryostatins are defensive (and therefore part of an important mutualism) and that they are produced by “*Ca. E. sertula*”. Neither cultivation of the bacteria nor expression of the complete pathway has yet been achieved, leaving a missing link. The expressed proteins, while providing promising data, are generic to trans-AT PKS pathways.

2.4 Defense against predation: The ecteinascidins

The ecteinascidins are classical examples of defensive marine natural products. Notably, ecteinascidin-743 (ET-743, **8**) is the only compound covered in this review to receive approval by the US Food and Drug Administration (for specific sarcomas).^{67, 68} ET-743 and congeners were isolated from the tunicate *Ecteinascidia turbinata*.^{69, 70} An early study showed that unpalatable substances and orange coloration in larvae of *E. turbinata* protected the hosts from predators, such as fishes.⁷¹ Biosynthesis was initially studied using

incorporation of radiolabeled precursors, indicating that ET-743 might be assembled from two halves.^{72, 73} Related compounds were also isolated from sponges (e.g., renieramycins),⁷⁴ mollusks (e.g., jorumycin),⁷⁵ and bacteria (e.g., saframycin A (**9**) and safracins),^{71, 76, 77} leading to the speculation that these might be microbial compounds.

A significant advance in the field was the sequencing of the saframycin MX1 biosynthetic pathway, from the bacterium *Myxococcus xanthus*.^{78, 79} This work demonstrated that saframycins originate from a nonribosomal peptide synthetase (NRPS) gene cluster. Because this was done very early in the sequencing era, a number of errors accrued (including only about 90% sequencing accuracy), leading to a misassignment of some of the functions. For example, the last two modules are actually two separate proteins, as found in other saframycin clusters. Subsequently, based upon this early work many clusters were sequenced and analyzed, with the culmination being the identification of a stuttering NRPS module incorporating two tyrosine derivatives and being responsible for “Pictet-Spenglerase” activity essential to formation of the alkaloid core.⁸⁰ The Pictet-Spenglerase mechanism was at odds with the results from radiolabeling in the tunicate, since rather than resulting from two halves it derives from four amino acids.

Initial progress with ET-743 biosynthesis was made using microbiological methods. An uncultivated γ -proteobacterium, “*Candidatus Endoecteinascidia frumentensis*”, was proposed to produce ET-743 based upon its presence in ET-containing tunicates.^{81, 82} Like the majority of bacteria, “*Ca. E. frumentensis*” could not be cultured (at least so far), making it difficult to experimentally validate the correlative data. In 2011, Sherman *et al.* applied a shotgun metagenomic and metaproteomic approach to find the putative ET-743 biosynthetic genes.⁸³ The Sherman group analysis relied on the previous work of Oikawa *et al.*, who generated a synthetic precursor **10** that was incorporated into **11** and advanced intermediates **12** en route to saframycin A in cultivated bacteria (Figure 1).⁸⁰ Sherman and co-workers used the same synthetic substrate **10**, which would not be expected to be the native substrate for ET-743 biosynthesis, and showed that a homologous reductase / Pictet-Spenglerase from the putative ET-743 cluster catalyzed the reduction to yield **11**. While the combined evidence that “*Ca. E. frumentensis*” produces ET-743 is exceptionally strong, the potential for reductases to exhibit promiscuous activity, and indeed to demonstrate potentially non-native activity in this case, leaves room for debate.

Subsequently, the complete genome of “*Ca. E. frumentensis*” was assembled to find additional genes, which were missing in the previous shotgun sequencing data.⁵ In addition, the genomic features including the small genome size (631 kb) and the low GC content (23.3%) suggested that “*Ca. E. frumentensis*” became an obligate symbiont through genome reduction.⁸⁴ The gene cluster of ET-743 (173 kb) occupied 27% of the small genome and survived the extreme genome reduction, implying its essential role in the tunicate-symbiont association. This placed it among a small number of other symbionts where genome reduction has focused on an essential specialized metabolite.^{85, 86} Among these, it represents the most striking example so far. In genome reduction, genes that are essential to symbiosis survive, while a rapid accumulation of mutations leads to loss of non-essential genes. Thus, the survival of the ET-743 and other biosynthetic genes within a reduced background represents experimental evidence for the essentiality of the biosynthetic pathway. Since

ET-743 is known to be defensive, chemical defense may be essential to the survival of the holobiont (animal + symbiont).

3. Experimentally defined nutritional mutualisms with a natural products component

There are many known mutualisms where bacteria and hosts exchange nutrients with each other. A particularly common example is cyanobacteria, providing fixed carbon usually in exchange for protection, nitrogen and so on. This section describes nutritional mutualisms where there is also a compelling story about how symbiotic microbes make bioactive natural products in the host animal, but there is as yet no experimental evidence to show that the natural products themselves contribute to the mutualism.

3.1 Cyanobacterial-sponge mutualism: The polybrominated diphenyl ethers

Polybrominated diphenyl ethers (PBDEs), e.g. **13** and **14**, are common marine metabolites found in marine animals including fishes,⁸⁷ sponges,⁸⁸ and mammals.^{89–91} Because PBDEs are widely used as fire retardants, industrial contamination was thought to cause its bioaccumulation. Contrary to expectation, ¹⁴C analyses of marine mammals indicated that methoxylated analogs of PBDEs are not anthropogenic chemicals but natural products,⁹² but the producing organism was unknown. It has long been known that marine γ -proteobacteria, *Pseudoalteromonas* spp., produce halogenated compounds structurally related to PBDEs,^{93–97} leading to the potential for a symbiotic origin for such compounds.

Among the most abundant sources of PBDEs are sponges in the family of *Dysideidae*, where PBDEs account for more than 10% of dried sponge tissues.^{88, 98, 99} The sponge *Dysidea herbacea* is always found associated with a filamentous cyanobacterial symbiont *Hormosilla spongelliae* (formerly *Oscillatoria spongelliae*), which is highly abundant in the sponge mesohyl.¹⁰⁰ The rate of photosynthesis and the abundance of the symbionts, in addition to many biological studies strongly imply a nutritional role for the cyanobacteria in the association.¹⁰¹

In 1994, Unson *et al.* isolated PBDEs from *H. spongelliae* within its host sponge. The observation provided the first insight into the metabolic association between *D. herbacea* and its symbiotic cyanobacterium.⁸⁸ However, compound localization does not necessarily coincide with the biogenic origin of the compound.¹⁰² Later, Ridley *et al.* found four chemotypes of *H. spongelliae*-harboring sponges, which contain either PBDEs, chlorinated peptides, or no halogenated small molecules.¹⁰³ A related early advance was the use of FISH to localize biosynthetic genes for chlorinated peptides to cyanobacteria.¹⁰⁴ In contrast to the bryostatins symbiosis,⁶² phylogenetic trees of the host sponges and the cyanobacterial symbionts were correlated with each other, suggesting the presence of cospeciation between *H. spongelliae* and the sponge hosts.^{103, 105}

In 2014, Agarwal *et al.* discovered and characterized PDBE biosynthesis in *Pseudoalteromonas*.⁹⁷ The pathway did not feature large modular enzymes such as PKSs and NRPSs described above, but instead contained a series of small proteins that modified substrates starting from chorismate. In *Dysideidae* sponges, a shotgun metagenomics

approach used insights from the bacterial enzymological studies to identify a candidate gene cluster in the sponge.¹⁰⁵ The cluster was validated by expressing genes and transforming synthetic substrate **15** into natural products **13** and **14**, providing an exceptional evidence level for function. This gene cluster originated from *H. spongeliae*, which was by far the dominant cyanobacterium in the sponges studied. The PBDE biosynthetic genes of the cyanobacterial symbionts¹⁰⁵ shared no to very low DNA sequence similarity with those from γ -proteobacteria,⁹⁷ but the core biosynthetic proteins are homologous. Different sponge/symbiont clades sequenced in the study contained different PDBEs or even no PDBEs.¹⁰⁵ The presence of PDBE biosynthetic genes absolutely correlated with the presence of compounds in the sponges.

This study provided a very high level of evidence that PDBEs are produced by symbiotic cyanobacteria, but there remains no functional evidence for a role in the sponges for the compounds. Beyond the halogenated compounds, it is likely that symbiotic cyanobacteria contribute many other metabolites to sponges. A good example is the localization of the excitatory amino acid dysiherbaine to *Synechocystis* cyanobacterial cells in sponges.¹⁰⁶

3.2 Cyanobacterial-tunicate mutualism: Peptide and other natural products

Cyanobacteria of the genus *Prochloron* form specific symbiotic associations with tunicates in the Family Didemnidae.¹⁰⁷ There, they are known to provide photosynthate and to recycle nitrogen, often providing a majority of the animals' nutritional requirements.^{108–110} In many instances, *Prochloron* has been shown to be vertically transmitted from parent to progeny, although horizontal transfer is evident in other cases and must occasionally occur based on genomic studies.^{24, 111, 112} The biology of *Prochloron* and associated tunicates has been extensively reviewed.^{113, 114} Despite this long history, *Prochloron* has yet to be stably cultivated, so that all experimental methods either use the whole animal, partially purified *Prochloron* cells from the animal, or metagenomic techniques.

Prochloron-bearing tunicates often contain nitrogenous compounds, especially cyanobactin family cyclic peptides such as patellamide A (**16**) and trunkamide (**17**).^{115–119} There are no studies defining the ecological roles of cyanobactins. Several cyanobactins exhibit cytotoxicity to mammalian cell lines or reverse multidrug resistance,¹²⁰ leading to speculation that they could be defensive in nature.^{121, 122} In addition, several are known to bind to metals specifically, although this may not be a universal property of cyanobactins.^{118, 119, 123} Studies demonstrate that certain patellamides bind to transition metals within cells, for example.¹²⁴ Thus, biological evidence is suggestive but not yet convincingly connected to a function in nature.

Many reviews describe cyanobactin structure, biosynthesis, and function,^{125–129} but here the focus will be on evidence underlying their symbiotic origin. It was long speculated that the cyanobactins and perhaps other compounds originated in *Prochloron*, and not in the animals themselves.^{115, 130} It should be emphasized that, while some of these early speculations are supported by current data, others have been disproven. For example, the polyketide patellazoles do not derive from cyanobacteria.⁸⁶ To test these hypotheses, cell separation studies were undertaken. In one lab, cyanobactins were localized to *Prochloron* cells,¹²⁰ while in another the compounds were distributed throughout the tunicate.¹³¹ This example

serves to reemphasize the fact that compounds in nature are often not localized to the producing cell type.

A metagenomic approach was used to determine the origin of patellamide cyanobactins.¹³² Two biogenetic hypotheses were likely: the compounds could be made by a modular NRPS mechanism, as found in the ET-743 example described above; alternatively, they could be made using the pathway type now referred to as the Ribosomally synthesized and Posttranslationally modified Peptide (RiPP) pathway. Both biogenetic origins had some precedent, although the patellamides were quite different from other biochemically characterized compounds. At the time, it was relatively trivial to investigate NRPS biosynthesis using homology cloning, while testing the RiPP hypothesis would require shotgun sequencing using expensive Sanger technology. Therefore, the NRPS hypothesis was examined first.¹⁴⁴ *Prochloron* cells were enriched using physical methods and used to prepare a fosmid library. Subsequently, degenerate PCR was used to screen the fosmid library for the presence of NRPS genes. Based upon the negative PCR results despite substantial effort, the NRPS hypothesis seemed unlikely, as no convincing pathway was discovered.

Therefore, to discover a non-NRPS cluster, tunicate cells were partially separated by physical means, and the resulting *Prochloron*-enriched fraction was sequenced using Sanger methods.¹³³ Using BLAST searching in all eight frames, a RiPP biosynthetic pathway was discovered to encode the exact sequences of patellamides A (**16**) and C on a single precursor peptide (Figure 3). In RiPP biochemistry, the sequences that directly encode products are referred to as “core peptides”.¹³⁴ In addition to the precursor peptide, a series of genes were annotated that we predicted to 1) heterocyclize cysteine and serine; 2) oxidize the resulting thiazole moieties; 3) proteolytically release the free N-termini of the core peptides; and 4) cleave the C-termini of the core peptides in tandem with macrocyclization. Because two core peptides were embedded in a single precursor, they could be aligned, showing that the precursor peptide contained two putative enzyme “recognition sequences” dictating the N- and C-termini of each core peptide. This observation was later exploited in generating recombinant libraries.¹³⁵

The pathway was expressed both by cloning just the *pat* biosynthetic gene cluster and expressing it in *Escherichia coli*, as well as by expressing the pathway as found on the previously reported fosmid library.¹³³ The production level was low, especially in the fosmid library where detection was barely possible. Because robust genome sequencing methods were used, it was clear that *pat* originated in *Prochloron*, and not in the host.^{24, 133}

At about the same time, Long *et al.* reported the shotgun cloning of enriched *Prochloron* DNA into a BAC library, describing a relatively high yield of patellamides produced in *E. coli*.¹³⁶ For unavoidable non-scientific reasons, it was not possible to sequence the producing BACs at the time of the initial study. It would still be of interest to determine this sequence and to further validate production, since the un-optimized level was relatively high as reported and would still be a valuable contribution to the biotechnological literature.

Subsequently, many related cyanobactin biosynthetic gene clusters were found in tunicates.^{24, 137–141} These biosynthetic gene clusters fall into several clades leading to different types of chemical structures.¹⁴¹ Two features were apparent. First, within each clade, the enzymes are generally identical, or nearly so, within tunicates collected in many different locations.¹³⁷ By contrast, the precursor peptides encoding products are hypervariable, leading to a natural combinatorial library of products. This implied that the biosynthetic enzymes are highly substrate tolerant: although the cores change drastically, the enzymes are still able to modify those cores. Overall, about 70 cyanobactin products have been found encoded in various core peptides from the tunicate metagenome. Understanding of the natural biosynthetic promiscuity was used to generate a non-native compound, eptidemnamide.¹³⁷ Second, between clades, enzymes were often identical, except that horizontally acquired new enzymes encoded new posttranslational modifications.¹³⁸ These new enzymes often were accompanied by changes in the recognition sequences embedded in the precursor peptide.¹⁴¹ This led to the hypothesis that recognition sequences could be paired with specific enzymes to generate libraries with unnatural combinations of posttranslational modifications.

When a core peptide was identified in the *Prochloron* metagenome, the product was identified in the host.^{139, 142} When a new core peptide sequence was discovered, the predicted compound was later identified in the host animal, further demonstrating this relationship.¹³⁹ Cyanobactins have also been discovered as a major component of free-living cyanobacteria,^{141, 143–145} although that is beyond the scope of this review.

There is exceptionally strong functional evidence for the *Prochloron*-cyanobactin connection. In addition to the heterologous expression experiments described above, several research groups have examined cyanobactin biosynthesis, including those of Schmidt, Dittmann, Haygood, Nair, Naismith, Jaspars, Suga, Sivonen, Liu, and Ruffner.^{133, 134, 137, 138, 141, 143–184} Every required enzyme in the cyanobactin *pat*-like and *tru*-like pathway groups have been crystallized and functionally analyzed. Moreover, many unnatural derivatives, including those using non-proteinogenic amino acids, have been synthesized *in vivo* and *in vitro*. Finally, stable high-yield expression of the cyanobactin trunkamide pathway was recently achieved using rational engineering and empirical methods.¹⁷⁹ Therefore, the assignment of the cyanobactin pathways to *Prochloron* is made with high confidence.

Although the natural role of cyanobactins is unknown, we would argue that the existing evidence indicates that they are clearly important to the symbiosis. Although *Prochloron* and *Lissoclinum patella* do not show evidence of co-speciation, there is a strong correlation between animal phylogeny and the presence of specific cyanobactins (regardless of geography or habitat), implying that something about the compounds may determine which *Prochloron* cells inhabit which host animal.¹⁴² By contrast, other *Prochloron* genomic markers did not correlate well with host phylogeny, (with caveats described²⁴). Most recently, transcriptomic analysis of *Prochloron* in *L. patella* reveals that the patellamide biosynthetic genes are among the most highly expressed genes in *Prochloron*, and that they are transcribed constitutively in *Prochloron* living within the animal.¹⁸⁵ This result further indicates the likely importance of cyanobactins to *Prochloron* and/or the host animals.

Beyond cyanobactins, several other compounds are made by *Prochloron*. Diverse biosynthetic pathways are encoded in the many *Prochloron*-tunicate metagenomes that have been sequenced. Among these are genes for MAA biosynthesis, described above, and many others.²⁴ Most of the resulting products have yet to be described, indicating a rich source for future small molecule discovery.

3.3 Bacteria-shipworm mutualism: Cellulases and natural products

Shipworms are historically famous animals.^{186, 187} The bivalve mollusks burrow in wood and are largely responsible for degrading wood in the oceans, including human structures. As such, they may have been indirectly responsible for the defeat of the Spanish Armada. In order to efficiently consume cellulose, most wood- or grass-eating animals exploit symbiotic microbes, which commonly release products such as acetate after digesting the sugars. These products are then absorbed by the animals. Such microbes generally inhabit the digestive tract, where they have direct access to wood or other cellulose-rich particles.

In 1983, Waterbury *et al.* isolated cellulolytic bacteria from the shipworm *Lyrodus pedicellatus*.¹⁸⁸ Distel *et al.* used FISH to show that the strain T7902 was localized to the gill as a true symbiont.¹⁸⁹ In 2002, shipworm gill isolates were formally described as a new genus and species of γ -proteobacteria, *Teredinibacter turnerae* (named after the renowned shipworm expert, Ruth Turner).¹⁹⁰ In addition to breaking down cellulose, *T. turnerae* is also a nitrogen fixer, which was deemed important because of the very nitrogen-limiting nature of the wood habitat. Nitrogen fixation could be precisely imaged within animal gills,¹⁹¹ further validating *T. turnerae* as the source of nitrogen for the host animals. It was considered remarkable that *T. turnerae* inhabited the gills of the shipworm, while wood digestion largely occurred in a specific organ, the cecum, where *T. turnerae* and other bacteria are essentially absent.¹⁹² This stood in contrast to the other cellulolytic symbioses previously described, where the symbionts coexisted with the food source in the gut. In 2014, Distel and co-workers demonstrated that *T. turnerae* cellulase proteins are highly abundant in the cecum, even though the bacteria are not.¹⁸⁷ This was proposed to provide a strategy in which symbiotic bacterial enzymes would liberate glucose, an energy-rich food source that would be directly consumed by the host animal. By contrast, in other symbioses the microbial symbionts consume glucose and release relatively nutrient-poor products.

Upon genome sequencing, it was clear that *T. turnerae* had an unusually large potential to produce secondary metabolites.¹⁹³ Isolated compounds include tartrolon E (**18**),¹⁸⁶ a member of a known family of antibiotics previously identified in cultivated actinomycetes,¹⁹⁴ and a novel siderophore.¹⁹⁵ In addition to the nutritional role, it was hypothesized that perhaps *T. turnerae* could provide some of the antibiotics that could keep the digestive organ free from bacteria. Indeed, tartrolons D and E were discovered by mass spectrometry in the cecum of shipworms, potentially reflecting this role.¹⁸⁶ This work also described the trans-AT PKS tartrolon gene cluster for the first time.

Another potential role for tartrolon or other shipworm antibiotics has to do with competition within gills. *T. turnerae* is not the sole occupant in gills, and indeed there are several related strains.^{192, 196, 197} Beyond wood, shipworms are also known to inhabit seagrasses¹⁹⁸ and even sulfide-rich sediments.¹⁹⁹ The underlying symbionts are equally diverse.

4. Compounds from symbiotic associations where there is no evidence of mutualism

There are several cases where evidence indicates that important and potent natural products are produced by symbionts of animals, but there is no conclusive demonstration of a mutualistic relationship between those organisms.

4.1 Entotheonella: The star biosynthetic symbionts in sponges

Many sponges contain “Entotheonella” (shorthand for “*Candidatus* Entotheonella spp.”), a star performer²⁰⁰ that synthesizes many different, structurally novel compounds within sponges.²⁰¹ The biological roles of these compounds can only be speculated at this point (defense), and the role of “Entotheonella” in sponges is unknown. Nonetheless, exceptionally broad data demonstrate “Entotheonella” to be the source of the compounds. Working with “Entotheonella” has been technically demanding, and much of the recent scientific advances have followed technical leaps, including single-cell sequencing²⁰¹ and innovative bioinformatics analysis of trans-AT PKSs.^{55, 56}

The story of “Entotheonella” begins with the eponymous sponge, *Theonella swinhoei*,²⁰² which contains a number of bioactive compounds with interesting novel scaffolds, including theonellamides (e.g., **19**),²⁰³ swinholides (e.g., **20**),^{204–206} polytheonamides (e.g., **21**),^{207–209} and pederin (**22**) derivatives such as onnamides (e.g., **23**).²¹⁰ Based on gross morphology and structure similarity between sponge-derived and cyanobacterial natural products, it was proposed that symbiotic cyanobacteria are likely origins of sponge-derived compounds.²¹¹ Indeed, *T. swinhoei* and related sponges contain a filamentous bacterium that grossly resembles cyanobacteria. However, in 1996, Bewley *et al.* showed that those filamentous bacteria were heterotrophic and not cyanobacteria.²¹² *T. swinhoei* was separated into four cell populations: sponge cells, unicellular cyanobacteria, unicellular heterotrophic bacteria, and filamentous heterotrophic bacteria, and analyzed the extracts from each population by HPLC and NMR. Contrary to the hypothesis, no major natural products were isolated from the sponge and cyanobacterial cells. Swinholide A (**20**), a cytotoxic macrolide, was detected from the unicellular heterotrophic bacteria. The filamentous heterotrophic bacteria contained an antifungal peptide, theopalauamide (**19**).²¹³ Subsequently, the theopalauamide-harboring symbiont in was identified as a novel filamentous δ -proteobacterium “*Candidatus* Entotheonella palauensis”.²¹⁴ Of note, genome-sequencing studies later described “Entotheonella” as a founding member of a novel phylum, not as proteobacteria.²⁰¹ In the study, 16S rRNA genes of the filamentous symbionts were isolated from different chemotypes of *T. swinhoei*. Genomic and chemical experiments revealed a potential correlation between symbiont 16S rRNA gene sequences and *T. swinhoei* chemotypes.

The first genomic evidence supporting the “sponge-symbiont hypothesis” was reported by Piel *et al.*²¹⁵ Piel started the work by identifying the pederin trans-AT PKS pathway in a terrestrial insect symbiont.²¹⁶ In the insect, strong prior evidence utilizing a “Koch’s postulates”-like approach indicated that specific bacteria made pederin (**22**).^{217, 218} In turn, pederin (**22**) was very similar in structure to onnamide A (**23**), previously identified in a

Japanese *T. swinhoei*.²¹⁰ The idea was that, by obtaining this gene cluster from a more experimentally validated system, the daunting sponge system would become approachable. A clone library was constructed from *T. swinhoei* and screened based on pederin genes. The resulting onnamide gene cluster was very similar in terms of sequence and organization to that from beetles, and it likely originated in “Entotheonella”. One interesting and convincing sequence feature was that onnamides contain an additional amino acid in comparison to pederin. Differences in the gene clusters could exactly explain this difference.²¹⁵ Functional evidence that the correct cluster was identified awaited expression of a small protein (methyltransferase) and characterization, such that like bryostatin the evidence remained sequence-based for several years.²¹⁹ Because of the specificity of the reaction, this provided among the best experimental evidence for function in a larger PKS or NRPS pathway from an uncultivated symbiont, in comparison to bryostatin or ET-743.

“Entotheonella” are also found in the genus *Discodermia*, among several others.^{220–222} Schirmer *et al.* suggested that “Entotheonella” spp. in *Discodermia dissoluta* might synthesize the antitumor compound discodermolide (**24**).²²³ Clone libraries were constructed from three cell fractions of *D. dissoluta*: total sponge tissue, unicellular bacteria, and filamentous bacteria.²²⁰ Interestingly, only filamentous bacteria and total sponge tissue contained NRPS and trans-AT PKS clusters, while unicellular bacteria did not. Although this work did not identify the pathway of discodermolide from the fraction of filamentous bacteria, 16S rRNA gene sequence and morphological analyses showed that the majority of the fraction was related to “*Ca. E. palauensis*” (97% identical over 1,335 bp of 16S rRNA). The observation was later supported by FISH imaging.²²¹ Similar to *D. dissoluta*, *Discodermia calyx*, also described as an “Entotheonella”-harboring sponge, contains many bioactive compounds, including calyculins^{224, 225} and calyxamides.²²² Wakimoto *et al.* identified the putative calyculin gene cluster in “Entotheonella”. In addition to trans-AT PKS-NRPS genes, the cluster encoded an additional phosphorylation event that created a calyculin A (**25**) prodrug.²²⁶ This novel compound was not previously detected because it is spontaneously enzymatically activated during tissue damage.^{227, 228}

There are a few examples implying the occurrence of horizontal gene transfer between “Entotheonella” spp. and other bacteria. For example, swinholide A (**20**), a macrolide detected in the unicellular heterotrophic bacteria fraction in *T. swinhoei*,²¹² was also isolated from a free-living cyanobacterium.²²⁹ This left several possibilities, including horizontal gene transfer between cyanobacteria and the heterotrophic bacteria, and compound secretion and transportation from cyanobacterial cells to other cells. Although the biosynthetic pathway of swinholide A was not yet described, Ueoka *et al.* identified three pathways of macrolides closely related to swinholide A (**20**).²³⁰ Based on fosmid-library screening and genomic analyses, putative biosynthetic gene clusters of misakinolide,²³¹ tolytoxin,²³² and luminaolide^{233, 234} were detected from three different bacteria: “Entotheonella” sp., a marine cyanobacterium, and a freshwater cyanobacterium, respectively. Comparative analysis of the three biosynthetic clusters suggested the possibility of horizontal gene transfer between bacteria. Recently, a swinholide biosynthetic gene cluster was described from terrestrial cyanobacteria, further indicating the widespread nature of this biosynthetic pathway in several different types of bacteria.²³⁵

Another example is that of kasumigamide (**26**), an anti-algal peptide originally isolated from a free-living freshwater cyanobacterium *Microcystis aeruginosa* NIES-87.²³⁶ Nakashima *et al.* also isolated kasumigamide (**26**) from a sponge *D. calyx*.²³⁷ Although it is common that highly similar structures are isolated from different organisms, this study was a less common case of exactly the same structure in different organisms. Using genetic and single-cell analyses, this work provided genetic evidence that “Entotheonella” spp. in *D. calyx* harbored an NRPS-PKS hybrid gene cluster (*kas*) that has an ability to produce kasumigamide (**26**). *M. aeruginosa* NIES-87, contained a similar NRPS-PKS cluster (*mkas*), which was identified and biochemically characterized as kasumigamide (**26**) biosynthetic genes. By contrast, *kas* was not readily characterized. Although the NRPS genes exhibited synteny in the two organisms, the PKS genes did not.²³⁸

Recently, “Entotheonella” was proposed as a new phylum.²⁰¹ Based on single-cell and metagenomic analyses, Wilson *et al.* assembled contigs representing large parts of the genomes from two closely related “Entotheonella” strains, from onnamide and polytheonamide chemotype *T. swinhoei*. Unlike many symbionts,^{5, 84, 86} the genome sizes of both “Entotheonella” were very large (>9 Mbp), with the average G+C content at 56%. Comparison of 16S rRNA gene sequences showed that these two strains were 97% identical to previously described “E. palauensis”.²¹⁴ Interestingly, one of the two strains “*Entotheonella factor* TSY1” carried almost all PKS and/or NRPS gene clusters for previously isolated natural products from its host. Further, genomic annotation and shotgun metaproteomics suggested that “Entotheonella” was facultative, anaerobic, and organotrophic.²³⁸ While the symbionts have the ability to produce most amino acids, they appeared to lack biosynthetic pathways for some vitamins. This information may prove useful in the ultimate stable cultivation of “Entotheonella”.

In addition to these genomic studies, the same group succeeded in identifying the biosynthetic pathway for polytheonamides, encoded in the “Entotheonella” genome.²³⁹ Polytheonamides are large and cytotoxic polypeptides, which are originally isolated from Japanese *T. swinhoei*.^{207–209, 240} Freeman *et al.* identified the polytheonamide precursor peptide as part of a RiPP pathway and characterized three enzymes using biochemical assays.²³⁹ Subsequently, expression and mutational studies characterized all of the enzymes required for biosynthesis.²⁴¹ Additional biochemical studies have fleshed out key details of this fascinating RiPP pathway.^{134, 242–247}

In sum, considerable evidence supports the versatility of “Entotheonella” in constructing numerous sponge compounds, including very strong and novel types of functional evidence. This includes a methylation reaction in the onnamide pathway, reconstitution and mutants generated in the polytheonamide pathway, and the forward prediction and biochemical analysis of additional and unexpected phosphorylation in the calyculin pathway. When functional studies are lacking, comparative studies are used, as illustrated with kasumigamide and onnamide. Because of their potent bioactivities, it is plausible that these “Entotheonella” metabolites work as defensive chemicals in their hosts, although the hypothesis has not been tested. In an early study that investigated anti-predatory activities of 17 Pacific sponge species, the extract of *T. swinhoei* appeared to show anti-predatory activities against fish and sea urchins.²⁴⁸

Recently, Keren *et al.* described an additional aspect of the sponge-“Entotheonella” association.²⁴⁹ Scanning electron and X-ray microscopes captured the images, which demonstrated that “Entotheonella” sp. in *T. swinhoei* accumulated arsenic and barium on intracellular membrane vesicles. The authors concluded that “Entotheonella” might help the sponge to detoxify metals, which is yet another possible foundation for mutualism.

4.2 Patellazole: A picomolar cytotoxin from intracellular tunicate symbionts

Like some of the “Entotheonella” compounds, patellazoles are extremely toxic to mammalian cells, with potencies in the picomolar range.²⁵⁰ Thus, patellazoles are of potential interest to cancer therapy, but due to several challenges, only the intriguing planar structure is known, leaving 12 undefined stereocenters for synthetic chemists. This left a potentially strong role for sequencing in drug development, since stereochemistry is often predictable in PKS clusters.

Patellazole A (**27**) is a polyketide-peptide hybrid first isolated from the tunicate *Lissoclinum patella*.²⁵¹ Since *L. patella* also houses *Prochloron*, it was speculated that *Prochloron* might make the compounds. However, genomic studies of *Prochloron*- and patellazole-containing tunicates indicated that *Prochloron* could not be the source of patellazoles.²⁴ Further metagenomic analysis revealed candidate trans-AT PKS-NRPS genes for patellazoles encoded in a newly discovered γ -proteobacterial symbiont, “*Ca. Endolissoclinum faulkneri*”.⁸⁶ However, these genes were not truly clustered, but instead were spread among four different loci on a relatively small (1.48 Mbp) genome. Normally, genome sequences are not completely assembled and instead reside on many distinct fragments known as “contigs”, as described for “Entotheonella” above. This is especially true in metagenomes, and in cases where the sequencing is done using high-throughput methodologies, such as Illumina. This often represents a major limitation of studies relying on sequencing. Remarkably, the “*E. faulkneri*” genome could be readily assembled into a single circular chromosome, despite resulting from Illumina sequencing. Because of this advantage, the genomic characteristics were obtained with high confidence, and all genes involved in biosynthesis could be discovered despite the lack of clustering.

Although the obtained PKSs closely matched the prediction for the patellazole structure, biochemical functional evidence was not obtained.⁸⁶ Instead, “a virtual knockout” was claimed based upon genomic correlation. Three closely related *L. patella* colonies were obtained. One of these, from Fiji, contained the patellazoles, “*E. faulkneri*”, and the patellazoles genes. Two others were found on the same reef in the Eastern Fields of Papua New Guinea in adjacent locations. Metagenomic and chemical analysis showed that the organisms were extremely similar, except that one lacked patellazoles, “*E. faulkneri*”, and the patellazoles genes, while the other contained all three.⁸⁶

Using FISH, “*E. faulkneri*” was localized to the *L. patella* blood stream, where they inhabit the tunicate equivalents of white blood cells.⁸⁶ The bacterial cells are shapeless, and in general resemble symbionts that have undergone substantial genome streamlining. Indeed, despite the size of about 1.5 Mbp, only about half of the genome is coding, and >10% of the coding genome encodes the single pathway to patellazoles.⁸⁶ This makes the symbiont in some ways quite similar to the ET-743 producer’s genome.⁵ It was proposed that this

genome reduction provided evidence that patellazoles are essential to the symbiosis, or at least to survival of the bacteria within the hosts. Further work to characterize the symbiosis revealed that “*E. faulkneri*” has been an obligate intracellular symbiont of *L. patella* for about 6–30 million years, a number that is complicated by the lack of a tunicate fossil record to calibrate the molecular clock.¹⁴²

Recently, the early stages of genome streamlining were observed in another polyketide-based, Gram-negative symbiont of didemnid tunicates.²⁵² The highly potent cytotoxic polyketides, mandelalides (e.g., **28**), are found in a South African *Lissoclinum* sp. tunicate, where they may deter predation. Kwan and coworkers sequenced the tunicate metagenome and identified “*Candidatus Didemnititus mandela*” as the likely true producer of mandelalides. Very early signs of genome streamlining were apparent, leading to an estimated age of the symbiosis of 30,000 years. Most interestingly, the biosynthetic gene cluster was found in seven copies, each in various stages of mutation, leading to 25.8% of the coding genomes involved in mandelalides production.

5. Cultivation: Are they mutualistic symbionts?

Cultivation of symbionts can be a powerful method to explore chemical mutualism. Good examples are described in sections 2.3 and 3.3 above. For the most part, as described in other sections, symbionts are still difficult to cultivate, necessitating molecular studies. However, while it would seem that cultivation simplifies the problem in comparison to molecular studies, often it does not, as will be shown in two further examples below.

5.1 Tetrodotoxin: What is the source?

Tetrodotoxin (**29**) is a widespread, potent sodium channel blocker with data supporting ecological roles including chemical defense.^{253–255} Tetrodotoxin is found in many different animal species on land and in the sea, leading to the hypothesis that it might be synthesized by symbiotic bacteria. Several excellent reviews summarize the recent literature in tetrodotoxin biosynthesis,^{256–258} biology,²⁵⁹ and potential symbiotic origin and production by bacteria.^{257, 258} In the oceans, the species most widely discussed in the context of potential symbiotic origin of tetrodotoxin are the pufferfish and blue-ringed octopus. No direct evidence yet links symbiosis to production of tetrodotoxin in these or other organisms.

More than 150 species of bacteria have been reported to produce tetrodotoxin, with the first identification more than 30 years ago.²⁵⁸ Most of these bacteria were isolated from tetrodotoxin-containing animals. Why, then, should there be any controversy about the origin of tetrodotoxin? Reports have universally relied on LC-MS, antibodies, and bioassay data.²⁵⁸ Although it would seem that fermentation would provide a good route to obtain valuable tetrodotoxin, the commercial source is still isolation from pufferfish (Sigma Aldrich). Isolation from bacteria in significant quantities seems impossible, so that NMR spectra are not reported. This has been proposed to result from the gradual loss of production over repeated culturing, starting from an initial low yield.²⁵⁸

Although low yields and loss of production are common problems in microbial fermentation, there are many known solutions, including mutational or biosynthetic

approaches. However, in the case of tetrodotoxin no biosynthetic gene(s) are known.^{260–262} Despite a potential biosynthetic relationship with saxitoxin, the biosynthesis of which is increasingly characterized,^{260–263} homologous genes have not been found for tetrodotoxin. Although tetrodotoxin and close analogs are taken up upon feeding to newts and pufferfish, putative intermediates or precursors are not, and thus even biosynthetic precursors remain mysterious.^{264, 265} In pufferfish, the current consensus is that tetrodotoxin is dietary in origin.²⁶⁵

Recently, advances by Yotsu-Yamashita and coworkers have provided key data that may ultimately be useful in discovering tetrodotoxin biosynthetic genes and proteins. In this work, putative intermediates in biosynthesis were isolated from newts using careful analytical chemistry studies.²⁶⁶ These results indicate that tetrodotoxin in newts may originate in a geranylated guanidinium species.

In summary, although there is a large amount of data, which is just briefly summarized here, skepticism arises because of the lack of chemical and DNA sequencing data that would be useful in supporting a microbial or symbiotic origin of tetrodotoxin. Data obtained in these areas would represent a major breakthrough in the field.

5.2 Polyketide pyrones: Found in bacteria and hosts

Mollusks have long been known to contain unsaturated polyketides, such as pyrones. For example, many sacoglossan mollusks contain polypropionate pyrones in their dorsal mantles.^{267, 268} The biological roles of these compounds are unknown. Radiolabeling evidence with ¹⁴C carbon dioxide revealed that the polypropionate pyrones are made de novo in the animals, via fixed carbon synthesized in chloroplasts within the mollusks.²⁶⁷ The true producers and biosynthetic pathways remain unknown.²⁶⁸

From cone snails, *Nocardiopsis* bacteria were cultivated and found to produce a series of γ -pyrones, the nocapyrones (e.g., **30**).²⁶⁹ An unusual PKS pathway was identified by sequencing and partially validated using biochemical analysis of an expressed methyltransferase. Because of the association of mollusks with pyrones, it was proposed that the nocapyrones might also be found in the snail host. Indeed, HPLC-MS studies revealed a high level of nocapyrone H (**30**) in *Nocardiopsis*-containing cone snails. A ¹³C-labeled internal standard was used to identify and quantify the compounds in snails, using the recombinant methyltransferase to incorporate the ¹³C label using ¹³C-*S*-adenosylmethionine. Further, the bacteria and biosynthetic genes from the bacteria could be associated with the snails by molecular methods. While the ecological role of the nocapyrones is not known, the compounds elicited action potentials in dorsal root ganglion neurons in mice, indicating a potential defensive or sensory role.²⁶⁹

6. Terminology

Much of the confusion around the symbiotic origin of important natural products stems from imprecise terms.

A *symbiosis* is defined as organisms that consistently live with each other. The roles of symbionts are not easily experimentally defined. A symbiosis may be obligate, where the organisms cannot live without each other; facultative, where the organisms do not absolutely require each other but where there is some demonstrated function underlying the interaction; or casual, where there is no strict association of the symbionts. The latter is sometimes not considered a symbiont because it is difficult to establish the consistency of the association, and *associate* may be a better term to use when consistency of an association is not known. Symbionts may also have a wide range of roles, from mutualists (described below) to pathogens, which harm the host. Commensalists are symbionts that rely on the host without harming it, but which provide nothing in return. As can be imagined, experimental evidence is usually lacking. In this field, symbiosis means a predictable relationship between organisms that can be followed over time and space.

The animal *microbiota* includes all of the associated microbes within an animal and does not make a judgment as to function. Conversely, a weakness is that “microbiota” does not imply any sort of functional importance, and many or most strains may be casual associates.

Uncultivated simply indicates that a strain is not yet in cultivation and makes no judgment as to whether it can or should be cultivated.

Mutualism is a subset of symbiosis where two organisms each benefit from a partnership. Mutualism is usually very difficult to demonstrate experimentally. In natural products, mutualism is usually assumed if one organism produces a compound that plays an important role (such as defensive) in the biology of another. Thus, here and in much of the current chemical mutualism literature the concept of mutualism is a proxy for a compound with an important role in the biology of a symbiosis.

7. Conclusion

7.1 Comparison: Establishing biosynthetic origin

There is significant debate in the natural products field about the origin of compounds. This debate rarely makes it into print, and when published is oblique. The controversies mainly result from the relative strengths of analytical methods.

Cultivation is a useful but risky way to establish source, as most cultivated isolates are casual associates that are unrelated to the biological question being posed. Examples of the successful use of cultivation, such as with shipworms, are usually also paralleled by very high level biological and sequencing studies, showing how the specific cultivated isolates in question contribute to the host through the life cycle of both symbiont and host. The case of isatin is also very convincing even without molecular data; the evidence is strong because the symbionts could be knocked out with antibiotics and reestablished. A weakness there is that those might not be the only players, or even the major players in nature, involved, since they preceded modern sequencing studies.

A common problem with using cultivation to assign source is that cultivated isolates show evidence of compound production that are not confirmed by methods accepted in chemistry

focused journals, and that the compounds are lost over repeated cultivation. This has seemingly complicated progress in understanding tetrodotoxin biosynthesis. Further, in these cases an absence of molecular evidence means that the importance of those cultivated microbes to production in the host cannot be confirmed. There are other examples of such instances where skepticism is still strong.

More often, symbionts are uncultivated or perhaps cannot be cultivated, necessitating molecular methods. The classic examples of using DNA library and sequencing methods to establish a symbiotic origin are described, with the bacterial 16S rRNA and PKS genes obtained for bryostatins and onnamide and other “Entotheonella” compounds. As with cultivated isolates, underlying these landmark projects on uncultivated bacteria is excellent biology, identifying specific types of bacteria and following them with biological tools, such as FISH. By performing these rigorous analyses, the origin of compounds could be unambiguously assigned. A weakness of those classical studies is that functional proof for the complex trans-AT PKS pathways was only slowly obtained and was indirect, using enzymes that may not be as specific as desired. Nonetheless, the totality of evidence is highly convincing.

Since those initial DNA library-based studies, it is now more common to see direct sequencing of samples, where a wide array of (meta)genomic sequencing and analysis methods have been adopted to various types of problems (single-cell sequencing, cellular enrichment sequencing, whole metagenome shotgun sequencing, etc.). Such methods were first applied to demonstrate the source of cyanobactins in *Prochloron*, and they have since been used in ET-743, the PDBEs, the MAAs, and various “Entotheonella” projects. Moreover, direct sequencing methods have greatly augmented later understanding of the early bryostatin and onnamide studies. Functional analysis remains somewhat problematic. The cyanobactin pathways were the first to be rationally discovered and expressed in *E. coli*, with later data for the “Entotheonella” polytheonamides, MAAs, and PDBEs being high convincing. What all of these pathways share is that they are relatively small and consist of small genes that are easier to synthesize and manipulate. At the time of this writing, modular genes such as PKSs and NRPSs still pose a challenge to the field, although recent advances are highly promising.²⁷⁰

What makes the evidence convincing, even when biochemical evidence is lacking, is strong comparative analysis based upon rigorously acquired sequencing data, where a pathway is predictable from principles such as colinearity and similarity to known pathways from cultivated isolates. Moreover, in such cases the pathway distribution exactly follows the detection of the compound in question, as found in the case of patellazoles.

7.2 Comparison: Establishing mutualism

These stories above provide good examples (albeit just a small number of such examples) of how the role of compounds in the environment can be determined, and how this supports the concept of an underlying symbiotic mutualism. Compounds such as MAAs, isatin, bryostatin, ET-743, and tetrodotoxin all have experimentally demonstrated defensive and other roles, albeit some with more evidence than others. Establishing a defensive role is greatly simplified by cultivation; if a compound originates in a cultivable symbiont, it may

be possible to eliminate or manipulate that symbiont, as has been done with all of the above examples except ET-743 and tetrodotoxin. In addition, the cultivation of shipworm symbionts provides an excellent model for testing the role of individual compounds in the symbiosis, although that has not yet been accomplished. It should be noted that, in all cases except where Koch's postulates can be met, molecular/genomic data has been essential to confirming a role. A good example of the necessity of such data can be found in the long history of debate about tetrodotoxin origin.

Proxies for mutualisms can sometimes be found. Certain kinds of genome reduction are good indications that a symbiosis may be obligate, and thus often mutualistic, as found in ET-743 and patellazoles. While some experimental evidence favors a defensive role for ET-743, no role is experimentally established for the highly toxic patellazoles. As the patellazoles pathway occupies >10% of a reduced genome and are highly potent, strong indirect evidence suggests that they are important to the symbiosis. Often, mutualisms are known via nutrition, such as the cellulolytic symbionts of shipworms and the photosynthetic symbionts of sponges and tunicates. In such cases, many excellent tools are available to unveil the biology and to show how nutrient from the environment move between symbiont and host, even absent cultivation.

7.3 Comparison: A framework for establishing symbiotic origin

As in the example of isatin, application of methods mimicking Koch's postulates provides strong initial evidence of a chemical mutualism, here defined as a symbiosis in which a compound plays an important role in biology (see Section 6). These principles were initially applied to determine the microbial cause of diseases, but the same ideas are exactly paralleled in establishing a chemical mutualism. Because of many limitations in Koch's postulates, microbiologists have designed further rules that are especially useful when genomics are applied. Here, we propose to extend these ideas to chemical mutualism and also describe proposed standards for the underlying data. These ideas are aimed at situations where classical Koch's postulates cannot be fulfilled, which covers most cases in the field.

1. Association of a putative chemical mutualist microbe with a host. The proposed producer should be consistently present when the compound in question is found in the whole animal, and absent when it is not. Exceptions should be explained via biological study. The bacteria should be localized within or on the host, preferably by visualization with FISH or similar methods, or they may be so abundant in the host that such studies are not necessary. Variation in the chemistry, such as different derivatives in different animals, should be explained by variation within the producing bacteria. Either the bacteria should be isolated and shown to produce the compounds, or genetic and biochemical methods may provide evidence in the absence of cultivation. The necessary chemical and genetic methods are described below.
2. Chemical data. Because weaknesses in chemical analysis have repeatedly complicated the field, the synthesis of a natural product by a cultivated microbe should be demonstrated using rigorous spectroscopic techniques, including both NMR and MS for key examples. The data should go beyond characterizing

known compounds, since these methods are notoriously prone to contamination by standards, and include isolation and characterization of new intermediates or designer products, or incorporation of stable isotopes followed by full characterization. This level of data should also be used to confirm genetic evidence, such as by heterologous expression of a pathway (below).

3. (Meta)genomic data. Except in especially strong cases of excellent chemical data using cultivated microbes, a symbiotic origin should not be claimed absent the characterization of biosynthetic genes responsible for synthesis of a compound. In metagenomic data, the highest level of evidence involves heterologous expression of a pathway to produce the natural products, unnatural derivatives, and labeled metabolites. Data can also be very strong when the product matches rules, such as colinearity, present in the genes; the genes are always and only present when the compound in question is also found in the host animal; biochemical data can be obtained for at least one predicted transformation *in vitro*.

Acknowledgments

The authors' work on symbiosis is funded by a fellowship from the Japan Society for the Promotion of Science to MM and by US National Institutes of Health research grants R35GM122521 and U01TW008163.

References

1. Bäckhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI. Science. 2005; 307:1915–1920. [PubMed: 15790844]
2. Gill SR, Pop M, DeBoy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, Gordon JI, Relman DA, Fraser-Liggett CM, Nelson KE. Science. 2006; 312:1355–1359. [PubMed: 16741115]
3. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. Nature. 2007; 449:804–810. [PubMed: 17943116]
4. Huttenhower C, Gevers D, Knight R, Abubucker S, Badger JH, Chinwalla AT, Creasy HH, Earl AM, Fitzgerald MG, Fulton RS, Giglio MG, Hallsworth-Pepin K, Lobos EA, Madupu R, Magrini V, Martin JC, Mitreva M, Muzny DM, Sodergren EJ, Versalovic J, Wollam AM, Worley KC, Wortman JR, Young SK, Zeng Q, Aagaard KM, Abolude OO, Allen-Verge E, Alm EJ, Alvarado L, Andersen GL, Anderson S, Appelbaum E, Arachchi HM, Armitage G, Arze CA, Ayvaz T, Baker CC, Begg L, Belachew T, Bhonagiri V, Bihan M, Blaser MJ, Bloom T, Bonazzi V, Paul Brooks J, Buck GA, Buhay CJ, Busam DA, Campbell JL, Canon SR, Cantarel BL, Chain PSG, Chen IMA, Chen L, Chhibba S, Chu K, Ciulla DM, Clemente JC, Clifton SW, Conlan S, Crabtree J, Cutting MA, Davidovics NJ, Davis CC, Desantis TZ, Deal C, Delehaunty KD, Dewhirst FE, Deych E, Ding Y, Dooling DJ, Dugan SP, Michael Dunne W, Scott Durkin A, Edgar RC, Erlich RL, Farmer CN, Farrell RM, Faust K, Feldgarden M, Felix VM, Fisher S, Fodor AA, Forney LJ, Foster L, Di Francesco V, Friedman J, Friedrich DC, Fronick CC, Fulton LL, Gao H, Garcia N, Giannoukos G, Giblin C, Giovanni MY, Goldberg JM, Goll J, Gonzalez A, Griggs A, Gujja S, Kinder Haake S, Haas BJ, Hamilton HA, Harris EL, Hepburn TA, Herter B, Hoffmann DE, Holder ME, Howarth C, Huang KH, Huse SM, Iizard J, Jansson JK, Jiang H, Jordan C, Joshi V, Katancik JA, Keitel WA, Kelley ST, Kells C, King NB, Knights D, Kong HH, Koren O, Koren S, Kota KC, Kovar CL, Kyrpides NC, La Rosa PS, Lee SL, Lemon KP, Lennon N, Lewis CM, Lewis L, Ley RE, Li K, Liolios K, Liu B, Liu Y, Lo CC, Lozupone CA, Dwayne Lunsford R, Madden T, Mahurkar AA, Mannon PJ, Mardis ER, Markowitz VM, Mavromatis K, McCorrison JM, McDonald D, McEwen J, McGuire AL, McInnes P, Mehta T, Mihindukulasuriya KA, Miller JR, Minx PJ, Newsham I, Nusbaum C, Ogloughlin M, Orvis J, Pagani I, Palaniappan K, Patel SM, Pearson M, Peterson J, Podar M, Pohl C, Pollard KS, Pop M, Priest ME, Proctor LM, Qin X, Raes J, Ravel J, Reid JG, Rhoades R, Rhodes R, Riehle KP, Rivera MC, Rodriguez-Mueller B, Rogers YH, Ross MC, Russ C, Sanka

RK, Sankar P, Fah Sathirapongsasuti J, Schloss JA, Schloss PD, Schmidt TM, Scholz M, Schriml L, Schubert AM, Segata N, Segre JA, Shannon WD, Sharp RR, Sharpton TJ, Shenoy N, Sheth NU, Simone GA, Singh I, Smillie CS, Sobel JD, Sommer DD, Spicer P, Sutton GG, Sykes SM, Tabbaa DG, Thiagarajan M, Tomlinson CM, Torralba M, Treangen TJ, Truty RM, Vishnivetskaya TA, Walker J, Wang L, Wang Z, Ward DV, Warren W, Watson MA, Wellington C, Wetterstrand KA, White JR, Wilczek-Boney K, Wu Y, Wylie KM, Wylie T, Yandava C, Ye L, Ye Y, Yooseph S, Youmans BP, Zhang L, Zhou Y, Zhu Y, Zoloth L, Zucker JD, Birren BW, Gibbs RA, Highlander SK, Methé BA, Nelson KE, Petrosino JF, Weinstock GM, Wilson RK, White O. *Nature*. 2012; 486:207–214. [PubMed: 22699609]

5. Schofield MM, Jain S, Porat D, Dick GJ, Sherman DH. *Environ Microbiol*. 2015; 17:3964–3975. [PubMed: 26013440]
6. Schmidt EW. *Nat Chem Biol*. 2008; 4:466–473. [PubMed: 18641627]
7. Piel J. *Nat Prod Rep*. 2004; 21:519–538. [PubMed: 15282634]
8. Piel J. *Nat Prod Rep*. 2009; 26:338–362. [PubMed: 19240945]
9. Flórez LV, Biedermann PH, Engl T, Kaltenpoth M. *Nat Prod Rep*. 2015; 32:904–936. [PubMed: 25891201]
10. Nakamura H, Kobayashi Ji, Hirata Y. *J. Chromatogr. A*. 1982; 250:113–118.
11. Shibata K. *Plant Cell Physiol*. 1969; 10:325–335.
12. Maragos, JE. Ph.D. Thesis. University of Hawaii; Honolulu: 1972.
13. Jokiel PL, York JRH. *Bull. Mar. Sci*. 1982; 32:301–315.
14. Dionisio-Sese L M, Ishikura M, Maruyama T, Miyachi S. *Mar. Biol*. 1997; 128:455–461.
15. Maruyama T, Hirose E, Ishikura M. *Biol. Bull*. 2003; 204:109–113. [PubMed: 12700141]
16. Hirose E, Ohtsuka K, Ishikura M, Maruyama T. *J. Mar. Biol. Assoc. U. K*. 2004; 84:789–794.
17. Hirose E, Hirabayashi S, Hori K, Kasai F, Watanabe MM. *Zoolog. Sci*. 2006; 23:57–63. [PubMed: 16547406]
18. Wängberg SA, Persson A, Karlson B. *J. Photochem. Photobiol. B*. 1997; 37:141–146.
19. Vernet M, Whitehead K. *Mar. Biol*. 1996; 127:35–44.
20. Portwich A, Garcia-Pichel F. *Phycologia*. 2003; 42:384–392.
21. Shick JM, Romaine-Lioud S, Ferrier-Pages C, Gattuso J-P. *Limnol. Oceanogr*. 1999; 44:1667–1682.
22. Balskus EP, Walsh CT. *Science*. 2010; 329:1653–1656. [PubMed: 20813918]
23. Asamizu S, Xie P, Brumsted CJ, Flatt PM, Mahmud T. *J. Am. Chem. Soc*. 2012; 134:12219–12229. [PubMed: 22741921]
24. Donia MS, Fricke WF, Partensky F, Cox J, Elshahawi SI, White JR, Phillippy AM, Schatz MC, Piel J, Haygood MG, Ravel J, Schmidt EW. *Proc. Natl. Acad. Sci. U S A*. 2011; 108:E1423–1432. [PubMed: 22123943]
25. Stochaj WR, Dunlap WC, Shick JM. *Mar. Biol*. 1994; 118:149–156.
26. Banaszak AT, Trench RK. *Protist*. 2001; 152:93–101. [PubMed: 11545441]
27. Shick JM, Dunlap WC, Pearse JS, Pearse VB. *Biol. Bull*. 2002; 203:315–330. [PubMed: 12480722]
28. Banaszak AT, Barba Santos MG, LaJeunesse TC, Lesser MP. *J. Exp. Mar. Biol. Ecol*. 2006; 337:131–146.
29. Starcevic A, Akthar S, Dunlap WC, Shick JM, Hranueli D, Cullum J, Long PF. *Proc. Natl. Acad. Sci. U S A*. 2008; 105:2533–2537. [PubMed: 18268342]
30. Shinzato C, Shoguchi E, Kawashima T, Hamada M, Hisata K, Tanaka M, Fujie M, Fujiwara M, Koyanagi R, Ikuta T, Fujiyama A, Miller DJ, Satoh N. *Nature*. 2011; 476:320–323. [PubMed: 21785439]
31. Plack PA, Fraser NW, Grant PT, Middleton C, Mitchell AI, Thomson RH. *Biochem. J*. 1981; 199:741–747. [PubMed: 7200360]
32. Grant PT, Middleton C, Plack PA, Thomson RH. *Comp. Biochem. Physiol. B, Biochem. Mol. Biol*. 1985; 80:755–759.

33. Chioccaro F, Zeuli L, Novellino E. *Comp. Biochem. Physiol. B, Biochem. Mol. Biol.* 1986; 85:459–461.
34. Osborn AR, Almabruk KH, Holzwarth G, Asamizu S, LaDu J, Kean KM, Karplus PA, Tanguay RL, Bakalinsky AT, Mahmud T. *Elife.* 2015; 4
35. Gil-Turnes MS, Hay ME, Fenical W. *Science.* 1989; 246:116–118. [PubMed: 2781297]
36. Gil-Turnes W S, Fenical W. *Biol. Bull.* 1992; 182:105–108. [PubMed: 29304709]
37. Venis MA, Stoessl A. *Biochem. Biophys. Res. Commun.* 1969; 36:54–56. [PubMed: 5796755]
38. Pettit GR, Herald CL, Doubek DL, Herald DL, Arnold E, Clardy J. *J. Am. Chem. Soc.* 1982; 104:6846–6848.
39. Pettit GR, Herald CL, Kamano Y. *J. Org. Chem.* 1983; 48:5354–5356.
40. Pettit GR, Herald CL, Kamano Y, Gust D, Aoyagi R. *J. Nat. Prod.* 1983; 46:528–531.
41. Pettit GR, Kamano Y, Herald CL, Tozawa M. *J. Am. Chem. Soc.* 1984; 106:6768–6771.
42. Cragg LH, Andreeff M, Feldman E, Roberts J, Murgo A, Winning M, Tombes MB, Roboz G, Kramer L, Grant S. *Clin. Cancer Res.* 2002; 8:2123–2133. [PubMed: 12114412]
43. Roberts JD, Smith MR, Feldman EJ, Cragg L, Grant S. *Clin. Lymphoma.* 2002; 3:184–188. [PubMed: 12521398]
44. Mayer AMS, Glaser KB, Cuevas C, Jacobs RS, Kem W, Little RD, McIntosh JM, Newman DJ, Potts BC, Shuster DE. *Trends Pharmacol. Sci.* 2010; 31:255–265. [PubMed: 20363514]
45. Haygood MG, Davidson SK. *Appl. Environ. Microbiol.* 1997; 63:4612–4616. [PubMed: 9361448]
46. Davidson SK, Allen SW, Lim GE, Anderson CM, Haygood MG. *Appl. Environ. Microbiol.* 2001; 67:4531–4537. [PubMed: 11571152]
47. Lim GE, Haygood MG. *Appl. Environ. Microbiol.* 2004; 70:4921–4929. [PubMed: 15294832]
48. Lopanik N, Lindquist N, Targett N. *Oecologia.* 2004; 139:131–139. [PubMed: 14747940]
49. Sharp KH, Davidson SK, Haygood MG. *ISME J.* 2007; 1:693–702. [PubMed: 18059493]
50. de Vries DJ, Herald CL, Pettit GR, Blumberg PM. *Biochem. Pharmacol.* 1988; 37:4069–4073. [PubMed: 3190746]
51. Wender PA, Cribbs CM, Koehler KF, Sharkey NA, Herald CL, Kamano Y, Pettit GR, Blumberg PM. *Proc. Natl. Acad. Sci. U S A.* 1988; 85:7197–7201. [PubMed: 3174627]
52. Hildebrand M, Waggoner LE, Liu H, Sudek S, Allen S, Anderson C, Sherman DH, Haygood M. *Chem. Biol.* 2004; 11:1543–1552. [PubMed: 15556005]
53. Sudek S, Lopanik NB, Waggoner LE, Hildebrand M, Anderson C, Liu H, Patel A, Sherman DH, Haygood MG. *J. Nat. Prod.* 2007; 70:67–74. [PubMed: 17253852]
54. Cheng YQ, Tang GL, Shen B. *Proc. Natl. Acad. Sci. U S A.* 2003; 100:3149–3154. [PubMed: 12598647]
55. Piel J. *Nat. Prod. Rep.* 2010; 27:996–1047. [PubMed: 20464003]
56. Helfrich EJN, Piel J. *Nat. Prod. Rep.* 2016; 33:231–316. [PubMed: 26689670]
57. Miller IJ, Vanee N, Fong SS, Lim-Fong GE, Kwan JC. *Appl. Environ. Microbiol.* 2016; 82:6573–6583. [PubMed: 27590822]
58. Lopanik NB, Shields JA, Buchholz TJ, Rath CM, Hothersall J, Haygood MG, Håkansson K, Thomas CM, Sherman DH. *Chem. Biol.* 2008; 15:1175–1186. [PubMed: 19022178]
59. Buchholz TJ, Rath CM, Lopanik NB, Gardner NP, Håkansson K, Sherman DH. *Chem. Biol.* 2010; 17:1092–1100. [PubMed: 21035732]
60. Clark MA, Moran NA, Baumann P, Wernegreen JJ. *Evolution.* 2000; 54:517–525. [PubMed: 10937228]
61. Sauer C, Stackebrandt E, Gadau J, Holldobler B, Gross R. *Int. J. Syst. Evol. Microbiol.* 2000; 50:1877–1886. [PubMed: 11034499]
62. Lim-Fong GE, Regali LA, Haygood MG. *Appl. Environ. Microbiol.* 2008; 74:3605–3609. [PubMed: 18390670]
63. McGovern TM, Hellberg ME. *Mol. Ecol.* 2003; 12:1207–1215. [PubMed: 12694284]
64. Trindade-Silva AE, Lim-Fong GE, Sharp KH, Haygood MG. *Curr. Opin. Biotechnol.* 2010; 21:834–842. [PubMed: 20971628]

65. Linneman J, Paulus D, Lim-Fong G, Lopanik NB. PLoS One. 2014; 9:e108783. [PubMed: 25275632]
66. McCutcheon JP, Moran NA. Nat. Rev. Microbiol. 2012; 10:13–26.
67. Delaloge S, Yovine A, Taamma A, Riofrio M, Brain E, Raymond E, Cottu P, Goldwasser F, Jimeno J, Misset JL, Marty M, Cvitkovic E. J. Clin. Oncol. 2001; 19:1248–1255. [PubMed: 11230465]
68. Casali PG, Sanfilippo R, D'Incalci M. Curr. Opin. Oncol. 2010; 22:342–346. [PubMed: 20489618]
69. Lichter W, Lopez DM, Wellham LL, Sigel MM. Proc. Soc. Exp. Biol. Med. 1975; 150:475–478. [PubMed: 1208565]
70. Rinehart L K, Holt G T, Fregeau L N, Stroh G J, Keifer A P, Sun L F, H Li, Martin G D. J. Org. Chem. 1990; 55:4512–4515.
71. Young CM, Bingham BL. Mar. Biol. 1987; 96:539–544.
72. Kerr G R, Miranda F N. J. Nat. Prod. 1995; 58:1618–1621.
73. Jeedigunta S, Krenisky M J, Kerr G R. Tetrahedron. 2000; 56:3303–3307.
74. Suwanborirux K, Amnuoypol S, Plubrukarn A, Pummangura S, Kubo A, Tanaka C, Saito N. J. Nat. Prod. 2003; 66:1441–1446. [PubMed: 14640515]
75. Fontana A, Cavaliere P, Wahidulla S, Naik CG, Cimino G. Tetrahedron. 2000; 56:7305–7308.
76. Takahashi K, Kubo A. J. Antibiot. 1977; 30:1015–1018. [PubMed: 591455]
77. Ikeda Y, Matsuki H, Ogawa T, Munakata T. J. Antibiot. 1983; 36:1284–1289. [PubMed: 6643278]
78. Pospiech A, Cluzel B, Bietenhader J, Schupp T. Microbiology. 1995; 141:1793–1803. [PubMed: 7551044]
79. Pospiech A, Bietenhader J, Schupp T. Microbiology. 1996; 142:741–746. [PubMed: 8936303]
80. Koketsu K, Watanabe K, Suda H, Oguri H, Oikawa H. Nat. Chem. Biol. 2010; 6:408–410. [PubMed: 20453862]
81. Moss C, Green DH, Pérez B, Velasco A, Henriquez R, McKenzie JD. Mar. Biol. 2003; 143:99–110.
82. Pérez-Matos AE, Rosado W, Govind NS. Antonie van Leeuwenhoek. 2007; 92:155–164. [PubMed: 17265101]
83. Rath CM, Janto B, Earl J, Ahmed A, Hu FZ, Hiller L, Dahlgren M, Kreft R, Yu F, Wolff JJ, Kweon HK, Christiansen MA, Håkansson K, Williams RM, Ehrlich GD, Sherman DH. ACS Chem. Biol. 2011; 6:1244–1256. [PubMed: 21875091]
84. Moran NA, McCutcheon JP, Nakabachi A. Annu. Rev. Genet. 2008; 42:165–190. [PubMed: 18983256]
85. Nakabachi A, Ueoka R, Oshima K, Teta R, Mangoni A, Gurgui M, Oldham NJ, Van Echten-Deckert G, Okamura K, Yamamoto K, Inoue H, Ohkuma M, Hongoh Y, Miyagishima SY, Hattori M, Piel J, Fukatsu T. Curr. Biol. 2013; 23:1478–1484. [PubMed: 23850282]
86. Kwan JC, Donia MS, Han AW, Hirose E, Haygood MG, Schmidt EW. Proc. Natl. Acad. Sci. U S A. 2012; 109:20655–20660. [PubMed: 23185008]
87. Watanabe I, Kashimoto T, Tatsukawa R. Chemosphere. 1987; 16:2389–2396.
88. Unson MD, Holland ND, Faulkner DJ. Mar. Biol. 1994; 119:1–11.
89. Marsh G, Athanasiadou M, Athanassiadis I, Bergman A, Endo T, Haraguchi K. Environ. Sci. Technol. 2005; 39:8684–8690. [PubMed: 16323763]
90. Vetter W, Scholz E, Gaus C, Müller JF, Haynes D. Arch. Environ. Contam. Toxicol. 2001; 41:221–231. [PubMed: 11462147]
91. Corsolini S, Borghesi N, Schiamone A, Focardi S. Environ. Sci. Pollution Res. 2007; 14:421–429.
92. Teuten EL, Xu L, Reddy CM. Science. 2005; 307:917–920. [PubMed: 15705850]
93. Burkholder PR, Pfister RM, Leitz FH. Appl. Microbiol. 1966; 14:649–653. [PubMed: 4380876]
94. Andersen RJ, Wolfe MS, Faulkner DJ. Mar. Biol. 1974; 27:281–285.
95. Holmström C, Kjelleberg S. FEMS Microbiol. Ecol. 1999; 30:285–293. [PubMed: 10568837]
96. Isnansetyo A, Kamei Y. Antimicrob. Agents Chemother. 2003; 47:480–488. [PubMed: 12543647]
97. Agarwal V, Li J, Rahman I, Borgen M, Aluwihare LI, Biggs JS, Paul VJ, Moore BS. Environ. Sci. Technol. 2015; 49:1339–1346. [PubMed: 25559102]

98. Carté B, Faulkner DJ. *Tetrahedron*. 1981; 37:2335–2339.
99. Norton RS, Croft KD, Wells RJ. *Tetrahedron*. 1981; 37:2341–2349.
100. Hinde R, Pironet F, Borowitzka MA. *Mar Biol*. 1994; 119:99–104.
101. Usher KM. *Mar. Ecol*. 2008; 29:178–192.
102. Simmons TL, Coates RC, Clark BR, Engene N, Gonzalez D, Esquenazi E, Dorrestein PC, Gerwick WH. *Proc. Natl. Acad. Sci. U S A*. 2008; 105:4587–4594. [PubMed: 18250337]
103. Ridley CP, Bergquist PR, Harper MK, Faulkner DJ, Hooper JN, Haygood MG. *Chem. Biol*. 2005; 12:397–406. [PubMed: 15797223]
104. Flatt PM, Gautschi JT, Thacker RW, Musafija-Girt M, Crews P, Gerwick WH. *Mar. Biol*. 2005; 147:761–774.
105. Agarwal V, Blanton JM, Podell S, Taton A, Schorn MA, Busch J, Lin Z, Schmidt EW, Jensen PR, Paul VJ, Biggs JS, Golden JW, Allen EE, Moore BS. *Nat. Chem. Biol*. 2017; 13:537–543. [PubMed: 28319100]
106. Sakai R, Yoshida K, Kimura A, Koike K, Jimbo M, Koike K, Kobiyama A, Kamiya H. *ChemBioChem*. 2008; 9:543–551. [PubMed: 18236479]
107. Withers NW, Alberte RS, Lewin RA. *Proc. Natl. Acad. Sci. U S A*. 1978; 75:2301–2305. [PubMed: 16592528]
108. Koike I, Yamamuro M, Pollard PC. *Mar. Freshw. Res*. 1993; 44:173–182.
109. Post AF, Bullerjahn GS. *FEMS Microbiol. Rev*. 1994; 13:393–413.
110. Yellowlees D, Rees TAV, Leggat W. *Plant Cell Environ*. 2008; 31:679–694. [PubMed: 18315536]
111. Hirose E, Fukuda T. *Zoolog. Sci*. 2006; 23:669–674. [PubMed: 16971784]
112. Hirose E. *Genesis*. 2015; 53:121–131. [PubMed: 24700539]
113. Hirose, E., Neilan, BA., Schmidt, EW., Murakami, A. *Handbook on Cyanobacteria*. Gault, PM., Marler, HJ., editors. Nova Science Publishers; New York: 2009. p. 161–189.
114. Lewin, RA., Cheng, L., editors. *Prochloron: A Microbial Enigma*. Chapman and Hall; New York: 1989.
115. Ireland C, Scheuer PJ. *J. Am. Chem. Soc*. 1980; 102:5688–5691.
116. Ireland CM, Durso AR, Newman RA, Hacker MP. *J. Org. Chem*. 1982; 47:1807–1811.
117. Carroll AR, Coll JC, Bourne DJ, MacLeod JK, Zabriskie TM, Ireland CM, Bowden BF. *Aust. J. Chem*. 1996; 49:659–667.
118. Wipf P, Uto Y. *Tetrahedron Lett*. 1999; 40:5165–5169.
119. Wipf P, Uto Y. *J. Org. Chem*. 2000; 65:1037–1049. [PubMed: 10814052]
120. Williams AB, Jacobs RS. *Cancer Lett*. 1993; 71:97–102. [PubMed: 8364904]
121. Degnan BM, Hawkins CJ, Lavin MF, McCaffrey EJ, Parry DL, Van Den Brenk AL, Watters DJ. *J. Med. Chem*. 1989; 32:1349–1354. [PubMed: 2724305]
122. Degnan BM, Hawkins CJ, Lavin MF, McCaffrey EJ, Parry DL, Watters DJ. *J. Med. Chem*. 1989; 32:1354–1359. [PubMed: 2724306]
123. Morris LA, Jaspars M, Kettenes-Van Den Bosch JJ, Versluis K, Heck AJR, Kelly SM, Price NC. *Tetrahedron*. 2001; 57:3185–3197.
124. Comba P, Eisenschmidt A, Gahan LR, Herten DP, Nette G, Schenk G, Seefeld M. *Chem. Eur. J*. 2017; 23:12264–12274. [PubMed: 28339125]
125. Schmidt EW, Donia MS. *Curr. Opin. Biotechnol*. 2010; 21:827–833. [PubMed: 21050742]
126. Czekster CM, Ge Y, Naismith JH. *Curr. Opin. Chem. Biol*. 2016; 35:80–88. [PubMed: 27639115]
127. Michael JP, Pattenden G. *Angew. Chem. Intl. Ed*. 1993; 32:1–23.
128. Bertram A, Pattenden G. *Nat. Prod. Rep*. 2007; 24:18–30. [PubMed: 17268606]
129. Schmidt, EW., Donia, MS. *Comprehensive Natural Products II*. Mander, L., Liu, H-W., editors. 2010.
130. Sings HL, Rinehart KL. *J. Ind. Microbiol. Biotechnol*. 1996; 17:385–396.
131. Salomon CE, Faulkner DJ. *J. Nat. Prod*. 2002; 65:689–692. [PubMed: 12027743]
132. Schmidt EW, Sudek S, Haygood MG. *J. Nat. Prod*. 2004; 67:1341–1345. [PubMed: 15332852]

133. Schmidt EW, Nelson JT, Rasko DA, Sudek S, Eisen JA, Haygood MG, Ravel J. *Proc. Natl. Acad. Sci. U S A.* 2005; 102:7315–7320. [PubMed: 15883371]
134. Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, Camarero JA, Campopiano DJ, Challis GL, Clardy J, Cotter PD, Craik DJ, Dawson M, Dittmann E, Donadio S, Dorrestein PC, Entian KD, Fischbach MA, Garavelli JS, Göransson U, Gruber CW, Haft DH, Hemscheidt TK, Hertweck C, Hill C, Horswill AR, Jaspars M, Kelly WL, Klinman JP, Kuipers OP, Link AJ, Liu W, Marahiel MA, Mitchell DA, Moll GN, Moore BS, Müller R, Nair SK, Nes IF, Norris GE, Olivera BM, Onaka H, Patchett ML, Piel J, Reaney MJT, Rebuffat S, Ross RP, Sahl HG, Schmidt EW, Selsted ME, Severinov K, Shen B, Sivonen K, Smith L, Stein T, Süßmuth RD, Tagg JR, Tang GL, Truman AW, Vederas JC, Walsh CT, Walton JD, Wenzel SC, Willey JM, Van Der Donk WA. *Nat. Prod. Rep.* 2013; 30:108–160. [PubMed: 23165928]
135. Ruffner DE, Schmidt EW, Heemstra JR. *ACS Synth. Biol.* 2015; 4:482–492. [PubMed: 25140729]
136. Long PF, Dunlap WC, Battershill CN, Jaspars M. *ChemBioChem.* 2005; 6:1760–1765. [PubMed: 15988766]
137. Donia MS, Hathaway BJ, Sudek S, Haygood MG, Rosovitz MJ, Ravel J, Schmidt EW. *Nat. Chem. Biol.* 2006; 2:729–735. [PubMed: 17086177]
138. Donia MS, Ravel J, Schmidt EW. *Nat. Chem. Biol.* 2008; 4:341–343. [PubMed: 18425112]
139. Donia MS, Fricke WF, Ravel J, Schmidt EW. *PLoS One.* 2011; 6
140. Donia MS, Ruffner DE, Cao S, Schmidt EW. *ChemBioChem.* 2011; 12:1230–1236. [PubMed: 21542088]
141. Donia MS, Schmidt EW. *Chem. Biol.* 2011; 18:508–519. [PubMed: 21513887]
142. Kwan JC, Schmidt EW. *PLoS One.* 2013; 8:e80822. [PubMed: 24324632]
143. Dittmann E, Wiegand C. *Mol. Nutr. Food Res.* 2006; 50:7–17. [PubMed: 16304634]
144. Sudek S, Haygood MG, Youssef DT, Schmidt EW. *Appl. Environ. Microbiol.* 2006; 72:4382–4387. [PubMed: 16751554]
145. Sivonen K, Leikoski N, Fewer DP, Jokela J. *Appl. Microbiol. Biotechnol.* 2010; 86:1213–1225. [PubMed: 20195859]
146. Ziemert N, Ishida K, Quillardet P, Bouchier C, Hertweck C, de Marsac NT, Dittmann E. *Appl. Environ. Microbiol.* 2008; 74:1791–1797. [PubMed: 18245249]
147. Lee J, McIntosh J, Hathaway BJ, Schmidt EW. *J. Am. Chem. Soc.* 2009; 131:2122–2124. [PubMed: 19166292]
148. Leikoski N, Fewer DP, Sivonen K. *Appl. Environ. Microbiol.* 2009; 75:853–857. [PubMed: 19047393]
149. McIntosh JA, Donia MS, Schmidt EW. *Nat. Prod. Rep.* 2009; 26:537–559. [PubMed: 19642421]
150. Schmidt EW, Donia MS. *Methods Enzymol.* 2009; 458:575–596. [PubMed: 19374999]
151. Houssen WE, Wright SH, Kalverda AP, Thompson GS, Kelly SM, Jaspars M. *ChemBioChem.* 2010; 11:1867–1873. [PubMed: 20715266]
152. Leikoski N, Fewer DP, Jokela J, Wahlsten M, Rouhiainen L, Sivonen K. *Appl. Environ. Microbiol.* 2010; 76:701–709. [PubMed: 20008171]
153. McIntosh JA, Donia MS, Schmidt EW. *J. Am. Chem. Soc.* 2010; 132:4089–4091. [PubMed: 20210311]
154. McIntosh JA, Robertson CR, Agarwal V, Nair SK, Bulaj GW, Schmidt EW. *J. Am. Chem. Soc.* 2010; 132:15499–15501. [PubMed: 20961047]
155. McIntosh JA, Schmidt EW. *ChemBioChem.* 2010; 11:1413–1421. [PubMed: 20540059]
156. Schmidt EW. *BMC Biol.* 2010; 8:83. [PubMed: 20594290]
157. McIntosh JA, Donia MS, Nair SK, Schmidt EW. *J. Am. Chem. Soc.* 2011; 133:13698–13705. [PubMed: 21766822]
158. Agarwal V, Pierce E, McIntosh J, Schmidt EW, Nair SK. *Chem. Biol.* 2012; 19:1411–1422. [PubMed: 23177196]
159. Koehnke J, Bent A, Houssen WE, Zollman D, Morawitz F, Shirran S, Vendome J, Nneoyiegbe AF, Trembleau L, Botting CH, Smith MC, Jaspars M, Naismith JH. *Nat. Struc. Molec. Biol.* 2012; 19:767–772.

160. Schmidt EW. Chem. Biol. 2012; 19:1501–1502. [PubMed: 23261594]
161. Tianero MD, Donia MS, Young TS, Schultz PG, Schmidt EW. J Am Chem Soc. 2012; 134:418–425. [PubMed: 22107593]
162. Bent AF, Koehnke J, Houssen WE, Smith MC, Jaspars M, Naismith JH. Acta Crystallogr. F. 2013; 69:618–623.
163. Koehnke J, Bent AF, Zollman D, Smith K, Houssen WE, Zhu X, Mann G, Lebl T, Scharff R, Shirran S, Botting CH, Jaspars M, Schwarz-Linek U, Naismith JH. Angew. Chem. Intl. Ed. 2013; 52:13991–13996.
164. Koehnke J, Morawitz F, Bent AF, Houssen WE, Shirran SL, Fuszard MA, Smellie IA, Botting CH, Smith MC, Jaspars M, Naismith JH. ChemBioChem. 2013; 14:564–567. [PubMed: 23483642]
165. Leikoski N, Liu L, Jokela J, Wahlsten M, Gugger M, Calteau A, Permi P, Kerfeld CA, Sivonen K, Fewer DP. Chem. Biol. 2013; 20:1033–1043. [PubMed: 23911585]
166. McIntosh JA, Lin Z, Tianero MD, Schmidt EW. ACS Chem. Biol. 2013; 8:877–883. [PubMed: 23411099]
167. Houssen WE, Bent AF, McEwan AR, Pieiller N, Tabudravu J, Koehnke J, Mann G, Adaba RI, Thomas L, Hawas UW, Liu H, Schwarz-Linek U, Smith MC, Naismith JH, Jaspars M. Angew. Chem. Intl. Ed. 2014; 53:14171–14174.
168. Mann G, Koehnke J, Bent AF, Graham R, Houssen W, Jaspars M, Schwarz-Linek U, Naismith JH. Acta Crystall. F. 2014; 70:1597–1603.
169. Koehnke J, Mann G, Bent AF, Ludewig H, Shirran S, Botting C, Lebl T, Houssen W, Jaspars M, Naismith JH. Nat. Chem. Biol. 2015; 11:558–563. [PubMed: 26098679]
170. Sardar D, Lin Z, Schmidt EW. Chem. Biol. 2015; 22:907–916. [PubMed: 26165156]
171. Sardar D, Pierce E, McIntosh JA, Schmidt EW. ACS Synth. Biol. 2015; 4:167–176. [PubMed: 24625112]
172. Bent AF, Mann G, Houssen WE, Mykhaylyk V, Duman R, Thomas L, Jaspars M, Wagner A, Naismith JH. Acta Crystall. D. 2016; 72:1174–1180.
173. Bras NF, Ferreira P, Calixto AR, Jaspars M, Houssen W, Naismith JH, Fernandes PA, Ramos MJ. Chemistry. 2016; 22:13089–13097. [PubMed: 27389424]
174. Hao Y, Pierce E, Roe D, Morita M, McIntosh JA, Agarwal V, Cheatham TE 3rd, Schmidt EW, Nair SK. Proc. Natl. Acad. Sci. U S A. 2016; 113:14037–14042. [PubMed: 27872314]
175. Oueis E, Jaspars M, Westwood NJ, Naismith JH. Angew. Chem. Intl. Ed. 2016; 128:5936–5939.
176. Parajuli A, Kwak DH, Dalponte L, Leikoski N, Galica T, Umeobika U, Trembleau L, Bent A, Sivonen K, Wahlsten M, Wang H, Rizzi E, De Bellis G, Naismith J, Jaspars M, Liu X, Houssen W, Fewer DP. Angew. Chem. Intl. Ed. 2016; 55:3596–3599.
177. Sardar D, Schmidt EW. Curr. Opin. Chem. Biol. 2016; 31:15–21. [PubMed: 26709871]
178. Sardar D, Tianero MD, Schmidt EW. Methods Enzymol. 2016; 575:1–20. [PubMed: 27417922]
179. Tianero MD, Pierce E, Raghuraman S, Sardar D, McIntosh JA, Heemstra JR, Schonrock Z, Covington BC, Maschek JA, Cox JE, Bachmann BO, Olivera BM, Ruffner DE, Schmidt EW. Proc. Natl. Acad. Sci. U S A. 2016; 113:1772–1777. [PubMed: 26831074]
180. Alexandru-Crivac CN, Umeobika C, Leikoski N, Jokela J, Rickaby KA, Grilo AM, Sjo P, Plowright AT, Idress M, Siebs E, Nneoyi-Egbe A, Wahlsten M, Sivonen K, Jaspars M, Trembleau L, Fewer DP, Houssen WE. Chem. Commun. 2017; 53:10656–10659.
181. Gu W, Schmidt EW. Acc. Chem. Res. 2017; 50:2569–2576. [PubMed: 28891639]
182. Oueis E, Nardone B, Jaspars M, Westwood NJ, Naismith JH. ChemistryOpen. 2017; 6:11–14. [PubMed: 28168143]
183. Oueis E, Stevenson H, Jaspars M, Westwood NJ, Naismith JH. Chem. Commun. 2017; doi: 10.1039/c7cc06550g
184. Sardar D, Hao Y, Lin Z, Morita M, Nair SK, Schmidt EW. J. Am. Chem. Soc. 2017; 139:2884–2887. [PubMed: 28195477]
185. Behrendt L, Raina JB, Lutz A, Kot W, Albertsen M, Halkjaer-Nielsen P, Sorensen SJ, Larkum AW, Kuhl M. ISME J. 2017; doi: 10.1038/ismej.2017.192

186. Elshahawi SI, Trindade-Silva AE, Hanora A, Han AW, Flores MS, Vizzoni V, Schrago CG, Soares CA, Concepcion GP, Distel DL, Schmidt EW, Haygood MG. *Proc. Natl. Acad. Sci. U S A*. 2013; 110:E295–304. [PubMed: 23288898]
187. O'Connor RM, Fung JM, Sharp KH, Benner JS, McClung C, Cushing S, Lamkin ER, Fomenkov AI, Henrissat B, Londer YY, Scholz MB, Posfai J, Malfatti S, Tringe SG, Woyke T, Malmstrom RR, Coleman-Derr D, Altamia MA, Dedrick S, Kaluziak ST, Haygood MG, Distel DL, McFall-Ngai MJ. *Proc. Natl. Acad. Sci. U S A*. 2014; 111
188. Waterbury JB, Calloway CB, Turner RD. *Science*. 1983; 221:1401–1403. [PubMed: 17759016]
189. Distel DL, DeLong EF, Waterbury JB. *Appl. Environ. Microbiol.* 1991; 57:2376–2382. [PubMed: 1722662]
190. Distel DL, Morrill W, MacLaren-Toussaint N, Franks D, Waterbury J. *Int. J. Syst. Evol. Microbiol.* 2002; 52:2261–2269. [PubMed: 12508896]
191. Lechene CP, Luyten Y, McMahon G, Distel DL. *Science*. 2007; 317:1563. [PubMed: 17872448]
192. Betcher MA, Fung JM, Han AW, O'Connor R, Seronay R, Concepcion GP, Distel DL, Haygood MG. *PLoS One*. 2012; 7:e45309. [PubMed: 23028923]
193. Yang JC, Madupu R, Durkin AS, Ekborg NA, Pedamallu CS, Hostetler JB, Radune D, Toms BS, Henrissat B, Coutinho PM, Schwarz S, Field L, Trindade-Silva AE, Soares CAG, Elshahawi S, Hanora A, Schmidt EW, Haygood MG, Posfai J, Benner J, Madinger C, Nove J, Anton B, Chaudhary K, Foster J, Holman A, Kumar S, Lessard PA, Luyten YA, Slatko B, Wood N, Wu B, Teplitski M, Mougous JD, Ward N, Eisen JA, Badger JH, Distel DL. *PLOS One*. 2009; 4:e6085. [PubMed: 19568419]
194. Pérez M, Crespo C, Schleissner C, Rodríguez P, Zúñiga P, Reyes F. J. *Nat. Prod.* 2009; 72:2192–2194. [PubMed: 19968258]
195. Han AW, Sandy M, Fishman B, Trindade-Silva AE, Soares CAG, Distel DL, Butler A, Haygood MG. *PLOS One*. 2013; 8:e76151. [PubMed: 24146831]
196. Distel DL, Beaudoin DJ, Morrill W. *Appl. Environ. Microbiol.* 2002; 68:6292–6299. [PubMed: 12450854]
197. Altamia MA, Wood N, Fung JM, Dedrick S, Linton EW, Concepcion GP, Haygood MG, Distel DL. *Mol. Ecol.* 2014; 23:1418–1432. [PubMed: 24765662]
198. Shipway JR, O'Connor R, Stein D, Cragg SM, Korshunova, Martynov A, Haga T, Distel DL. *PLoS One*. 2016; 11
199. Distel DL, Altamia MA, Lin Z, Shipway JR, Han A, Forteza I, Antemano R, Limbaco MGJP, Teboe AG, Dechavez R, Albano J, Rosenberg G, Concepcion GP, Schmidt EW, Haygood MG. *Proc. Natl. Acad. Sci. U S A*. 2017; 114:E3652–E3658. [PubMed: 28416684]
200. Bewley CA, Faulkner DJ. *Angew. Chem. Intl. Ed.* 1998; 37:2162–2178.
201. Wilson MC, Mori T, Rückert C, Uria AR, Helf MJ, Takada K, Gernert C, Steffens UA, Heycke N, Schmitt S, Rinke C, Helfrich EJ, Brachmann AO, Gurgui C, Wakimoto T, Kracht M, Crüsemann M, Hentschel U, Abe I, Matsunaga S, Kalinowski J, Takeyama H, Piel J. *Nature*. 2014; 506:58–62. [PubMed: 24476823]
202. Gray J, Edward. *Proc. Zool. Soc. Lond.* 1868
203. Matsunaga S, Fusetani N, Hashimoto K, Wächli M. J. *Am. Chem. Soc.* 1989; 111:2582–2588.
204. Carmely S, Kashman Y. *Tetrahedron Lett.* 1985; 26:511–514.
205. Kobayashi M, Tanaka Ji, Katori T, Matsuura M, Kitagawa I. *Tetrahedron Lett.* 1989; 30:2963–2966.
206. Kitagawa I, Kobayashi M, Katori T, Yamashita M, Tanaka JI, Doi M, Ishida T. J. *Am. Chem. Soc.* 1990; 112:3710–3712.
207. Hamada T, Sugawara T, Matsunaga S, Fusetani N. *Tetrahedron Lett.* 1994; 35:609–612.
208. Hamada T, Sugawara T, Matsunaga S, Fusetani N. *Tetrahedron Lett.* 1994; 35:719–720.
209. Hamada T, Matsunaga S, Yano G, Fusetani N. J. *Am. Chem. Soc.* 2005; 127:110–118. [PubMed: 15631460]
210. Sakemi S, Ichiba T, Kohmoto S, Saucy G, Higa T. J. *Am. Chem. Soc.* 1988; 110:4851–4853.
211. Kobayashi J, Ishibashi M. *Chem. Rev.* 1993; 93:1753–1769.
212. Bewley CA, Holland ND, Faulkner DJ. *Experientia*. 1996; 52:716–722. [PubMed: 8698116]

213. Schmidt EW, Bewley CA, Faulkner DJ. *J. Org. Chem.* 1998; 63:1254–1258.
214. Schmidt EW, Obratsova AY, Davidson SK, Faulkner DJ, Haygood MG. *Mar. Biol.* 2000; 136:969–977.
215. Piel J, Hui D, Wen G, Butzke D, Platzer M, Fusetani N, Matsunaga S. *Proc. Natl. Acad. Sci. U S A.* 2004; 101:16222–16227. [PubMed: 15520376]
216. Piel J. *Proc. Natl. Acad. Sci. U S A.* 2002; 99:14002–14007. [PubMed: 12381784]
217. Kellner RLL. *Chemoecology.* 2001; 11:127–130.
218. Kellner RLL. *Insect Biochem. Mol. Biol.* 2002; 32:389–395. [PubMed: 11886773]
219. Zimmermann K, Engeser M, Blunt JW, Munro MHG, Piel J. *J. Am. Chem. Soc.* 2009; 131:2780–2781. [PubMed: 19206228]
220. Schirmer A, Gadkari R, Reeves CD, Ibrahim F, DeLong EF, Hutchinson CR. *Appl. Environ. Microbiol.* 2005; 71:4840–4849. [PubMed: 16085882]
221. Brück WM, Sennett SH, Pomponi SA, Willenz P, McCarthy PJ. *ISME J.* 2008; 2:335–339. [PubMed: 18256706]
222. Kimura M, Wakimoto T, Egami Y, Tan KC, Ise Y, Abe I. *J. Nat. Prod.* 2012; 75:290–294. [PubMed: 22276742]
223. Gunasekera SP, Gunasekera M, Longley RE, Schulte GK. *J. Org. Chem.* 1990; 55:4912–4915.
224. Kato Y, Fusetani N, Matsunaga S, Hashimoto K, Fujita S, Furuya T. *J. Am. Chem. Soc.* 1986; 108:2780–2781.
225. Matsunaga S, Fusetani N. *Tetrahedron Lett.* 1991; 32:5605–5606.
226. Egami Y, Wakimoto T, Abe I. *Bioorg. Med. Chem. Lett.* 2014; 24:5150–5153. [PubMed: 25442302]
227. Wakimoto T, Egami Y, Nakashima Y, Wakimoto Y, Mori T, Awakawa T, Ito T, Kenmoku H, Asakawa Y, Piel J, Abe I. *Nat. Chem. Biol.* 2014; 10:648–655. [PubMed: 24974231]
228. Wakimoto T, Egami Y, Abe I. *Nat. Prod. Rep.* 2016; 33:751–760. [PubMed: 26923942]
229. Andrianasolo EH, Gross H, Goeger D, Musafija-Girt M, McPhail K, Leal RM, Mooberry SL, Gerwick WH. *Org. Lett.* 2005; 7:1375–1378. [PubMed: 15787510]
230. Ueoka R, Uria AR, Reiter S, Mori T, Karbaum P, Peters EE, Helfrich EJ, Morinaka BI, Gugger M, Takeyama H, Matsunaga S, Piel J. *Nat. Chem. Biol.* 2015; 11:705–712. [PubMed: 26236936]
231. Kato Y, Fusetani N, Matsunaga S, Hashimoto K, Sakai R, Higa T, Kashman Y. *Tetrahedron Lett.* 1987; 28:6225–6228.
232. Carmeli S, Moore RE, Patterson GML. *J. Nat. Prod.* 1990; 53:1533–1542. [PubMed: 2128517]
233. Kitamura M, Schupp PJ, Nakano Y, Uemura D. *Tetrahedron Lett.* 2009; 50:6606. [PubMed: 20119494]
234. Maru N, Inuzuka T, Yamamoto K, Kitamura M, Schupp PJ, Yamada K, Uemura D. *Tetrahedron Lett.* 2013; 54:4385–4387.
235. Humisto A, Jokela J, Liu L, Wahlsten M, Wang H, Permi P, Machado JP, Antunes A, Fewer DP, Svionen K. *Appl. Environ. Microbiol.* 2017; doi: 10.1128/AEM.02321-17
236. Ishida K, Murakami M. *J. Org. Chem.* 2000; 65:5898–5900. [PubMed: 10987919]
237. Nakashima Y, Egami Y, Kimura M, Wakimoto T, Abe I. *PLoS One.* 2016; 11:e0164468. [PubMed: 27732651]
238. Lackner G, Peters EE, Helfrich EJ, Piel J. *Proc. Natl. Acad. Sci. U S A.* 2017; 114:E347–E356. [PubMed: 28049838]
239. Freeman MF, Gurgui C, Helf MJ, Morinaka BI, Uria AR, Oldham NJ, Sahl HG, Matsunaga S, Piel J. *Science.* 2012; 338:387–390. [PubMed: 22983711]
240. Iwamoto M, Shimizu H, Muramatsu I, Oiki S. *FEBS Lett.* 2010; 584:3995–3999. [PubMed: 20699099]
241. Freeman MF, Helf MJ, Bhushan A, Morinaka BI, Piel J. *Nat. Chem.* 2017; 9:387–395. [PubMed: 28338684]
242. Fuchs SW, Lackner G, Morinaka BI, Morishita Y, Asai T, Riniker S, Piel J. *Angew. Chem. Int. Ed. Engl.* 2016; 55:12330–12333. [PubMed: 27584723]

243. Parent A, Benjdia A, Guillot A, Kubiak X, Balty C, Lefranc B, Leprince J, Berteau O. J. Am. Chem. Soc. 2017; doi: 10.1021/jacs.7b08402
244. Helf MJ, Jud A, Piel J. ChemBioChem. 2017; 18:444–450. [PubMed: 27966282]
245. Parent A, Benjdia A, Chartier G, Leprince J, Berteau O. J. Am. Chem. Soc. 2016; 138:15515–15518. [PubMed: 27934015]
246. Morinaka BI, Vagstad AL, Helf MJ, Gugger M, Kegler C, Freeman MF, Bode HB, Piel J. Angew. Chem. Int. Ed. Engl. 2014; 53:8503–8507. [PubMed: 24943072]
247. Freeman MF, Vagstad AL, Piel J. Curr. Opin. Chem. Biol. 2016; 31:8–14. [PubMed: 26625171]
248. Burns E, Ifrach I, Carmeli S, Pawlik JR, Ilan M. Mar. Ecol. Prog. Ser. 2003; 252:105–114.
249. Keren R, Mayzel B, Lavy A, Polishchuk I, Levy D, Fakra SC, Pokroy B, Ilan M. Nat. Commun. 2017; 8
250. Richardson AD, Aalbersberg W, Ireland CM. Anticancer Drugs. 2005; 16:533–541. [PubMed: 15846119]
251. Zabriskie TM, Mayne CL, Ireland CM. J. Am. Chem. Soc. 1988; 110:7919–7920.
252. Lopera J, Miller IJ, McPhail KL, Kwan Jason C. mSystems. 2018; doi: 10.1128/mSystems.00096-17
253. Williams BL. Mar. Drugs. 2010; 8:381–398. [PubMed: 20411104]
254. Itoi S, Yoshikawa S, Asahina K, Suzuki M, Ishizuka K, Takimoto N, Mitsuoka R, Yokoyama N, Detake A, Takayanagi C, Eguchi M, Tatsuno R, Kawane M, Kokubo S, Takanashi S, Miura A, Suitoh K, Takatani T, Arakawa O, Sakakura Y, Sugita H. Toxicon. 2014; 78:35–40. [PubMed: 24279996]
255. Bucciarelli GM, Bradley Shaffer H, Green DB, Kats LB. Sci. Rep. 2017; 7
256. Bane V, Lehane M, Dikshit M, O'Riordan A, Furey A. Toxins. 2014; 6:693–755. [PubMed: 24566728]
257. Jal S, Khora SS. J. Appl. Microbiol. 2015; 119:907–916. [PubMed: 26178523]
258. Magarlamov TY, Melnikova DI, Chernyshev AV. Toxins. 2017; 9
259. Lorentz MN, Stokes AN, Rößler DC, Lötters S. Curr. Biol. 2016; 26:R870–R872. [PubMed: 27728785]
260. Kalaitzis JA, Chau R, Kohli GS, Murray SA, Neilan BA. Toxicon. 2010; 56:244–258. [PubMed: 19761784]
261. Chau R, Kalaitzis JA, Neilan BA. Aquat. Toxicol. 2011; 104:61–72. [PubMed: 21543051]
262. Chau R, Kalaitzis JA, Wood SA, Neilan BA. Mar. Drugs. 2013; 11:2695–2712. [PubMed: 23917066]
263. Tsuchiya S, Cho Y, Yoshioka R, Konoki K, Nagasawa K, Oshima Y, Yotsu-Yamashita M. Angew. Chem. Intl. Ed. 2017; 56:5327–5331.
264. Kono M, Matsui T, Furukawa K, Yotsu-Yamashita M, Yamamori K. Toxicon. 2008; 51:1269–1273. [PubMed: 18420245]
265. Kudo Y, Chiba C, Konoki K, Cho Y, Yotsu-Yamashita M. Toxicon. 2017; 137:78–82. [PubMed: 28734983]
266. Kudo Y, Yasumoto T, Mebs D, Cho Y, Konoki K, Yotsu-Yamashita M. Angew. Chem. Intl. Ed. 2016; 55:8728–8731.
267. Ireland C, Scheuer PJ. Science. 1979; 205:922–923. [PubMed: 17813086]
268. Gavagnin M, Mollo E, Cimino G. Rev. Bras. Farmacogn. 2015; 25:588–591.
269. Lin Z, Torres JP, Ammon MA, Marett L, Teichert RW, Reilly CA, Kwan JC, Huguen RW, Flores M, Tianero MD, Peraud O, Cox JE, Light AR, Villaraza AJL, Haygood MG, Concepcion GP, Olivera BM, Schmidt EW. Chem. Biol. 2013; 20:73–81. [PubMed: 23352141]
270. Boyhüyük KA, Fleischhacker F, Linck A, Wesche F, Tietze A, Niesert C-P, Bode HB. Nat. Chem. 2018

Figure 1.
Defensive natural products isolated from marine animals.

Figure 2. Functional evidence for ET-743 biosynthesis

Saframycin A (9) is a bacterial compound shown to be made via an NRPS module SfmC2. Feeding precursor 10 to apo-SfmC2 enabled the reductase to function, affording 11, while the holo protein was able to fully incorporate 10 into the tetrahydroquinoline core, e.g. 12. The same analog 10 was used with putative ET-743 reductive domain EtuA2(RE) from the tunicate metagenome, although 10 is not identical to the putative precursor for 8; 10 was reduced to 11.

Figure 3. Functional evidence for bromophenyl ether biosynthesis

Natural sponge compounds **13** and **14** could be made from synthetic precursor **15** using expressed genes *bmp7* and *bmp7-12*, respectively. These genes were identified from symbiotic cyanobacteria living in a sponge.

Natural products **16** and **17**, first isolated from a tunicate, could be produced in *E. coli* using the patellamide and trunkamide biosynthetic pathways. These pathways are nearly identical (green: 100% identity). The major differences (red: >60% identity) mostly reflect regions where genes or domains with new functions have been imported. For example, the large red block at left reflects a change in chemoselectivity of the heterocyclase. The precursor peptides for each compound are shown, each containing two core peptides shown in black. The core peptide encoding the natural products are shown in bold.

tartralon E (18)

Figure 5.
An antibacterial natural product isolated from a shipworm.

Figure 6.

Natural products originally isolated from *Theonella swinhoei*. Onnamide A (**23**) is shown with the structurally-related natural product, pederin (**22**).

Figure 7.
Natural products from the genus *Discodermia*.

Figure 8.

The structures of patellazole A (27), mandelalide A (28), tetrodotoxin (29), and nocapyrone H (30).

Scheme 1.

Biosynthesis of shinorine (1). The shinorine gene cluster from *Anabaena variabilis* and the enzymatic route from sedoheptulose 7-phosphate (3) to shinorine (1).