Metabolomics Approaches to Dereplicate Natural Products from Coral-Derived Bioactive Bacteria

Jessica M. Deutsch[1], Paige Mandelare-Ruiz[2], Yingzhe Yang[1], Gabriel Foster[1], Apurva Routhu[1], Jay Houk[2], Yesmarie T. De La Flor[2], Blake Ushijima[2,3], Julie L. Meyer[4], Valerie J. Paul[2], Neha Garg[1,*]

[1]School of Chemistry and Biochemistry, Engineered Biosystems Building, Center for Microbial Dynamics and Infection, Georgia Institute of Technology, Atlanta, GA, USA

[2]Smithsonian Marine Station, Smithsonian Institution, Fort Pierce, FL, USA

[3]Department of Biology and Marine Biology, University of North Carolina Wilmington, Wilmington, NC

[4]Department of Soil and Water Sciences, University of Florida, Gainesville, FL, USA

Correspondence:

*Neha Garg neha.garg@chemistry.gatech.edu

Dedicated to Dr. William H. Gerwick, University of California at San Diego, for his pioneering work on bioactive natural products.

Abstract

Stony corals (Scleractinia) are invertebrates that form symbiotic relationships with eukaryotic algal endosymbionts and the prokaryotic microbiome. The microbiome has the potential to produce bioactive natural products providing defense and resilience to the coral host against pathogenic microorganisms, but this potential has not been extensively explored. Bacterial pathogens can pose a significant threat to corals, with some species implicated in primary and opportunistic infections of various corals. In response, probiotics have been proposed as a potential strategy to protect corals in the face of increased incidence of disease outbreaks. In this study, we screened bacterial isolates from healthy and diseased corals for antibacterial activity. The bioactive extracts were analyzed using untargeted metabolomics. Herein, UpSet plot and hierarchical clustering analyses were performed to identify isolates with the largest number of unique metabolites. These isolates also displayed different antibacterial activities. Through application of *in silico* and experimental approaches coupled with genome analysis, we dereplicated natural products from these coral-derived bacteria from Florida's coral reef environments. The metabolomics approach highlighted in this study serves as a useful resource to select probiotic candidates and enables insights into natural product-mediated chemical ecology in holobiont symbiosis.

Coral reefs are important ecosystems hosting many marine species and are well known hotspots of biodiversity. The coral holobiont is comprised of the coral animal, endosymbiotic microalgae (Symbiodiniaceae) and a diverse microbiome including bacteria, archaea, fungi, protists, and viruses. Shifts in the coral microbiome can have positive effects on the coral by contributing to host fitness and resilience, or negative consequences resulting in disease during periods of environmental stress such as increased ocean temperature and colonization by opportunistic pathogens such as *Vibrio shiloi* in the coral *Oculina patagonica*.[1, 2] In this regard, the coral holobiont can be broadly divided into three distinct compartments, surface mucus layer (SML), the coral tissue, and the underlying skeleton, all of which have differing assemblages of bacteria communities. The dominant microbial phyla observed within corals include Proteobacteria (Alpha and Gamma), Actinobacteria, and Cyanobacteria, while some other phyla, Bacteroidetes and Firmicutes, may be transient and important in the onset of microbialization in coral reefs. Within the three distinct layers of a coral holobiont, SML-associated bacteria can be dominated by Gammaproteobacteria. For example, some studies have found the SML of the corals *Pocillopora damicornis* and *Acropora palmata* dominated by the family Vibrionaceae that accounted for 40% of the microbiome including, but not limited to, *Photobacterium* sp., *Vibrio harveyii*, and *Vibrio coralliilyticus*.[3, 4] The broad range of taxa present in the SML are presumed to be mutualistic and aid in the protection of the corals by inhibiting the growth of pathogens.[3] Tissue-associated bacteria are strongly driven by the coral and the Symbiodiniaceae genotypes.[5] The skeleton-associated bacteria are dominated by the cyanobacteria and some anaerobic and anoxygenic photosynthetic microbial communities.[6]

The coral probiotic hypothesis links environmental adaptation to modulation of the coral-associated microbiome in order to obtain the most advantageous microbiota that contribute to nutrient cycling, intercellular communication, and protection against pathogens.[7] One potential aspect in the selection of microbiota with respect to protection against pathogens involves the ability of beneficial bacteria (probiotics) to produce natural products with biological properties

that can shape the diversity and richness of the coral microbiome while warding off pathogens.[8] Ritchie showed that coral mucus and 20% of the bacteria isolated from mucus of healthy colonies inhibited pathogenic *Vibrio* spp. in disk diffusion assays while bacteria from mucus of apparently healthy coral colonies during a bleaching event failed to inhibit pathogens.[3] Prolonged thermal stress and acidification of the holobiont is implicated in the diminished ability of the microbial population to fight off pathogens, resulting in dysbiosis, a disruption of the microbiome.[1, 9] The necessity for development of coral probiotics in reef restoration and prevention of disease cannot be emphasized enough with the rise and spread of several diseases of corals. Currently, coral reefs, particularly in Florida and the Caribbean, are experiencing an unprecedented disease outbreak, referred to as stony coral tissue loss disease (SCTLD).[10] Since 2014, SCTLD has spread within the Florida Reef Tract and has affected over 388,000,000 m² of reefs (Florida Department of Environmental Protection, Stony Coral Tissue Loss Disease Response), which includes at least 20 of the 45 Florida reef-building coral species. One coral pathogen that appears to be involved in coinfections with SCTLD, *V. coralliilyticus*, is resistant to several different classes of antibiotics, including tetracyclines, aminoglycosides, and beta-lactams.[11, 12] Various *Vibrio* spp. are implicated as primary or opportunistic coral pathogens and are often primary colonizers of stressed corals.[3, 8, 13] This dysbiosis can contribute to metabolic shifts within the coral microbiome and allow additional pathogenic microorganisms to colonize. Understanding the pathogenesis of SCTLD and development of probiotic treatments for this disease are urgent areas of research.

Development of probiotics can be guided by studying microbiomes of resilient coral colonies. Oftentimes during a coral disease outbreak, including during SCTLD, select coral colonies are not affected by disease despite being surrounded by diseased colonies.[14-17] Thus, these resistant corals serve as a potentially important source for the isolation of beneficial bacteria. Recently, a repository of cultured coral bacteria was established with the goal of promoting probiotic development while also enabling research aimed at understanding the role

of the bacterial fraction in coral holobiont health.[18] This study highlighted that phylogenetically diverse bacteria can be isolated from the coral holobiont by varying the growth conditions. In coral microbiome research, culture-based and untargeted metabolomics approaches have lagged behind culture-independent microbiome profiling and metagenome sequencing-based approaches. The mass spectral repositories are sparse both in availability of untargeted metabolomics data as well as in mass spectral data of natural products isolated from coral-derived bacteria. Many of the approaches that address this issue have been developed by Professor William Gerwick, to whom this special issue has been dedicated.[19-23] In this study, we highlight how metabolomics approaches can be applied to dereplicate the natural product potential of cultured coral-associated bacteria from Florida corals and further guide the selection of probiotic candidates. We also investigate metabolomics of known pathogenic strains such as *V. coralliilyticus* isolated from Florida corals.[11] Such datasets will also serve as valuable resources for natural product chemists investigating the ecological roles of bacterial natural products. The metabolomics infrastructure such as Reanalysis of Data User (ReDU) interface allows co-analysis of public datasets.[24] Thus, data from healthy and diseased corals can be effectively co-analyzed with data from cultured bacteria to probe the role of overlapping metabolites in health and disease. The successful implementation of such data-driven approaches necessitates availability of appropriate datasets with dereplicated natural product profiles.

In this study, we provide the first such dataset for cultured bacteria from different species of stony corals along with genome sequencing data of select strains and highlight dereplication strategies for natural products from several different bacterial species including several pathogenic species. These bacterial isolates displayed varying levels of antimicrobial activity against coral disease-associated bacteria. Using UpSet plots, we select individual strains with the most diverse metabolomes. One of these isolates was the only isolate displaying complete inhibition of *Leisingera* sp., a putative pathogenic species involved in SCTLD. The analyses

described in this manuscript further highlight the limited availability of mass spectral data from coral-associated bacteria in mass spectral libraries. Rapid dereplication of natural products from coral-derived bacteria will become increasingly expedient as centralized repositories and *in silico* tools are developed to aid natural product annotation. Here we highlight the applicability of a recently developed *in silico* tool, MolDiscovery. Such approaches will not only allow selection of probiotic candidates for coral reef restoration, but also enable insights into natural product mediated chemical ecology of complex symbiotic relationships.

Results and Discussion

Disk Diffusion Assay. The extracts generated from lyophilized and wet cellular fractions of bacterial cultures isolated from various coral species (Figure 1) were subjected to disk diffusion growth inhibition assays for screening of antibacterial activity against coral pathogens *V. coralliilyticus* OfT6-21 and OfT7-21[11] and against *Leisingera* sp. McT4-56, a representative isolate of the Rhodobacterales suspected to play a role in SCTLD[25] (Table 1). The antibiotic nalidixic acid was chosen as a positive control owing to its low minimum inhibitory concentration against *V. coralliilyticus*.[11] Overall, the extracts generated using lyophilized cells displayed higher antibacterial activities than the extracts generated using wet cell pellets. In some cases, no antibacterial activity was observed with extracts from wet cell pellets though extracts from their lyophilized cell counterparts displayed activity. The improvement in extraction efficiency when lyophilized cells are extracted is also supported through metabolomics data (see below). Only one strain of *V. coralliilyticus* (Cn52-H1) and the only strain of *Escherichia* sp. (Cnat2-31) investigated in this study displayed activity against *Leisingera* sp. McT4-56, while several strains displayed varying levels of antibacterial activity against *V. coralliilyticus* OfT6-21 and OfT7-21.

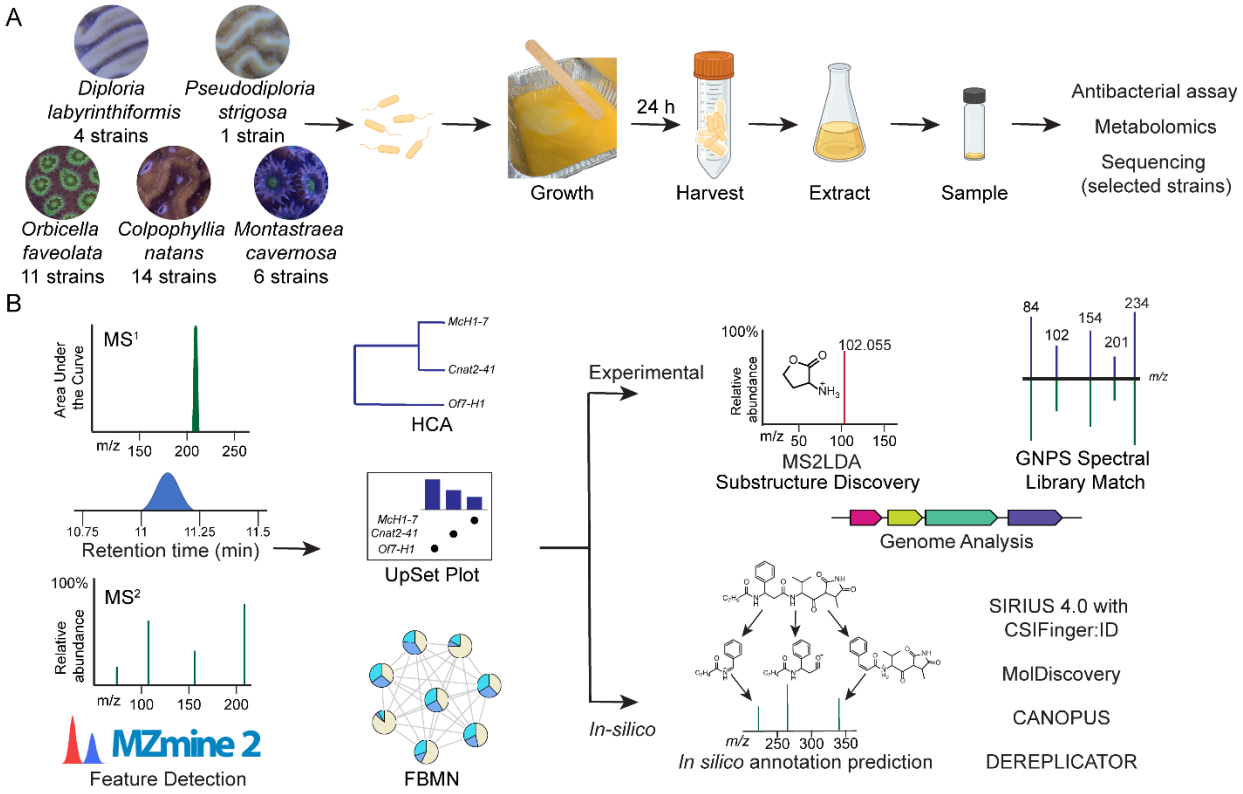

Figure 1. **Workflow for bacterial Isolation, sample preparation, data acquisition and analyses.** (A) Bacteria were isolated from five coral species. The number of bacterial strains isolated from each coral species is listed. Bacteria were grown on seawater agar in pans, harvested and extracted. (B) The area under the chromatogram for each metabolite was extracted using MZmine 2 and processed to remove metabolites detected in media and solvent controls. Metabolomic profiling was performed using hierarchical clustering analysis (HCA) and UpSet plots while compound annotation at level 2 was performed using molecular networking and substructure discovery. *In silico* tools were employed to enhance dereplication of known natural products. The annotations were confirmed using standards when available and through MS^2 comparison with published spectra.

Table 1: Average disk diffusion zones of inhibition (ZOIs, mm) of extracts generated from coral-derived bacteria (n=3) against putative pathogens, *V. coralliilyticus* (OfT6-21, OfT7-21) and *Leisingera* sp. (McT4-56). Complete inhibition is indicated by C, while P indicates partial growth within the zone of inhibition after 48 h of incubation (Figure S1A).

Samples:	16S rRNA ID:	Dosage (µg):	*Vibrio coralliiilyticus* (OfT6-21) (Mean ZOI ± SEM, mm)	*Vibrio coralliiilyticus* (OfT7-21) (Mean ZOI ± SEM, mm)	*Leisingera* sp. (McT4-56) (Mean ZOI ± SEM, mm)
Nalidixic Acid, Positive	-	62.5	C, 4.4 ± 0.1	C, 4.7 ± 0.1	C, 2.4 ± 0.1
Methanol, Negative	-	-	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
CnMc7-37, Ly	*Pseudoalteromonas* sp.	125.0	P, 3.4 ± 0.1	P, 2.6 ± 0.2	0.0 ± 0.0
Cnat2-18, Wet	*Pseudoalteromonas* sp.	125.0	P, 1.2 ± 0.1	P, 1.3 ± 0.3	0.0 ± 0.0
Cnat2-18, Ly	*Pseudoalteromonas* sp.	125.0	P, 2.6 ± 0.1	P, 2.8 ± 0.1	0.0 ± 0.0
CnH1-48, Wet	*Pleionea* sp.	125.0	P, 2.3 ± 0.1	P, 1.6 ± 0.3	0.0 ± 0.0
CnH1-48, Ly	*Pleionea* sp.	125.0	P, 3.4 ± 0.2	P, 2.4 ± 0.1	0.0 ± 0.0
DL2H-2.2, Wet	*Pseudoalteromonas* sp.	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
DL2H-2.2, Ly	*Pseudoalteromonas* sp.	125.0	P, 4.0 ± 0.4	P, 2.2 ± 0.3	0.0 ± 0.0
Cn5-1, Ly	*Tenacibaculum* sp.	125.0	P, 2.7 ± 0.1	P, 2.1 ± 0.1	0.0 ± 0.0
Cn5-1, Wet	*Tenacibaculum* sp.	125.0	P, 5.2 ± 0.4	P, 6.4 ± 0.6	0.0 ± 0.0
Cn5-12, Ly	*Halomonas* sp.	125.0	P, 2.4 ± 0.1	P, 2.6 ± 0.1	0.0 ± 0.0
Cn5-12, Wet	*Halomonas* sp.	125.0	0.0 ± 0.0	P, 10.6 ± 1.9	0.0 ± 0.0
Cn5-15, Ly	*Thalassobius* sp.	125.0	P, 2.8 ± 0.1	P, 2.3 ± 0.1	0.0 ± 0.0
Cn5-15, Wet	*Thalassobius* sp.	125.0	P, 11.2 ± 1.1	P, 9.4 ± 0.6	0.0 ± 0.0
Cn5-34, Ly	*Tenacibaculum* sp.	125.0	P, 2.3 ± 0.1	P, 2.7 ± 0.2	0.0 ± 0.0
Cn5-34, Wet	*Tenacibaculum* sp.	125.0	P, 13.8 ± 0.6	P, 13.7 ± 0.3	0.0 ± 0.0
Cnat2-28, Ly	*Ruegeria* sp.	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Cnat2-28, Wet	*Ruegeria* sp.	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Cnat2-31, Ly	*Escherichia* sp.	125.0	P, 3.5 ± 0.1	0.0 ± 0.0	0.0 ± 0.0
Cnat2-31, Wet	*Escherichia* sp.	125.0	0.0 ± 0.0	0.0 ± 0.0	P, 1.1 ± 0.2
Cnat3-28, Ly	*Alteromonas* sp.	125.0	P, 3.5 ± 0.1	0.0 ± 0.0	0.0 ± 0.0
Cnat2-41, Ly	*Pseudoalteromonas* sp.	125.0	P, 7.4 ± 0.4	P, 3.0 ± 0.2	0.0 ± 0.0
Cnat2-8, Ly	*Alteromonas* sp.	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
DL2-H6, Ly	*Pseudoalteromonas* sp.	125.0	P, 4.3 ± 0.1	P, 3.0 ± 0.4	0.0 ± 0.0
DL2-H6, Wet	*Pseudoalteromonas* sp.	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
DL2-H1, Ly	*Pseudoalteromonas* sp.	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
DL2-H1, Wet	*Pseudoalteromonas* sp.	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
DL2-H3, Ly	*Pseudoalteromonas* sp.	125.0	P, 2.9 ± 0.6	P, 1.3 ± 0.4	0.0 ± 0.0
DL2-H3, Wet	*Pseudoalteromonas* sp.	125.0	P, 1.7 ± 0.4	P, 0.9 ± 0.01	0.0 ± 0.0
Of5-H5, Ly	*Pseudoalteromonas* sp.	125.0	P, 2.2 ± 0.1	0.0 ± 0.0	0.0 ± 0.0
Of5-H5, Wet	*Pseudoalteromonas* sp.	125.0	P, 2.1 ± 0.8	0.0 ± 0.0	0.0 ± 0.0
Of7-H1, Ly	*Pseudoalteromonas* sp.	125.0	P, 3.8 ± 0.1	0.0 ± 0.0	0.0 ± 0.0
Of7-H1, Wet	*Pseudoalteromonas* sp.	125.0	P, 2.2 ± 0.8	0.0 ± 0.0	0.0 ± 0.0
Ofav3-42, Ly	*Ruegeria* sp.	125.0	P, 2.5 ± 0.1	0.0 ± 0.0	0.0 ± 0.0
Ofav3-42, Wet	*Ruegeria* sp.	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
McH1-25, Ly	*Halomonas* sp.	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
McH1-7, Ly	*Pseudoalteromonas* sp.	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Cn5-10, Ly	*Vibrio* sp.	125.0	P, 2.1 ± 0.1	0.0 ± 0.0	0.0 ± 0.0
Cn5-10, Wet	*Vibrio* sp.	125.0	P, 17.9 ± 0.7	P, 21.1 ± 2.0	0.0 ± 0.0
Of5-H9, Ly	*Vibrio* sp.	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Of5-H9, Wet	*Vibrio* sp.	125.0	P, 1.7 ± 0.3	0.0 ± 0.0	0.0 ± 0.0
Of7-H2, Ly	*Vibrio* sp.	125.0	P, 1.5 ± 0.1	P, 1.0 ± 0.1	0.0 ± 0.0
Of7-H2, Wet	*Vibrio* sp.	125.0	P, 1.1 ± 0.1	P, 0.9 ± 0.1	0.0 ± 0.0
Cn26-H1, Ly	*Vibrio coralliilyticus*	125.0	P, 6.6 ± 0.4	P, 2.8 ± 0.1	0.0 ± 0.0
Cn26-H1, Wet	*Vibrio coralliilyticus*	125.0	P, 2.3 ± 0.2	P, 3.0 ± 0.6	0.0 ± 0.0
Cn52-H1, Ly	*Vibrio coralliilyticus*	125.0	P, 5.2 ± 0.1	P, 5.4 ± 0.2	C, 1.7 ± 0.1
Cn52-H1, Wet	*Vibrio coralliilyticus*	125.0	P, 1.4 ± 0.1	P, 2.1 ± 0.4	0.0 ± 0.0
MmMcT2-4, Ly	*Vibrio coralliilyticus*	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
MmMcT2-4, Wet	*Vibrio coralliilyticus*	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Ofav2-20, Ly	*Vibrio* sp.	125.0	P, 7.1 ± 0.6	P, 4.2 ± 0.9	0.0 ± 0.0
Ofav2-20, Wet	*Vibrio* sp.	125.0	P, 1.6 ± 0.6	P, 1.4 ± 0.2	0.0 ± 0.0
Ofav2-48, Ly	*Vibrio* sp.	125.0	P, 1.3 ± 0.2	P, 1.6 ± 0.1	0.0 ± 0.0
Ofav2-48, Wet	*Vibrio* sp.	125.0	P, 2.2 ± 0.4	P, 1.8 ± 0.1	0.0 ± 0.0
Ofav3-11, Ly	*Vibrio* sp.	125.0	P, 3.6 ± 0.3	P, 4.0 ± 0.7	0.0 ± 0.0
Ofav3-11, Wet	*Vibrio* sp.	125.0	P, 2.4 ± 0.3	P, 3.0 ± 0.2	0.0 ± 0.0
MCA25, Ly	*Vibrio coralliilyticus*	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
MCA25, Wet	*Vibrio coralliilyticus*	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
MCA32, Ly	*Vibrio coralliilyticus*	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
MCA32, Wet	*Vibrio coralliilyticus*	125.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Ps1-H3, Ly	*Vibrio* sp.	125.0	P, 0.8 ± 0.1	P, 0.9 ± 0.4	0.0 ± 0.0
Ps1-H3, Wet	*Vibrio* sp.	125.0	P, 0.8 ± 0.1	P, 1.4 ± 0.1	0.0 ± 0.0
OfT6-17, Ly	*Vibrio coralliilyticus*	125.0	P, 2.4 ± 0.1	0.0 ± 0.0	0.0 ± 0.0
OfT6-17, Wet	*Vibrio coralliilyticus*	125.0	P, 3.1 ± 0.1	P, 3.1 ± 0.1	0.0 ± 0.0

Analyses of Untargeted Metabolomics and Phylogenetic Relatedness. The untargeted metabolomics data acquired via data independent acquisition were processed with Mzmine 2 to extract the area under the chromatographic peak for each metabolite feature detected. The features detected in the solvent controls and growth media blanks were removed from the resulting feature table. The metabolite features were then submitted to several metabolome profiling and dereplication workflows as shown in Figure 1. The Feature Based Molecular Networking (FBMN) analysis and substructure discovery using MS2LDA allowed compound annotation via matching to experimental spectra deposited in libraries and available in published literature. The *in silico* analyses tools namely MolDiscovery, Sirius 4 with CSI:FingerID, and DEREPLICATOR provided putative annotations, which were confirmed using MS^2 spectral comparison with literature reported spectra, with data acquired on analytical standards, and manual annotation of MS^2 fragments resulting in level 2 compound annotations. A list of annotations and the bacterial strains in which these compounds were detected is provided in Table S2. Where possible (Table S1, Figure S1C), the genome sequences were searched for the presence of biosynthetic gene clusters to further support compound annotations.

The metabolome diversity in extracts of cultured bacterial strains was queried by plotting the total number of unique features detected for each bacterial genus using UpSet plots. Among all bacterial isolates analyzed, *Pseudoalteromonas* spp. and *Vibrio* spp. displayed the largest number of unique features (Figure 2A). This observation may be pronounced by the fact that we analyzed the largest number of isolates from these two bacterial genera. Next, we queried whether select isolates of *Vibrio* and *Pseudoalteromonas* had greater metabolome diversity than others (Figure 2B and Figure S1). Herein, the largest number of unique features were detected in *Vibrio coralliilyticus* Cn52-H1 (Figure S1B). We employed *in silico* tools to aid dereplication of natural products produced exclusively by Cn52-H1, which aided annotation of moiramides and

andrimid in extracts of Cn52-H1 (*vide infra*). Similarly, the largest number of unique features among *Pseudoalteromonas* spp. were detected in extracts of *Pseudoalteromonas* sp. Cnat2-18 (Figure 2B). In the HCA analysis of metabolomics data, this strain appeared as a separate clade suggesting large metabolomic variation (Figure 2C). This strain slowed disease progression in *Colpophyllia natans* colonies affected with SCTLD in aquaria testing and is also currently being tested in the field as a probiotic treatment for SCTLD.[26] The clustering of bacterial strains observed in the HCA plot suggests that metabolomics data can be used as a proxy to gauge the diversity of isolated bacterial genera in culture-based studies and can be employed to prioritize strains for genome sequencing. While bacterial strains clustered by genus in the phylogenetic tree based on 71 single copy bacterial genes as expected (Figure S1C and Figure 2B), the HCA plot shows clustering of strains based on their metabolomics profile that did not always match their phylogenetic relatedness (Figure 2). For example, *Pseudoalteromonas* Cnat2-18 and 2-41 isolated from the same *C. natans* colony are present in immediately neighboring clades in the phylogenetic tree, but they are present in distant clades in the metabolomics-based HCA plot. Furthermore, neither the phylogenetic tree nor the metabolomics tree showed clustering based on the coral species. Thus, on the whole, metabolomic relatedness, which reflects chemical relationships between different strains, is not represented in their phylogenetic relatedness or the source of isolation. Select natural products that explain metabolome-based clustering are described below.

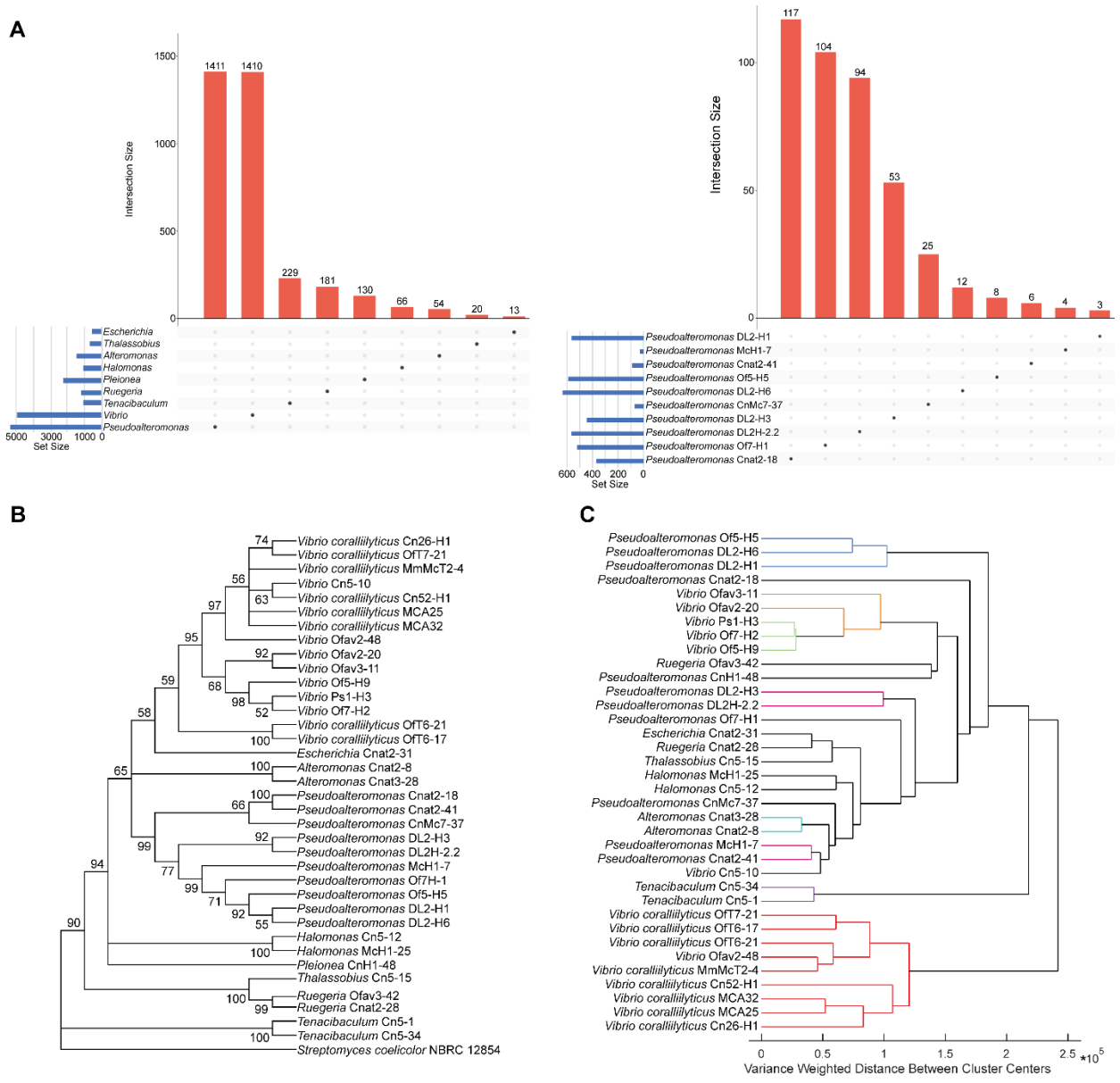

Figure 2. **Metabolome profiling using UpSet plots and HCA and phylogenomic analysis for select strains.** (A) The number of metabolite features uniquely detected in each bacterial genus is shown using an UpSet plot (left). The UpSet plot (right) shows the distribution of unique metabolite features among the *Pseudoalteromonas* spp. and the UpSet plot for distribution of metabolite features among *Vibrio* spp. is shown in Figure S1B. Intersection size refers to the number of unique metabolite features detected in each isolate, while set size refers to the total number of metabolite features detected in each isolate. (B) The phylogenetic tree based on 16S sequencing reveals clustering by genera. Isolates identified as *V. coralliilyticus* were fully sequenced and the full genomes are closely related to a type strain for *V. coralliilyticus*, therefore the identification to the species level is provided. C) The hierarchical

clustering analysis based on the acquired metabolomic data is shown. The colored branches are used to highlight clades where the data from the same genus clustered together.

Hybrid NRP/PK-Derived Andrimid and Moiramide. We first focused on annotation of natural products detected in *V. coralliilyticus* Cn52-H1 guided by UpSet plot analysis shown in Figure S1B. A set of features detected only in Cn52-H1 with *m/z* and retention time (*m/z*_t_R) 454.234_9.5 min, 496.244_10 min, and 480.249_10.6 min were proposed as moiramide B, moiramide C, and andrimid, respectively, using the *in silico* dereplication workflow via MolDiscovery (Figure 3A). The dereplication of these features was further supported by MS^2 spectral match with a published spectrum of an andrimid analog discovered via evolution of biosynthetic machinery[27] (Figure 3B). Herein, the MS^2 spectrum of andrimid is not available in any of the databases, while MS^2 spectrum of an analog with Ile in place of Val was available.[27] The biosynthetic gene cluster for andrimid was also detected in the genome of *V. coralliilyticus* Cn52-H1 isolated from *C. natans* (Figure 3C). The only missing gene was a homolog of *admU*, however, this gene product is not suspected to be involved in andrimid biosynthesis.[28] Thus, *in silico* tools paired with analysis of biosynthetic gene clusters enable annotation of known natural products when experimental spectra are sparse in databases. Compound annotations were propagated via shifts in *m/z* of characteristic fragment ions (Figure S2-S4). In addition to detection of biosynthetic intermediates lacking the methylsuccinimide ring, the substitutions in annotated analogs include variation in chain length and saturation of acyl tail and substitution of Val with Ile/Leu (Figure S2-S4). The antibacterial compound andrimid was first reported in cultures of *Enterobacter* sp., a planthopper symbiont, and the structurally similar moiramides were first characterized from *Pseudomonas fluorescens*.[29, 30] These compounds have since been shown to be produced by bacteria isolated from various hosts such as sponges,[31] tunicates,[30] and plants.[28] The methylsuccinimide substructure is critical for binding to the target, which is conserved among propagated analogs, while substitutions in the fatty acid are better tolerated.[30, 32] In addition to its broad spectrum antibacterial activity, andrimid was recently

shown to induce production of secondary metabolites in *Photobacterium galatheae* further highlighting the proposed ecological roles of this class of compounds.[33] Andrimid is reported to be produced by Indo-Pacific *V. coralliilyticus* isolates.[34, 35] It is interesting to note that andrimid and moiramides were only detected in one of the Atlantic isolates of *Vibrio* spp. in this study and this strain, Cn52-H1, is the only strain that resulted in complete inhibition of *Leisingera* sp. McT4-56 (Table 1). The Cn52-H1 extracts also displayed activity against *V. coralliilyticus* isolates OfT6-21 and OfT7-21 in the bioassay, which may be due to the putative resistance genes (*admQ* and *admT*) being located within the biosynthetic gene cluster[28] that are not present in these target strains. The workflow described for annotation of these compounds highlights that metabolomics approaches utilizing *in silico* tools can identify natural products wherein potential ecological roles can be proposed and then validated through experimental studies, use of genetic knock outs or direct detection in organic extracts of field-collected coral samples with relevant phenotypes.

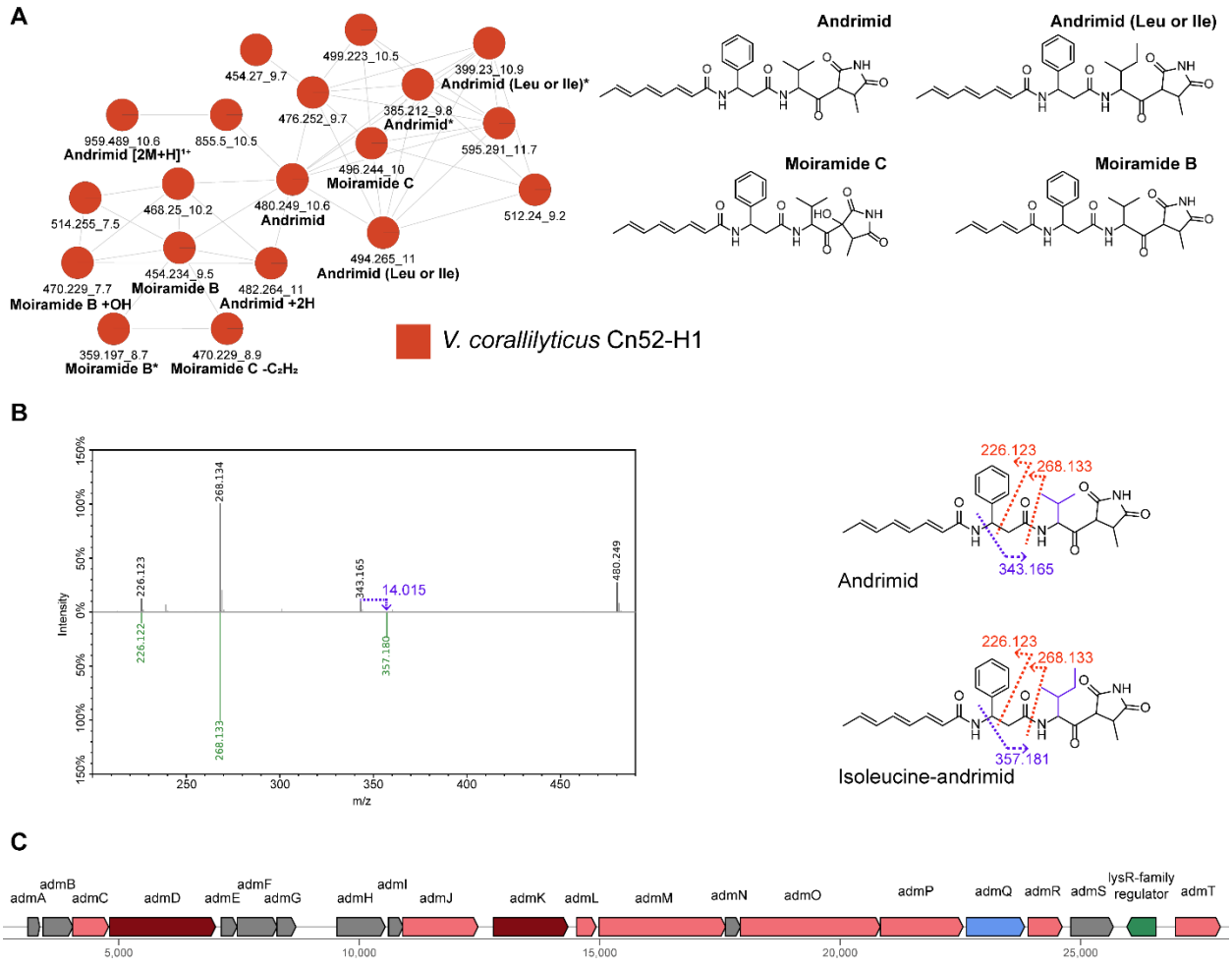

Figure 3. ***In silico* annotation of moiramides, andrimid and analogs.** (A) The cluster of nodes detected only in *V. coralliilyticus* Cn52-H1 corresponding to moiramides B, C, and andrimid are shown. The MS^2 spectral analysis of these compounds is shown in Figures S2-S4. (B) The mirror plot of andrimid (current dataset) and a known andrimid analog with Ile in place of Val is shown. [27]The * represents the biosynthetic intermediates lacking the methylsuccinimide unit. (C) The biosynthetic gene cluster for andrimid in *V. coralliilyticus* Cn52-H1 is shown, supporting the *in silico* annotation. The coding sequences are depicted to scale with respect to each other and the base pairs indicated by the numbers are located below the figure. The color coding is based upon the antiSMASH coding, with dark red as core biosynthetic genes, pink as additional biosynthetic genes, blue as transport-related genes, green as regulatory genes, and grey as other genes.

Hybrid NRP/PK Polycyclic Tetramate Macrolactams. A feature at *m/z*_t_R of 479.290_14.5 min was suggested as the polycyclic tetramate macrolactam ikarugamycin by two *in silico* dereplication workflows, namely MolDiscovery and SIRIUS 4.0 with CSI:Finger ID (Figure 4).

The polycyclic tetramate macrolactams are a class of antibiotics produced by a large diversity of

bacteria isolated from marine and terrestrial environments and the biosynthetic pathway is also shown to be conserved.[36, 37] Ikarugamycin has been shown to have strong antiprotozoal activity, antifungal activity, and antibiotic activity.[38, 39] The MS^2 spectra of this compound was also not available in spectral databases. This annotation was supported by MS^2 spectral and retention time matching with an analytical standard (Figure 4B). While an *in silico* annotation of ngercheumicin in *Pseudoalteromonas* by MolDiscovery was found to be incorrect, accurate annotation of structurally complex ikarugamycin supports applicability of these *in silico* tools in careful dereplication of complex natural products. Additional annotations were propagated to nodes connected to ikarugamycin via the shift in *m/z* of connected nodes (Figure 4A). A manual search of the molecular network for protonated molecules corresponding to known polycyclic tetramate macrolactams[40] enabled annotation of a singleton node with *m/z*_t_R 511.280_11.2 min as alteramide A, which was further supported through MS^2 spectral comparison with MS^2 from literature[41] (Figure 4A and B). Alteramide A was first reported in an *Alteromonas* sp. living symbiotically with a marine sponge and displayed cytotoxic activity.[40] This compound was also reported in a *Pseudoalteromonas* sp. isolated from necrotic tissue of a soft coral with light-dependent antifungal activity.[41] In this study, alteramide A was detected in *Pseudoalteromonas* sp. isolates Of5-H5 and Of7-H1, isolated from *Orbicella faveolata*, and all four strains isolated from *Diploria labyrinthiformis,* including DL2-H6, DL2-H3, DL2H-2.2, and DL2-H1. Ikarugamycin and its analogs were detected in all of these strains except DL2-H3 suggesting shared biosynthetic pathways and prevalence of these natural products in coral microbiomes (Figure 4).

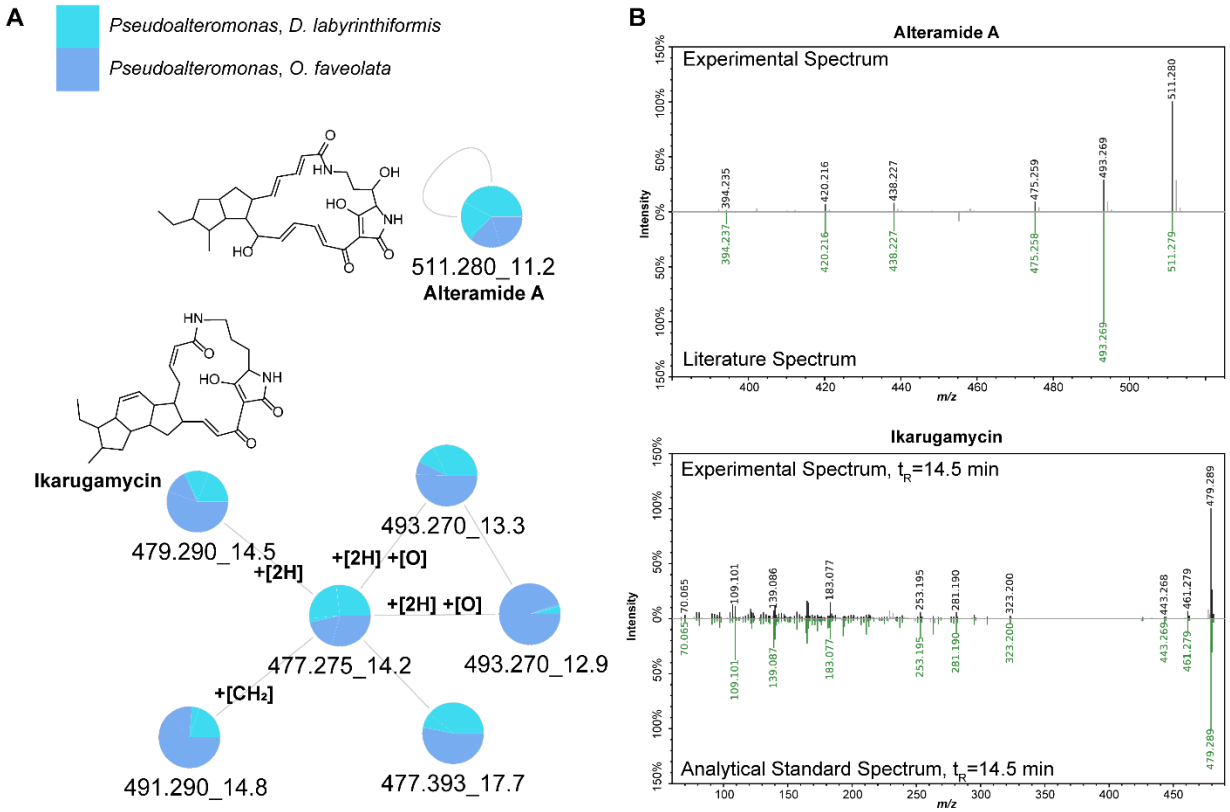

Figure 4. ***In silico* annotation of polycyclic tetramate macrolactams detected in *Pseudoalteromonas* sp. isolates.** (A) The network of nodes that cluster with ikarugamycin, as well as the singleton node annotated as alteramide A is shown. Ikarugamycin was detected in isolates Of5-H5, Of7-H1, DL2-H6, DL2H-2.2, and DL2-H1. Alteramide A was detected in isolates Of5-H5, Of7-H1, DL2-H6, DL2H-2.2, DL2-H1, and DL2-H3. (B) The MS² spectrum from literature[41] (bottom trace) is compared to the experimental MS² spectrum (top trace) for alteramide A. The MS² mirror match of ikarugamycin detected in the isolates (top trace) and an analytical standard of ikarugamycin (bottom trace) is shown. No MS² spectra was available in the literature.

***N*-acyl Homoserine Lactones.** *N*-acyl homoserine lactones (AHLs) are signaling molecules involved in bacterial quorum sensing, a population density driven process that allows bacteria to coordinate group behavior and synchronized production of secondary metabolites related to activities such as biofilm formation and virulence.[42] AHLs vary in the number of carbons in the acyl chain (C4–C20), in substitutions at C-3 (hydrogen, hydroxy, and oxo groups) and in the degrees of unsaturation in the acyl chain. Among these, longer chain length AHLs are common among marine bacteria and an acyl chain length of 20 carbons was recently reported in the marine bacterium *Rhodovulum sulfidophilum*.[43, 44] AHLs have been previously extensively

studied in terrestrial bacteria, but a limited knowledge base exists regarding AHL production and their biological roles in marine bacteria except for shorter chain length AHLs in *Vibrio* spp.

A feature with *m/z*_t_R 366.300_16.1 min was identified as C18:1-AHL via a match to MS^2 spectra in the GNPS database (Figure 5A). Because this feature did not cluster with any other features within the FBMN, we used substructure discovery via MS2LDA to search for additional AHLs. A motif 433, which contains the characteristic fragment peak 102.055 corresponding to the AHL homoserine lactone ring (Figure 5C), was shared between features with *m/z*_t_R 366.300_16.1 min and 384.311_14.9 min (Figure 5B). The feature at 384.311_14.9 min was annotated as hydrolyzed C18:1-AHL based on the fragment peak 120.066, which corresponds to the hydrolyzed homoserine lactone ring (Figure S5). Features that clustered with hydrolyzed C18:1-AHL were annotated as hydrolyzed AHLs containing C16-C17 fatty acyl chains with varying degrees of unsaturation (Figure 5B). The identification of these AHLs will be missed if tools such as MS2LDA are not employed. While the C18:1-AHL was detected in isolates of *Ruegeria* spp. Ofav3-42 and Cnat2-28, *Thalassobius* sp. Cn5-15, and *Escherichia* sp. Cnat2-31, the hydrolyzed AHLs were detected in many additional isolates including *Pseudoalteromonas* (DL2H-2.2, DL2-H6, DL2-H3, and DL2-H1 from *D. labyrinthiformis*, Of7-H1 and Of5-H5 from *O. faveolata*, CnMc7-37 from *Montastraea cavernosa*, and Cnat2-41 and Cnat2-18 from *C. natans*), *Alteromonas* (Cnat2-8 and Cnat3-28 from *C. natans*), *Pleionea* (CnH1-48 from *C. natans*), and *Halomonas* (McH1-25 from *M. cavernosa*) (Figure 5B). A recent study based on genome analysis also highlighted a high prevalence of genes for biosynthetic genes for AHLs in a variety of marine bacterial species.[18] Together these findings suggest that AHLs may play an important role in symbiosis in diverse coral microbiomes and warrants systematic investigation of the role of longer chain length AHLs in marine microbiomes.

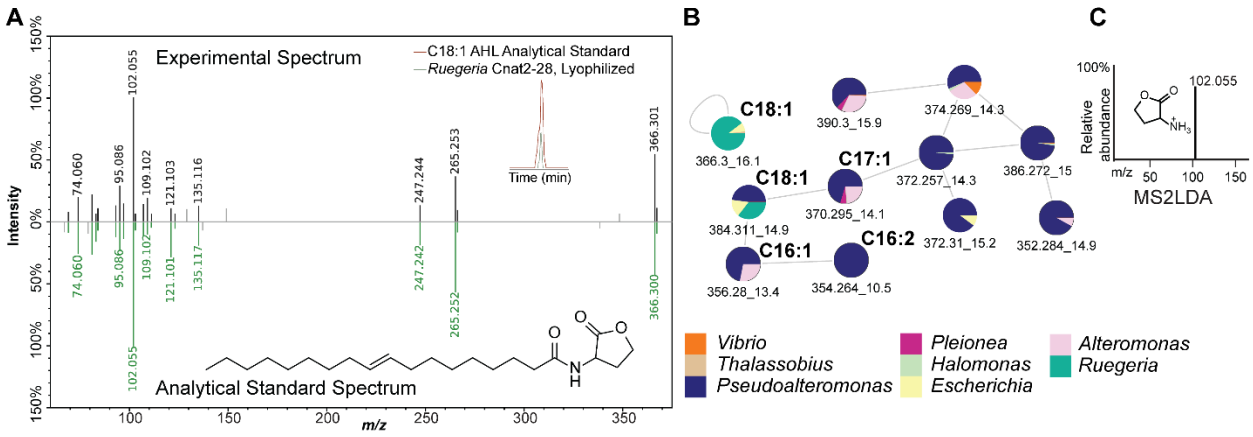

Figure 5. **Annotation of AHLs.** (A) The mirror plot of the experimental MS² spectrum (top) and the MS² spectrum acquired on the commercial analytical standard of C18:1-AHL (bottom). The inset shows the overlay of the extracted ion chromatogram of C18:1-AHL analytical standard and the experimental spectrum acquired on the extract of *Ruegeria* Cnat2-28. (B) The singleton node observed for C18:1-AHL and the hydrolyzed AHLs annotated in the molecular network are shown. The AHLs were detected in isolates Ofav3-42, Cnat2-28, Cn5-15, Cnat2-31, DL2H-2.2, DL2-H6, DL2-H3, DL2-H1, Of7-H1, Of5-H5, CnMc7-37, Cnat2-41, Cnat2-18, Cnat2-8, Cnat3-28, CnH1-48, and McH1-25 (Table S2). C) The MS2LDA motif enabling annotation of hydrolyzed AHLs is shown. The acyl chain length and saturation of annotated AHLs are labeled on the nodes. Representative MS² spectra of hydrolyzed AHLs is shown in panel B.

Prodigiosin and Analogs. Prodiginines are a class of red pigments with broad anti-inflammatory and antimicrobial properties.[45] These pigments were identified from *Serratia marcescens* and reported to be produced within a select temperature range.[46] A variety of bacterial genera including *Pseudoalteromonas* are known to produce prodigiosin and derivatives.[45, 47] Recently, cycloprodigiosin was also demonstrated as a compound involved in attachment and settlement of coral *Leptastrea purpurea* larvae.[48] Thus, a broad range of biological activities are reported for prodiginines. In our study, features at *m/z*_t_R 324.207_12.1 min and 322.191_11.1 min were identified via GNPS library matches to prodigiosin and cycloprodigiosin, respectively (Figure 6). Additional prodiginine derivatives were identified through propagation of annotation of connected nodes in FBMN (Figure 6 and Figures S6 and S7).[48, 49] The annotated pigments were detected in two isolates of *Pseudoalteromonas* spp., namely DL2-H3 and DL2H-2.2 isolated from *D. labyrinthiformis*. Additionally, the prodiginine annotations allowed for comparison of the extraction efficiency of the wet and lyophilized

methods. The bacterial cell pellets were either left wet or lyophilized overnight to complete dryness prior to extraction with 2:2:1 EtOAc:MeOH:H_2O. The annotated pigments were detected at relatively higher abundances in the lyophilized bacterial extracts (Figure S8). These were not the only examples of metabolites detected at higher abundance in the lyophilized extracts, thus the lyophilization prior to extraction can be incorporated as a standard step in microbial metabolomics workflows. The antimicrobial activities of prodigiosin, 2-methyl-3-hexyl-prodiginine, and cycloprodigiosin were recently compared to various antibiotics and found to have comparable activity against *Escherichia coli*, *Candida albicans, Salmonella typhi*, and *Staphylococcus aureus*.[49] Currently, amoxicillin is being tested in the field to treat SCTLD lesions.[50] The pathogenic *V. coralliilyticus* including the isolate *V. coralliilyticus* OfT6-21 can be resistant to amoxicillin.[11] This strain was isolated from lesions with SCTLD that appeared to progress more rapidly and may be acting as an opportunistic pathogen, exacerbating the detrimental results of SCLTD.[11] Both *Pseudoalteromonas* sp. isolates producing prodiginines were active against the strain OfT6-21 (Table 1). Thus, the limitation of antibiotic resistant pathogens implicated in SCTLD progression may be addressed by the use of probiotic bacteria capable of producing a variety of natural products. The rapid dereplication of natural product profiles will aid in selection of probiotic candidates. The annotation of prodigiosin and its derivatives demonstrates the value of a centralized repository for rapid dereplication of metabolites that contribute towards the bioactivity of candidate probiotic isolates.

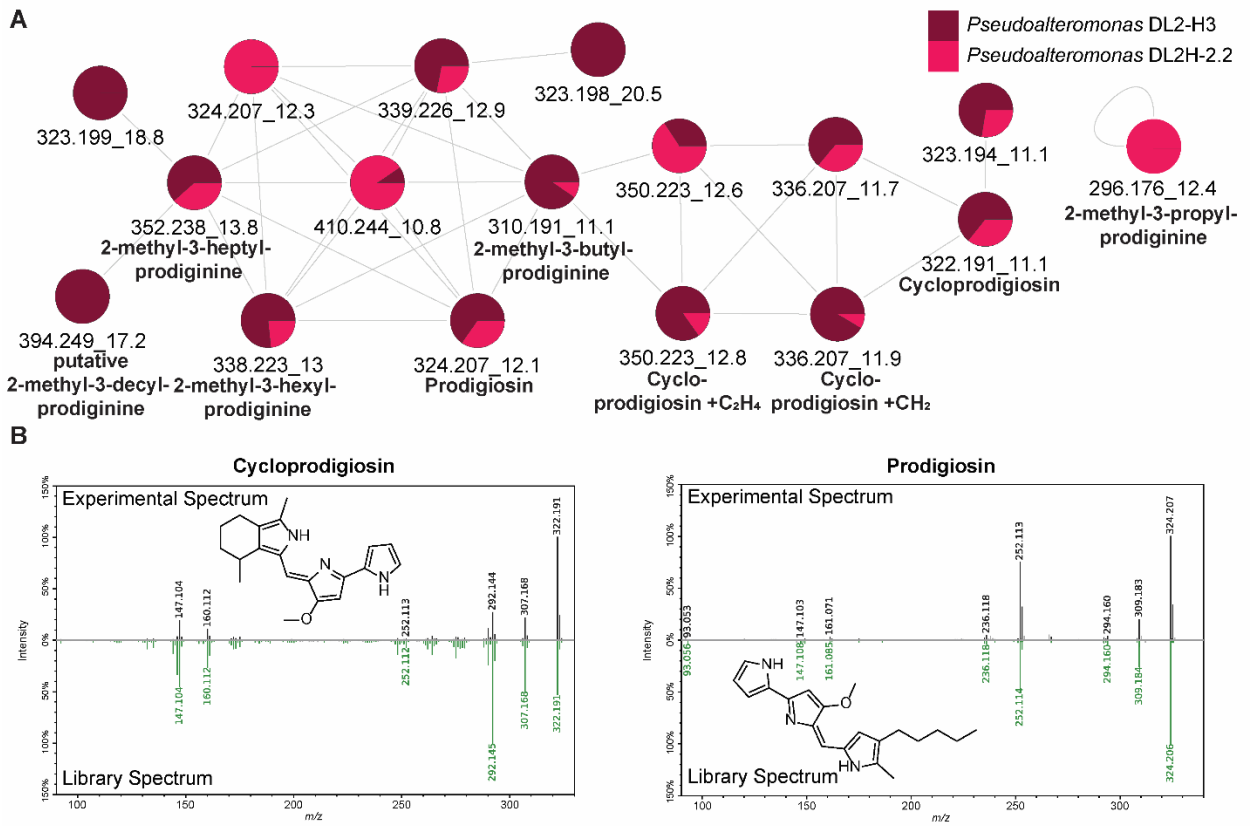

Figure 6. **Prodiginine analogs.** (A) The network containing features annotated as prodiginine analogs is shown. The features at *m/z*_t_R 324.207_12.1 min and 322.191_11.1 min match prodigiosin and cycloprodigiosin MS² spectra available in the GNPS library. The prodiginines were detected in isolates DL2-H3 and DL2H-2.2. (B) The MS² mirror plot of cycloprodigiosin (left) is shown where the MS² spectrum acquired on the extract (top) is compared with the MS² spectrum available in the GNPS library (bottom). The MS² mirror plot of prodigiosin (right) is shown where the MS² spectrum acquired on the extract (top) is compared with the MS² spectrum available in the GNPS library (bottom). Annotation of additional analogs is provided in Figures S6 and S7.

Siderophores. Using FBMN and MS2LDA, various features detected in several of the cultured strains were identified as siderophores. The representative members of these siderophores include promicroferrioxamine C13, desferrioxamine E, escherichelin, bisucaberin, esdehydroxynocardamine, and amphibactin (Figures S9-S12, 7 and 8). As described in supplemental text and Figure S9-S11, use of MS2LDA enabled discovery of a substructure motif containing fragment peaks corresponding to the hydroxamate backbone. Thus, various hydroxamate siderophores were identified using this approach and were found to be produced a

number of bacterial isolates (Table S2), the spectral annotations for which were not observed as database hits.

Interestingly, a siderophore HPTzTnCOOH, previously reported from terrestrial bacteria was observed as a GNPS hit. HPTzTnCOOH, also known as escherichelin and HPTT-COOH, has been previously reported from *Burkholderia* sp. bacteria, *Pseudomonas aeruginosa* and pathogenic *E. coli* from the urinary tract.[51-53] This compound is shown to be a biosynthetic intermediate in biosynthesis of siderophores such as pyochelin and yersiniabactin.[54] This biosynthetic intermediate is known to mediate iron uptake and was reported to compete with pyochelin.[52, 55] We looked for additional biosynthetic precursors, which are also detected during pyochelin biosynthesis.[54] A feature at *m/z*_t_R 210.058_6.5 min was annotated as aerugine, which is also a product of pyochelin biosynthesis,[56] through comparison with published MS^2 spectra[51] (Figure 7C and Figure S10). The feature annotated as HPTzTnCOOH at *m/z*_t_R 307.021_9 min was detected in *Vibrio* spp. (Cn52-H1 and Cn26-H1 from *C. natans,* OfT6-21, Ofav2-48, OfT6-17, and OfT7-21 from *O. faveolata,* and MmMcT2-4 from *M. cavernosa*) and in *Pseudoalteromonas* spp. (DL2-H6 and DL2-H1 from *D. labyrinthiformis*). The feature annotated as aerugine was detected in many of the same *Vibrio* spp. isolates OfT6-21, Cn26-H1, Ofav2-48, OfT6-17, OfT7-21, and MmMcT2-4. Thus, we searched for precedence of pyochelin-type siderophores in marine bacteria. A focused literature search directed us towards production of a structural homolog, piscibactin, in marine bacteria isolated from bacterial sources other than corals.[57] We searched our data for presence of a node representative of piscibactin spectra and found an observed *m/z*_t_R 454.093_8.3 for piscibactin and *m/z*_t_R 507.004_6.9 for ferri-piscibactin (Figure 7C and Figure S11B), which were detected in the same isolates as the biosynthetic precursors HPTzTnCOOH and aerugine. We report aerugine and HPTzTnCOOH as intermediates of piscibactin biosynthesis based on the detection pattern of these metabolites and the proposed piscibactin and yersiniabactin biosynthetic pathway.[57] Piscibactin has been reported from the fish pathogen *Photobacterium damselae* subsp. *piscicida*[57] and from *V.*

neptunius, where production of piscibactin was shown to correlate with pathogenicity of *V. neptunius* towards molluscs.[58] Previous *in silico* analyses of *Vibrio* spp. genomes revealed that *V. coralliilyticus* has the genomic potential to produce piscibactin.[58, 59] Indeed, six of the seven isolates in which piscibactin was detected in this study were identified as *V. coralliilyticus,* two of which are the putative coral pathogens OfT6-21 and OfT7-21. Thus, our data shows that *Vibrio* spp. capable of producing this known virulence factor in molluscs are also present in the microbiomes of multiple species of corals.

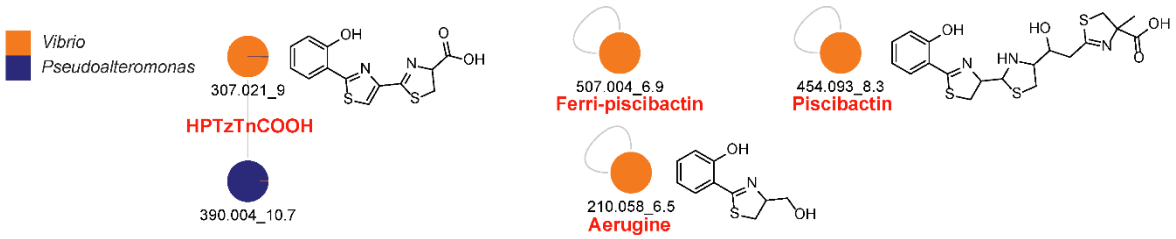

Figure 7. **Annotation of piscibactin and biosynthetic intermediates.** The singleton nodes for piscibactin, biosynthetic intermediates in piscibactin biosynthesis and their chemical structures are shown. These features were detected in isolates OfT6-21, DL2-H6, Cn26-H1, Ofav2-48, OfT6-17, Cn52-H1, OfT7-21, DL2-H1, and MmMcT2-4.

MolDiscovery also enabled the annotation of several amphibactins, a class of amphisiderophores which contain a polar peptidic head linked to an apolar fatty acid tail[60] (Figure 8A). These prevalent siderophores are known to be produced by a variety of *Vibrio* sp., including *V. coralliilyticus*. The node for amphibactin B was detected in several isolates of *Vibrio* spp. (OfT7-21, OfT6-17, OfT6-21, and Ofav2-48 from *O. faveolata*, Cn26-H1 and Cn52-H1 from *C. natans* and MCA32, MmMcT2-4, and MCA25 from *M. cavernosa*). We annotated another cluster of nodes as amphibactins as it was linked to the amphibactin B cluster via substructure motifs 459 and 528 from MS2LDA (Figure 8B). The motif 528 contains several peaks that correspond to *b* and *y* fragments corresponding to the peptidic head group of amphibactin. These annotations were further supported through comparison with published MS^2 spectra for several amphibactins (Figure S12).[61] In addition, we propagated annotations to features that

clustered with known amphibactins (Figure 8, labeled as Δ*m/z* on network edge). These novel analogs contained modifications within the acyl chain, with tail lengths varying between C14-C18. The widespread occurrence of the amphibactin gene cluster was reported by Galvis *et al*. in mollusc pathogens and hemolymph microbiota. This observation is further supported in our study by the detection of these siderophores in extracts of coral-derived *Vibrio* spp.

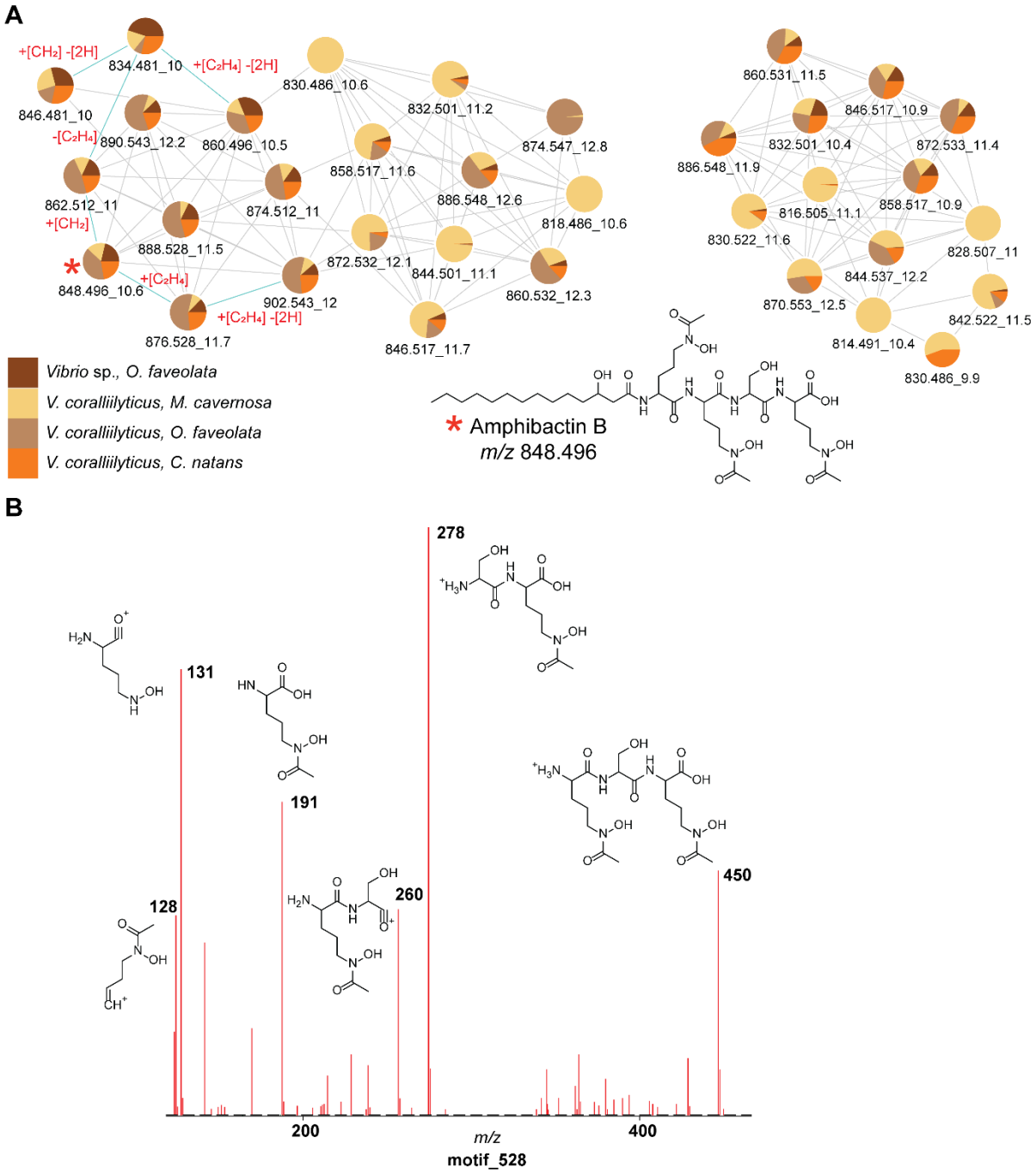

Figure 8. **Description of amphibactin molecular cluster aided by MolDiscovery-based *in silico* annotation and substructure discovery via MS2LDA.** (A) Two network clusters, where putative annotations as amphibactins were derived using MolDiscovery. These features were detected in *Vibrio* spp. isolates including OfT7-21, OfT6-17, OfT6-21, Ofav2-48, Cn26-H1, Cn52-H1, MCA32, MmMcT2-4, and MCA25. Dereplication was further supported through comparison with published MS^2 spectra as shown in Figure S12. (B) The spectrum for MS2LDA motif 528 which is shared between the network clusters in (A) is shown. The motif consists of several fragment peaks that correspond to *b* and *y* ions of the amphibactin peptidic head group.

Conclusion. With the unprecedented increase in coral reef diseases, it is well appreciated that innovative approaches aimed at disease prevention and coral reef restoration are needed. In this regard, bacteria associated with the mucus of healthy coral colonies were shown to inhibit the growth of potential pathogens.[3] The role of beneficial bacteria has also been linked to observed resistance or resilience to heat stress supporting their use as probiotics for microbiome manipulation.[14, 62, 63] The selection of probiotic candidates is typically envisaged through metagenome sequencing to decipher presence of beneficial genomic islands, or through disease transmission studies in aquaria. With the advancement in MS instruments and bioinformatics, metabolomics-based approaches can serve as complementary approaches in selection of probiotic candidates when coupled with bioactivity screening of cultured bacteria. The field of microbiology has the long-standing 'great plate anomaly'[64] where only approximately <10% of present microbes within a sample will grow on a particular growth medium, suggesting the need for a variety of culture conditions for any one sample to obtain a more complete survey of the present bacteria and evaluate their role as potential probiotics. In this regard, a literature-based review has suggested that by the application of a variety of culture conditions, taxonomically diverse bacteria can be isolated from the coral holobiont.[18] Herein, we highlight that by application of metabolomics profiling such as UpSet plots, one can identify metabolically diverse bacteria with both probiotic and pathogenic potential. In our case, the isolate Cn52-H1 with the largest number of unique metabolites from a known pathogenic genus, *Vibrio,* was identified as the only strain that completely inhibited the *Leisingera* sp. isolate, a potential

pathogen isolated from diseased *M. cavernosa*. The *Pseudoalteromonas* sp. isolate Cn2-18 with the largest number of unique metabolites was also established as a potential probiotic in aquaria-based disease transmission studies.[26] The natural product chemistry of this strain remains to be established. Using various recently developed tools, we describe strategies for dereplication of natural products. One such tool, namely MolDiscovery enabled annotation of several natural products such as andrimid, moiramide, alteramide, amphibactin, and ikarugamycin. The spectral matches to experimental spectra were not observed despite the availability of experimental spectra for some of these compounds in natural product spectral databases such as GNPS. These annotations are supported via the presence of biosynthetic gene clusters in the genome of the isolated strain, comparison with experimental spectra available in supplementary material of published work or through spectral matching with analytical standards when available. These observations support that metabolomics profiling and natural product annotations via both *in silico* and molecular network-based approaches that employ experimental spectra are indeed applicable in guiding selection of potential probiotic candidates, especially as these tools are further refined with availability of advanced training datasets, availability of experimental spectra, and advancement in scoring algorithms. Many of these natural products are known to be produced by several marine bacteria isolated from marine sediments or from organisms such as tunicates and marine sponges, highlighting that these studies will provide insights into the chemical ecology of natural products across complex marine holobionts. As we inventory coral microbiome-produced natural products and populate spectral databases such as GNPS, their detection in organic extracts generated directly from coral mucus and tissue will become feasible. Such studies will allow one to directly link and validate ecological functions and prevalence of microbial natural products in coral host derived samples. However, a larger fraction of coral host and algal endosymbiont biomass as compared to the bacterial biomass affects the limit of detection of bacterial natural products and presents an analytical challenge in metabolomics studies. Such challenges can be overcome through

development of extraction strategies whereby fractionation of bacterial cells and depletion of higher-abundance lipids can be employed to improve limits of detection of microbial natural products by mass spectrometry. In conjunction, microbial genomes assembled from metagenomes, referred to as MAGs (metagenome-assembled genomes), can be mined for the presence of biosynthetic gene clusters and correlated to disease phenotypes.

Experimental Section

Coral Collection and Sampling. Coral colonies with SCTLD lesions were identified, collected, and sampled from various locations throughout the Florida Keys and in Broward County (near Ft. Lauderdale), FL under appropriate state and federal permits from 2017-2020 (Table S1). Apparently healthy corals were obtained from NOAA Key West Nursery, in Biscayne National Park, and outside of Dry Tortugas National Park. The coral species sampled from Florida in this study include *M. cavernosa* (Mcav, Mc), *C. natans* (Cnat, Cn), *Pseudodiploria strigosa* (PSTR, Ps), *O. faveolata* (Ofav, Of), and *D. labyrinthiformis* (Dlab, DL). Approximately 18% of collected strains reported in this study were derived from nursery-held corals, 27% of strains were derived from apparently healthy corals from Dry Tortugas National Park, and 27% of strains were derived from apparently healthy corals from Biscayne Bay National Park (Table S1). The mucus and tissue samples were obtained using a sterile 30 mL syringe as described previously.[65] Briefly, the tip was used to agitate the area of the coral (healthy tissue site on apparently healthy corals; lesion site on diseased corals) while drawing in the mucus and tissue. Each 15 mL sample was transferred into a separate sterile conical tube for further studies.

Bacteria Growth Conditions. All bacteria in this study were grown on a seawater-based medium. Seawater broth (SWB) was prepared using 4 g/L of tryptone (Fisher, cat# BP1421-500) and 2 g/L of yeast extract (Fisher, cat# BP1422-500) in filtered seawater (FSW, 0.22 µm pore) with an addition of 15 g/L agar for seawater agar (SWA) (Teknova Supplier, cat#

50-841-063).[11] FSW was collected and prepared as stated in previous work.[11] All cultures were incubated at 28 °C.

Bacterial Isolation and Selection from Coral Sources. The coral mucus and tissue samples (15 mL) were vortexed vigorously for 2-3 min. A total of seven 10-fold serial dilutions of these samples were prepared in autoclaved FSW. A 50 µL aliquot of each dilution was plated using sterile glass beads onto SWA followed by incubation at 28 °C for 72 h. Following incubation, individual bacterial colonies with unique colony morphologies were streaked on SWA. The isolated colonies were streaked 2-3 times to ensure purity.

Metabolite Extraction. For extraction of metabolites, bacterial isolates were cultured in sterile 26.4 X 33.5 X 16.0 cm^3 aluminum pans containing 300 mL SWA and in 1 L SWB in 2 L culture flasks for 24-48 h. The bacterial cells from aluminum pans were scraped from the surface of the agar with a sterile tongue depressor (Fisher, cat# 22-363-154) and frozen. The liquid cultures were centrifuged at 11,440 x g for 10 min. The supernatant was discarded and the cells were frozen and extracted as described below. Half of the cellular fraction collected from solid cultures was lyophilized to complete dryness. The lyophilized and wet cell pellets were extracted three times with ~14 mL of a 2:2:1 mixture of EtOAc:MeOH:H_2O at 25 °C in prewashed 20 mL vials. Extractions were sonicated for 30-60 seconds to visibly disrupt the cell pellet, allowed to extract for 5-8 hours, and the extract was transferred with a glass pipet into prewashed 20 ml scintillation vials and concentrated on a ThermoSci Savant SpeedVac at 40 °C for 3-4 hours until all EtOAc was removed. Methanol:water mixture (3:1) was added to the vials, and the extract was transferred to a sterile 15 ml polypropylene conical. Cellular debris was pelletized via centrifugation at 4500 x g for 10 minutes and remaining supernatant was concentrated via rotary evaporation at 40° C. The final extract was transferred in MeOH to a prewashed 7 ml scintillation vial and concentrated to dryness on the SpeedVac.

Antimicrobial Assay. The growth inhibition activity of bacterial extracts was tested against putative coral pathogens, *V. coralliilyticus* OfT6-21 (GenBank Accession:JABSMZ010000000) and *V. coralliilyticus* OfT7-21 (GenBank Accession: JABSMY010000000)[11] isolated from diseased *O. faveolata*, and *Leisingera* sp. McT4-56 (GenBank Accession:JAJALF000000000) isolated from a diseased *M. cavernosa*, where both corals were apparently undergoing SCTLD. Extracts were solubilized in MeOH at a concentration of 31.25 mg/mL. A 4 µL (125 µg) aliquot of this extract was placed on a sterile paper disk in triplicate (Fisher, Whatman, cat#1001-125). A sterile paper disk with 4 µL methanol was used as a solvent control. A filter disc with 4 µL of nalidixic acid at 15.62 mg/mL (62.5 µg) was used as a positive control. Liquid molten SWA was cooled and seeded with 200 µL of each target strain grown to an optical density of 0.5. Plates were incubated for 24 h at 28 °C and antibacterial activity was determined in triplicate by measuring zones of inhibition (ZOIs). Zones with complete inhibition were clear while partial zones indicated growth within the zone of inhibition, forming a halo around the disk (Figure S1A).

Mass Spectrometry Data Acquisition and Analyses. The dried extracts were resuspended in 100% MeOH (Fisher Chemical, LC/MS grade, cat#A452-4) containing 1 µM of sulfadimethoxine as an internal standard. The samples were analyzed with an Agilent 1290 Infinity II UHPLC system (Agilent Technologies) using a Kinetex 1.7 µm C18 reversed-phase UHPLC column (50×2.1 mm, Phenomenex, cat#00B-4475-AN) coupled to an ImpactII ultrahigh resolution Qq-TOF mass spectrometer (Bruker Daltonics) equipped with an ESI source for MS/MS analysis. MS spectra were acquired in positive mode with an *m/z* range of 50–2000 Da. The eight most intense ions per MS^1 spectra were selected for further acquisition of MS^2 data. An active exclusion of two spectra was used, implying that an MS^1 ion would not be selected for fragmentation after two consecutive MS^2 spectra had been recorded for it in a 0.5 min time window. The exclusion was reconsidered and additional MS^2 spectra was acquired if five-fold

enhancement in intensity was observed. The chromatography solvent A: H_2O (Fisher Chemical, LC/MS Grade, cat#W6-4) + 0.1% v/v formic acid (Fisher Scientific, LC/MS, cat#A117-50) and solvent B: MeCN (Fisher Chemical, LC/MS Grade, cat#A955-4) + 0.1% v/v formic acid were employed for separation. Flow rate was held constant at 0.5 mL/min throughout the run. The gradient applied for chromatographic separation was 5% solvent B and 95% solvent A for 3 min, a linear gradient of 5% B–95% B over 17 min, held at 95% B for 3 min, 95% B–5% B in 1 min, and held at 5% B for 1 min, 5% B-95% B in 1 min, held at 95% B for 2 min, 95% B-5% B in 1 min, then held at 5% B for 2.5 min at a flow rate of 0.5 mL/min throughout. Following acquisition of data on 12 samples, data on a mixture of six-compounds (amitryptiline, sulfamethazine, sulfamethizole, sulfachloropyridazine, sulfadimethoxine, coumarin-314) was acquired as quality control step to ensure consistent instrument and column performance.

The raw data was converted to mzXML format using vendor software. Metabolite features were extracted using MZmine 2.53 to carry out steps for mass detection, chromatogram building, chromatogram deconvolution, isotopic grouping, retention time alignment, duplicate removal, and missing peak filling.[66] This processed data was submitted to the feature-based molecular networking (FBMN) workflow on the Global Natural Product Social Molecular Networking (GNPS) platform.[21] The output of MZmine includes information about LC-MS features across all samples containing the *m/z* value of each feature, retention time of each feature, the area under the peak for the corresponding chromatogram of each feature, and a unique identifier. The MS^2 spectral summary contains a list of MS^2 spectra, with one representative MS^2 spectrum per feature. The mapping information between the feature quantification table and the MS^2 spectral summary was stored in the output using the unique feature identifier and scan number, respectively. This information was used to relate LC-MS feature information to the molecular network nodes. The quantification table (.csv file) and the linked MS^2 spectra (.mgf file) were exported using the GNPS export module[66, 67] and the SIRIUS

4.0 export module.[66, 68] Feature Based Molecular Networking was performed using the MS^2 spectra (.mgf file) and the quantification table (.csv file).

The molecular network was generated as follows; the data was filtered by removing all MS^2 fragment ions within +/- 17 Da of the precursor *m/z*. MS^2 spectra were filtered by choosing only the top 6 fragment ions in the +/- 50 Da window throughout the spectrum. The precursor ion mass tolerance was set to 0.02 Da and the MS^2 fragment ion tolerance to 0.02 Da. A molecular network was then created where edges were filtered to have a cosine score above 0.7 and more than four matched peaks. Further, edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's respective top 10 most similar nodes. Finally, the maximum size of a molecular family was set to 100, and the lowest scoring edges were removed from molecular families until the molecular family size was below this threshold. The spectra in the network were then searched against GNPS spectral libraries.[21] The library spectra were filtered in the same manner as the input data. All matches kept between network spectra and library spectra were required to have a score above 0.7 and at least four matched peaks. The molecular network is available at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=559213a2d954436587516d9bf4382470. The workflow for DEREPLICATOR was also run during generation of the molecular network, the output is available at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=4e724351435a46ff8fbcf89754cf39e6.

The molecular network was visualized using Cytoscape v3.7.2.[69] The compound annotations follow the "level 2" annotation standard based upon spectral similarity with public spectral libraries or spectra published in the literature as proposed by the Metabolomics Society Standard Initiative.[70] Annotations were confirmed with commercial standards when available. The commercial analytical standards for ikarugamycin (cat#15386) and C18:1-AHL (cat#10012674) were purchased from Cayman Chemical Company. The MS2LDA analysis was

performed as previously described with default parameters.[51, 71] The output of this analysis is available at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=f06640039c744b39acf2402ab0e5e1e6. Prior to statistical analysis, blank subtraction was performed using a Jupyter notebook, available at https://github.com/Garg-Lab/Jupyter-Notebook-Blank-Subtraction/blob/main/Probiotic%20Extracts%205.29.21%20pos%20mode.ipynb. The mean area under the curve of each feature in the samples is compared to the mean of the feature in the solvent controls, media blanks, and LC-MS blanks. A feature is retained if its sample mean is greater than 0.25 × the mean of the sample in the blanks. The entire quantification table is exported, with each feature marked as "true" (feature is retained) or "false" (feature is not retained). Visualization of the molecular network, metadata mapping, and feature filtering was performed using Cytoscape v 3.7.2.[69] Within Cytoscape, the nodes corresponding to features present in solvent controls, media blanks, and LC-MS blanks were removed based on the exported quantification table from the Jupyter notebook. Hierarchical clustering analysis (HCA) was performed on PLS_Toolbox 8.8.1 using the Ward algorithm for clustering and pareto scaling.[72] UpSet plots were generated using the Intervene app.[73]

Additional compound annotations were performed by searching entries in databases such as The Natural Product Atlas[74] and MarinLit,[75] and by the use of *in silico* tools such as MolDiscovery (v.1.0.0),[76] SIRIUS 4.9.3[68] with CSI:FingerID[77] and CANOPUS[19] and through literature searches for compounds known to be produced by marine bacteria for untargeted analysis. MolDiscovery, available through the GNPS platform, compares *in silico* generated mass spectra of small molecules from a variety of databases with experimental MS^2 spectra and includes a similarity score for each reported match.[76] The .mgf file generated from Mzmine on the positive mode data was submitted to the MolDiscovery workflow. Proposed annotations were further supported either through comparison with published MS^2 spectra or through

analytical standards. The MolDiscovery job is available at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=814f6338cd0140528bba145e690c23c9. Compound annotations using the *in silico* tool CSI:FingerID were performed using default settings. The fragmentation trees were computed for submitted LC-MS features and predicted chemical formulas and structures that best matched the experimental spectra were ranked using machine learning.[68] The compound classes were assigned using CANOPUS, which uses MS/MS spectra and the chemical structures predicted by CSI:FingerID to propose chemical classes for a feature.[78, 79] The compound annotations were proposed at the confidence level 2 as recommended by the metabolomics standard initiative.[70] The MS^2 spectra of all of the proposed annotations was compared with MS^2 spectra available in literature and/or in mass spectral databases. All annotations reported in the manuscript were detected at <3 ppm mass error (Table S2).

Bacterial Genome Sequencing and Construction of Phylogenomic Trees. Genomic DNA was extracted from cultured probiotic strains and sequenced as previously described[11] on an Illumina Miseq with the 2 x 150 bp v. 2 cycle at the University of Florida Interdisciplinary Center for Biotechnology Research. Raw sequencing reads were quality-filtered with the Minoche filtering pipeline[80] in illumina-utils v. 2.3[81] and Illumina adapters and Nextera transposase sequences were removed with cutadapt v. 1.8.1.[82] Quality-filtered reads were assembled with SPAdes v. 3.13.0.[83] The sequenced genomes were uploaded to GenBank and accession numbers are reported in Table S1. A phylogenomic tree was created with anvio v. 6.2[84] from an alignment of 71 concatenated bacterial single-copy genes for bacterial strains with sequenced genomes. All bacterial strains described in this study were submitted to Genewiz, Inc. for 16S rRNA gene sequencing. For the phylogenetic tree based on 16S rRNA gene sequencing containing all isolates in this study, the phylogenetic history was inferred by using the Maximum Likelihood method and Tamura-Nei model.[85] The bootstrap consensus tree

inferred from 1000 replicates is taken to represent the phylogenetic history of the taxa analyzed.[86] Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches.[86] Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model, and then selecting the topology with superior log likelihood value. This analysis involved 37 nucleotide sequences. There were a total of 1893 positions in the final dataset. Evolutionary analyses were conducted in MEGA X.[87]

Acknowledgements

We thank T. DeMarco for assistance with bacterial culturing, S. Thompson and R. Wetzell for bacterial isolations, K. Pitts for coral collection and data management, J. Sneed and W. Lee and other Smithsonian personnel for laboratory support, and T. Vekich and K. Zimmerman for coral care. Many people assisted with coral collections and provided field assistance, including G. Aeby, O. Carmack, A. Looby, E. Bartels and other Mote Marine Laboratory staff, K. Banks and A. Rovira and other staff at Broward County Environmental Protection and Growth Management Department, and A. Bourque and other staff at Biscayne National Park. We thank the Florida Fish and Wildlife Conservation Commission for permitting this activity under the State of Florida Special Activity License Permits SAL-17-1702-SRP, SAL-18-1702-SRP, SAL-19-1702-SRP, the Florida Keys National Marine Sanctuary permits # FKNMS-2017-128-A2 and FKNMS-2019-160 and the National Park Service permit # BISC-2020-SCI-0003 for collections in Biscayne National Park. This is contribution #1170 of the Smithsonian Marine Station.

Funding

This study was supported by the NSF CAREER award to NG (award number 2047235), NSF RAPID award IOS-1728002 (VP), Environmental Protection Agency to NG, VP, and BU (award

number X7-01D00120 -0), the Florida Department of Environmental Protection Office of Resilience and Coastal Protection-Southeast Region awards to VP, BU and JM and the Smithsonian Hunterdon Oceanographic Fund. A research cruise to the Dry Tortugas National Park in January 2020 was also funded by the Florida Department of Environmental Protection Office of Resilience and Coastal Protection-Southeast Region. JMD is supported through NSF CAREER, awarded to NG.

References

1. Stévenne, C.; Micha, M.; Plumier, J.-C.; Roberty, S. Corals and Sponges Under the Light of the Holobiont Concept: How Microbiomes Underpin Our Understanding of Marine Ecosystems. *Frontiers in Marine Science* **2021**, 8:698853, DOI: 10.3389/fmars.2021.698853.

2. Webster, N. S.; Reusch, T. B. H. Microbial Contributions to the Persistence of Coral Reefs. *The ISME Journal* **2017,** *11* (10), 2167-2174, DOI: 10.1038/ismej.2017.66.

3. Ritchie, K. B. Regulation of Microbial Populations by Coral Surface Mucus and Mucus-Associated Bacteria. *Marine Ecology Progress Series* **2006,** *322*, 1-14, DOI: 10.3354/meps322001.

4. Gibbin, E.; Gavish, A.; Krueger, T.; Kramarsky-Winter, E.; Shapiro, O.; Guiet, R.; Jensen, L.; Vardi, A.; Meibom, A. *Vibrio coralliilyticu*s Infection Triggers a Behavioural Response and Perturbs Nutritional Exchange and Tissue Integrity in a Symbiotic Coral. *The ISME Journal* **2019,** *13* (4), 989-1003, DOI: 10.1038/s41396-018-0327-2.

5. Rohwer, F.; Seguritan, V.; Azam, F.; Knowlton, N. Diversity and Distribution of Coral-Associated Bacteria. *Marine Ecology Progress Series* **2002,** *243*, 1-10, DOI: 10.3354/meps243001.

6. Magnusson, S. H.; Fine, M.; Kuhl, M. Light Microclimate of Endolithic Phototrophs in the Scleractinian Corals *Montipora monasteriata* and *Porites cylindrica*. *Marine Ecology Progress Series* **2007,** *332*, 119-128, DOI: 10.3354/meps332119.

7. Reshef, L.; Koren, O.; Loya, Y.; Zilber-Rosenberg, I.; Rosenberg, E. The Coral Probiotic Hypothesis. *Environmental Microbiology* **2006,** *8* (12), 2068-2073, DOI: 10.1111/j.1462-2920.2006.01148.x.

8. Pogoreutz, C.; Voolstra, C. R.; Rädecker, N.; Weis, V.; Cárdenas, A.; Raina, J.-B., The Coral Holobiont Highlights the Dependence of Cnidarian Animal Hosts on Their Associated Microbes. In *Cellular Dialogues in the Holobiont*, CRC Press: 2020; pp 91-118.

9. Matthews, J. L.; Raina, J.-B.; Kahlke, T.; Seymour, J. R.; van Oppen, M. J. H.; Suggett, D. J. Symbiodiniaceae-Bacteria Interactions: Rethinking Metabolite Exchange in Reef-Building Corals as Multi-Partner Metabolic Networks. *Environmental Microbiology* **2020,** *22* (5), 1675-1687, DOI: 10.1111/1462-2920.14918.

10. Meyer, J. L.; Castellanos-Gell, J.; Aeby, G. S.; Häse, C. C.; Ushijima, B.; Paul, V. J. Microbial Community Shifts Associated with the Ongoing Stony Coral Tissue Loss Disease Outbreak on the Florida Reef Tract. *Frontiers in Microbiology* **2019**, 10:2244, DOI: 10.3389/fmicb.2019.02244.

11. Ushijima, B.; Meyer, J. L.; Thompson, S.; Pitts, K.; Marusich, M. F.; Tittl, J.; Weatherup, E.; Reu, J.; Wetzell, R.; Aeby, G. S.; Häse, C. C.; Paul, V. J. Disease Diagnostics and Potential Coinfections by *Vibrio coralliilyticus* During an Ongoing Coral Disease Outbreak in Florida. *Frontiers in Microbiology* **2020**, 11:569354, DOI: 10.3389/fmicb.2020.569354.

12. Vizcaino, M. I.; R, J. W.; Kimes, N. E.; Williams, K.; Torralba, M.; Nelson, K. E.; Smith, G. W.; Weil, E.; Moeller, P. D. R.; Morris, P. J. Antimicrobial Resistance of the Coral Pathogen *Vibrio coralliilyticus* and Caribbean Sister Phylotypes Isolated from a Diseased Octocoral. *Microbial Ecology* **2010,** *59* (4), 646-657, DOI: 10.1007/s00248-010-9644-3.

13. Gignoux-Wolfsohn, S. A.; Aronson, F. M.; Vollmer, S. V. Complex Interactions Between Potentially Pathogenic, Opportunistic, and Resident Bacteria Emerge During Infection on a Reef-Building Coral. *FEMS Microbiology Ecology* **2017,** *93* (7), fix080, DOI: 10.1093/femsec/fix080.

14. Rosales, S. M.; Miller, M. W.; Williams, D. E.; Traylor-Knowles, N.; Young, B.; Serrano, X. M. Microbiome Differences in Disease-Resistant vs. Susceptible *Acropora* Corals Subjected to Disease Challenge Assays. *Scientific Reports* **2019,** *9* (1), 18279, DOI: 10.1038/s41598-019-54855-y.

15. Kelley, E. R.; Sleith, R. S.; Matz, M. V.; Wright, R. M. Gene Expression Associated with Disease Resistance and Long-term Growth in a Reef-Building Coral. *Royal Society Open Science* **2021,** *8* (4), 210113, DOI: 10.1098/rsos.210113.

16. Muller, E. M.; Bartels, E.; Baums, I. B. Bleaching Causes Loss of Disease Resistance Within the Threatened Coral Species *Acropora cervicornis*. *eLife* **2018,** *7*, e35066, DOI: 10.7554/eLife.35066.

17. MacKnight, N. J.; Cobleigh, K.; Lasseigne, D.; Chaves-Fonnegra, A.; Gutting, A.; Dimos, B.; Antoine, J.; Fuess, L.; Ricci, C.; Butler, C.; Muller, E. M.; Mydlarz, L. D.; Brandt, M. Microbial Dysbiosis Reflects Disease Resistance in Diverse Coral Species. *Communications Biology* **2021,** *4* (1), 679, DOI: 10.1038/s42003-021-02163-5.

18. Sweet, M.; Villela, H.; Keller-Costa, T.; Costa, R.; Romano, S.; Bourne, D. G.; Cárdenas, A.; Huggett, M. J.; Kerwin, A. H.; Kuek, F.; Medina, M.; Meyer, J. L.; Müller, M.; Pollock, F. J.; Rappé, M. S.; Sere, M.; Sharp, K. H.; Voolstra, C. R.; Zaccardi, N.; Ziegler, M.; Peixoto, R.; Bouskill, N.; Speare, L.; Shore, A. Insights into the Cultured Bacterial Fraction of Corals. *mSystems* **2021,** *6* (3), e01249-20, DOI: 10.1128/mSystems.01249-20.

19. Dührkop, K.; Nothias, L.-F.; Fleischauer, M.; Reher, R.; Ludwig, M.; Hoffmann, M. A.; Petras, D.; Gerwick, W. H.; Rousu, J.; Dorrestein, P. C.; Böcker, S. Systematic Classification of Unknown Metabolites Using High-Resolution Fragmentation Mass Spectra. *Nature Biotechnology* **2021,** *39* (4), 462-471, DOI: 10.1038/s41587-020-0740-8.

20. Reher, R.; Kim, H. W.; Zhang, C.; Mao, H. H.; Wang, M.; Nothias, L.-F.; Caraballo-Rodriguez, A. M.; Glukhov, E.; Teke, B.; Leao, T.; Alexander, K. L.; Duggan, B. M.; Van Everbroeck, E. L.; Dorrestein, P. C.; Cottrell, G. W.; Gerwick, W. H. A Convolutional Neural Network-Based Approach for the Rapid Annotation of Molecularly Diverse Natural Products. *Journal of the American Chemical Society* **2020,** *142* (9), 4114-4120, DOI: 10.1021/jacs.9b13786.

21. Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W.-T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya P, C. A.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.; Nothias, L.-F.; Alexandrov, T.; Litaudon, M.; Wolfender, J.-L.;

Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D.-T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. Ø.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nature Biotechnology* **2016,** *34* (8), 828-837, DOI: 10.1038/nbt.3597.

22. Briand, E.; Bormans, M.; Gugger, M.; Dorrestein, P. C.; Gerwick, W. H. Changes in Secondary Metabolic Profiles of *Microcystis aeruginos*a Strains in Response to Intraspecific Interactions. *Environmental Microbiology* **2016,** *18* (2), 384-400, DOI: 10.1111/1462-2920.12904.

23. Boudreau, P. D.; Monroe, E. A.; Mehrotra, S.; Desfor, S.; Korobeynikov, A.; Sherman, D. H.; Murray, T. F.; Gerwick, L.; Dorrestein, P. C.; Gerwick, W. H. Expanding the Described Metabolome of the Marine Cyanobacterium *Moorea producens* JHB through Orthogonal Natural Products Workflows. *PLoS One* **2015,** *10* (7), e0133297, DOI: 10.1371/journal.pone.0133297.

24. Jarmusch, A. K.; Wang, M.; Aceves, C. M.; Advani, R. S.; Aguirre, S.; Aksenov, A. A.; Aleti, G.; Aron, A. T.; Bauermeister, A.; Bolleddu, S.; Bouslimani, A.; Caraballo Rodriguez, A. M.; Chaar, R.; Coras, R.; Elijah, E. O.; Ernst, M.; Gauglitz, J. M.; Gentry, E. C.; Husband, M.; Jarmusch, S. A.; Jones, K. L.; Kamenik, Z.; Le Gouellec, A.; Lu, A.; McCall, L.-I.; McPhail, K. L.; Meehan, M. J.; Melnik, A. V.; Menezes, R. C.; Montoya Giraldo, Y. A.; Nguyen, N. H.; Nothias, L. F.; Nothias-Esposito, M.; Panitchpakdi, M.; Petras, D.; Quinn, R. A.; Sikora, N.; van der Hooft, J. J. J.; Vargas, F.; Vrbanac, A.; Weldon, K. C.; Knight, R.; Bandeira, N.; Dorrestein, P. C. ReDu: a Framework to Find and Reanalyze Public Mass Spectrometry Data. *Nature Methods* **2020,** *17* (9), 901-904, DOI: 10.1038/s41592-020-0916-7.

25. Rosales, S. M.; Clark, A. S.; Huebner, L. K.; Ruzicka, R. R.; Muller, E. M. *Rhodobacterales* and *Rhizobiales* are Associated with Stony Coral Tissue Loss Disease and Its Suspected Sources of Transmission. *Frontiers in Microbiology* **2020**, 11:681, DOI: 10.3389/fmicb.2020.00681.

26. Paul, V. J.; Pitts, K. A.; Mandelare-Ruiz, P.; De La Flor, Y.; Deutsch, J. M.; Garg, N. *DEP Final Report 2021*; Development of Alternative In Situ Treatments for Stony Coral Tissue Loss Disease, 2021. available at https://floridadep.gov/rcp/coral/documents/development-alternative-situ-treatments-stony-coral-tissue-loss-disease

27. Evans, B. S.; Chen, Y.; Metcalf, W. W.; Zhao, H.; Kelleher, N. L. Directed Evolution of the Nonribosomal Peptide Synthetase AdmK Generates New Andrimid Derivatives *In Vivo*. *Chemistry & Biology* **2011,** *18* (5), 601-607, DOI: 10.1016/j.chembiol.2011.03.008.

28. Jin, M.; Fischbach, M. A.; Clardy, J. A Biosynthetic Gene Cluster for the Acetyl-CoA Carboxylase Inhibitor Andrimid. *Journal of the American Chemical Society* **2006,** *128* (33), 10660-10661, DOI: 10.1021/ja063194c.

29. Fredenhagen, A.; Tamura, S. Y.; Kenny, P. T.; Komura, H.; Naya, Y.; Nakanishi, K.; Nishiyama, K.; Sugiura, M.; Kita, H. Andrimid, a New Peptide Antibiotic Produced by an Intracellular Bacterial Symbiont Isolated from a Brown Planthopper. *Journal of the American Chemical Society* **1987,** *109* (14), 4409-4411.

30. Needham, J.; Kelly, M. T.; Ishige, M.; Andersen, R. J. Andrimid and Moiramides A-C, Metabolites Produced in Culture by a Marine Isolate of the Bacterium *Pseudomonas fluorescens*: Structure Elucidation and Biosynthesis. *The Journal of Organic Chemistry* **1994,** *59* (8), 2058-2063, DOI: 10.1021/jo00087a020.

31. Oclarit, J. M.; Okada, H.; Ohta, S.; Kaminura , K.; Yamaoka, Y.; Iizuka, T.; Miyashiro, S.; Ikegami, S. Anti-Bacillus Substance in the Marine Sponge, *Hyatella* Species, Produced by an Associated *Vibrio* Species Bacterium. *Microbios* **1994,** *78* (314), 7-16.

32. Pohlmann, J.; Lampe, T.; Shimada, M.; Nell, P. G.; Pernerstorfer, J.; Svenstrup, N.; Brunner, N. A.; Schiffer, G.; Freiberg, C. Pyrrolidinedione Derivatives as Antibacterial Agents with a Novel Mode of Action. *Bioorganic & Medicinal Chemistry Letters* **2005,** *15* (4), 1189-1192, DOI: 10.1016/j.bmcl.2004.12.002.

33. Buijs, Y.; Isbrandt, T.; Zhang, S.-D.; Larsen, T. O.; Gram, L. The Antibiotic Andrimid Produced by *Vibrio coralliilyticus* Increases Expression of Biosynthetic Gene Clusters and Antibiotic Production in *Photobacterium galatheae*. *Frontiers in Microbiology* **2020**, 11:622055, DOI: 10.3389/fmicb.2020.622055.

34. Wietz, M.; Mansson, M.; Gotfredsen, C. H.; Larsen, T. O.; Gram, L. Antibacterial Compounds from Marine *Vibrionaceae* Isolated on a Global Expedition. *Marine Drugs* **2010,** *8* (12), 2946-2960, DOI: 10.3390/md8122946.

35. Long, R. A.; Rowley, D. C.; Zamora, E.; Liu, J.; Bartlett, D. H.; Azam, F. Antagonistic Interactions among Marine Bacteria Impede the Proliferation of *Vibrio cholerae*. *Applied and Environmental Microbiology* **2005,** *71* (12), 8531-8536, DOI:10.1128/AEM.71.12.8531-8536.2005.

36. Zhang, G.; Zhang, W.; Zhang, Q.; Shi, T.; Ma, L.; Zhu, Y.; Li, S.; Zhang, H.; Zhao, Y. L.; Shi, R.; Zhang, C. Mechanistic Insights into Polycycle Formation by Reductive Cyclization in Ikarugamycin Biosynthesis. *Angewandte Chemie (International ed. in English)* **2014,** *53* (19), 4840-4844, DOI: 10.1002/anie.201402078.

37. Antosch, J.; Schaefers, F.; Gulder, T. A. M. Heterologous Reconstitution of Ikarugamycin Biosynthesis in *E. coli*. *Angewandte Chemie (International ed. in English)* **2014,** *53* (11), 3011-3014, DOI: 10.1002/anie.201310641.

38. Jomon, K.; Kuroda, Y.; Ajisaka, M.; Sakai, H. A New Antibiotic, Ikarugamycin. *The Journal of Antibiotics* **1972,** *25* (5), 271-280, DOI: 10.7164/antibiotics.25.271.

39. Lacret, R.; Oves-Costales, D.; Gómez, C.; Díaz, C.; de la Cruz, M.; Pérez-Victoria, I.; Vicente, F.; Genilloud, O.; Reyes, F. New Ikarugamycin Derivatives with Antifungal and Antibacterial Properties from *Streptomyces zhaozhouensis*. *Marine Drugs* **2014,** *13* (1), 128-140, DOI: 10.3390/md13010128.

40. Shigemori, H.; Bae, M. A.; Yazawa, K.; Sasaki, T.; Kobayashi, J. Alteramide A, a New Tetracyclic Alkaloid from a Bacterium *Alteromonas* sp. Associated with the Marine Sponge *Halichondria okadai*. *The Journal of Organic Chemistry* **1992,** *57* (15), 4317-4320, DOI: 10.1021/jo00041a053.

41. Moree, W. J.; McConnell, O. J.; Nguyen, D. D.; Sanchez, L. M.; Yang, Y.-L.; Zhao, X.; Liu, W.-T.; Boudreau, P. D.; Srinivasan, J.; Atencio, L.; Ballesteros, J.; Gavilán, R. G.; Torres-Mendoza, D.; Guzmán, H. M.; Gerwick, W. H.; Gutiérrez, M.; Dorrestein, P. C. Microbiota of Healthy Corals are Active Against Fungi in a Light-Dependent Manner. *ACS Chemical Biology* **2014,** *9* (10), 2300-2308, DOI: 10.1021/cb500432j.

42. Papenfort, K.; Bassler, B. L. Quorum Sensing Signal–Response Systems in Gram-Negative Bacteria. *Nature Reviews Microbiology* **2016,** *14* (9), 576-588, DOI: 10.1038/nrmicro.2016.89.

43. Doberva, M.; Stien, D.; Sorres, J.; Hue, N.; Sanchez-Ferandin, S.; Eparvier, V.; Ferandin, Y.; Lebaron, P.; Lami, R. Large Diversity and Original Structures of Acyl-Homoserine Lactones in Strain MOLA 401, a Marine *Rhodobacteraceae* Bacterium. *Frontiers in Microbiology* **2017,** *8* (1152), 8:1152, DOI: 10.3389/fmicb.2017.01152.

44. Arashida, N.; Shimbo, K.; Terada, T.; Okimi, T.; Kikuchi, Y.; Hashiro, S.; Umekage, S.; Yasueda, H. Identification of Novel Long Chain *N*-acylhomoserine Lactones of Chain Length C(20) from the Marine Phototrophic Bacterium *Rhodovulum sulfidophilum*. *Bioscience, Biotechnology, and Biochemistry* **2018,** *82* (10), 1683-1693, DOI: 10.1080/09168451.2018.1490168.

45. Vynne, N. G.; Månsson, M.; Nielsen, K. F.; Gram, L. Bioactivity, Chemical Profiling, and 16S rRNA-Based Phylogeny of *Pseudoalteromonas* Strains Collected on a Global Research Cruise. *Marine Biotechnology* **2011,** *13* (6), 1062-1073, DOI: 10.1007/s10126-011-9369-4.

46. Williams, R. P. Biosynthesis of Prodigiosin, a Secondary Metabolite of *Serratia marcescens*. *Applied Microbiology* **1973,** *25* (0003-6919 (Print)), 396-402, DOI: 10.1128/am.25.3.396-402.1973.

47. Fehér, D.; Barlow, R.; Lorenzo, P. S.; Hemscheidt, T. K. A 2-substituted Prodiginine, 2-(p-hydroxybenzyl)prodigiosin, from *Pseudoalteromonas rubra*. *Journal of Natural Products* **2008,** *71* (11), 1970-1972, DOI: 10.1021/np800493p.

48. Petersen, L.-E.; Kellermann, M. Y.; Nietzer, S.; Schupp, P. J. Photosensitivity of the Bacterial Pigment Cycloprodigiosin Enables Settlement in Coral Larvae—Light as an Understudied Environmental Factor. *Frontiers in Marine Science* **2021**, 8:749070, DOI: 10.3389/fmars.2021.749070.

49. Setiyono, E.; Adhiwibawa, M. A. S.; Indrawati, R.; Prihastyanti, M. N. U.; Shioi, Y.; Brotosudarmo, T. H. P. An Indonesian Marine Bacterium, *Pseudoalteromonas rubra*, Produces Antimicrobial Prodiginine Pigments. *ACS Omega* **2020,** *5* (9), 4626-4635, DOI: 10.1021/acsomega.9b04322.

50. Neely, K. L.; Shea, C. P.; Macaulay, K. A.; Hower, E. K.; Dobler, M. A. Short- and Long-Term Effectiveness of Coral Disease Treatments. *Frontiers in Marine Science* **2021**, 8:675349, DOI: 10.3389/fmars.2021.675349.

51. McAvoy, A. C.; Jaiyesimi, O.; Threatt, P. H.; Seladi, T.; Goldberg, J. B.; da Silva, R. R.; Garg, N. Differences in Cystic Fibrosis-Associated *Burkholderia* spp. Bacteria Metabolomes After Exposure to the Antibiotic Trimethoprim. *ACS Infectious Diseases* **2020,** *6* (5), 1154-1168, DOI: 10.1021/acsinfecdis.9b00513.

52. Ohlemacher, S. I.; Giblin, D. E.; d'Avignon, D. A.; Stapleton, A. E.; Trautner, B. W.; Henderson, J. P. Enterobacteria Secrete an Inhibitor of *Pseudomonas* Virulence During Clinical Bacteriuria. *Journal of clinical investigation* **2017,** *127* (11), 4018-4030, DOI: 10.1172/JCI92464.

53. Xu, G.; Guo, H.; Lv, H. Metabolomics Assay Identified a Novel Virulence-Associated Siderophore Encoded by the High-Pathogenicity Island in Uropathogenic *Escherichia coli*. *Journal of Proteome Research* **2019,** *18* (5), 2331-2336, DOI: 10.1021/acs.jproteome.9b00190.

54. Jaiyesimi, O. A.; McAvoy, A. C.; Fogg, D. N.; Garg, N. Metabolomic Profiling of *Burkholderia cenocepacia* in Synthetic Cystic Fibrosis Sputum Medium Reveals Nutrient Environment-Specific Production of Virulence Factors. *Scientific Reports* **2021,** *11* (1), 21419, DOI: 10.1038/s41598-021-00421-4.

55. Mislin, G. L. A.; Hoegy, F.; Cobessi, D.; Poole, K.; Rognan, D.; Schalk, I. J. Binding Properties of Pyochelin and Structurally Related Molecules to FptA of *Pseudomonas aeruginosa*. *Journal of Molecular Biology* **2006,** *357* (5), 1437-1448, DOI: 10.1016/j.jmb.2006.01.080.

56. Kaplan, A. R.; Musaev, D. G.; Wuest, W. M. Pyochelin Biosynthetic Metabolites Bind Iron and Promote Growth in *Pseudomonads* Demonstrating Siderophore-Like Activity. *ACS Infectious Diseases* **2021,** *7* (3), 544-551, DOI: 10.1021/acsinfecdis.0c00897.

57. Souto, A.; Montaos, M. A.; Rivas, A. J.; Balado, M.; Osorio, C. R.; Rodríguez, J.; Lemos, M. L.; Jiménez, C. Structure and Biosynthetic Assembly of Piscibactin, a Siderophore from *Photobacterium damselae* subsp. *piscicida*, Predicted from Genome Analysis. *European Journal of Organic Chemistry* **2012,** *2012* (29), 5693-5700, DOI: 10.1002/ejoc.201200818.

58. Galvis, F.; Ageitos, L.; Rodríguez, J.; Jiménez, C.; Barja, J. L.; Lemos, M. L.; Balado, M. *Vibrio neptunius* Produces Piscibactin and Amphibactin and Both Siderophores Contribute Significantly to Virulence for Clams. *Frontiers in Cellular and Infection Microbiology* **2021**, 11:750567, DOI: 10.3389/fcimb.2021.750567.

59. Osorio, C. R.; Rivas, A. J.; Balado, M.; Fuentes-Monteverde, J. C.; Rodríguez, J.; Jiménez, C.; Lemos, M. L.; Waldor, M. K. A Transmissible Plasmid-Borne Pathogenicity Island Confers Piscibactin Biosynthesis in the Fish Pathogen *Photobacterium damselae* subsp. *piscicida*. *Applied and Environmental Microbiology* **2015,** *81* (17), 5867-5879, DOI: 10.1128/AEM.01580-15.

60. Martinez, J. S.; Carter-Franklin, J. N.; Mann, E. L.; Martin, J. D.; Haygood, M. G.; Butler, A. Structure and Membrane Affinity of a Suite of Amphiphilic Siderophores Produced by a Marine Bacterium. *Proceedings of the National Academy of Sciences* **2003,** *100* (7), 3754, DOI: 10.1073/pnas.0637444100.

61. Galvis, F.; Ageitos, L.; Martínez-Matamoros, D.; Barja, J. L.; Rodríguez, J.; Lemos, M. L.; Jiménez, C.; Balado, M. The Marine Bivalve Molluscs Pathogen *Vibrio neptunius* Produces the Siderophore Amphibactin, Which is Widespread in Molluscs Microbiota. *Environmental Microbiology* **2020,** *22* (12), 5467-5482, DOI: 10.1111/1462-2920.15312.

62. Peixoto, R. S.; Rosado, P. M.; Leite, D. C. d. A.; Rosado, A. S.; Bourne, D. G. Beneficial Microorganisms for Corals (BMC): Proposed Mechanisms for Coral Health and Resilience. *Frontiers in Microbiology* **2017**, 8:341, DOI: 10.3389/fmicb.2017.00341.

63. Santoro, E. P.; Borges, R. M.; Espinoza, J. L.; Freire, M.; Messias, C. S. M. A.; Villela, H. D. M.; Pereira, L. M.; Vilela, C. L. S.; Rosado, J. G.; Cardoso, P. M.; Rosado, P. M.; Assis, J. M.; Duarte, G. A. S.; Perna, G.; Rosado, A. S.; Macrae, A.; Dupont, C. L.; Nelson, K. E.; Sweet, M. J.; Voolstra, C. R.; Peixoto, R. S. Coral Microbiome Manipulation Elicits Metabolic and Genetic Restructuring to Mitigate Heat Stress and Evade Mortality. *Science Advances* **2021,** *7* (33), eabg3088, DOI: 10.1126/sciadv.abg3088.

64. Staley, J. T.; Konopka, A. Measurement of In Situ Activities of Nonphotosynthetic Microorganisms in Aquatic and Terrestrial Habitats. *Annual Review of Microbiology* **1985,** *39* (1), 321-346, DOI: 10.1146/annurev.mi.39.100185.001541.

65. Deutsch, J. M.; Jaiyesimi, O. A.; Pitts, K. A.; Houk, J.; Ushijima, B.; Walker, B. K.; Paul, V. J.; Garg, N. Metabolomics of Healthy and Stony Coral Tissue Loss Disease Affected *Montastraea cavernosa* Corals. *Frontiers in Marine Science* **2021**, 8:714778, DOI: 10.3389/fmars.2021.714778.

66. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. Mzmine 2: Modular Framework for Processing, Visualizing, and Analyzing Mass Spectrometry-Based Molecular Profile Data. *BMC Bioinformatics* **2010,** *11* (1), 395, DOI: 10.1186/1471-2105-11-395.

67. Nothias, L.-F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; Aicheler, F.; Aksenov, A. A.; Alka, O.; Allard, P.-M.; Barsch, A.; Cachet, X.; Caraballo-Rodriguez, A. M.; Da Silva, R. R.; Dang, T.; Garg, N.; Gauglitz, J. M.; Gurevich, A.; Isaac, G.; Jarmusch, A. K.; Kameník, Z.; Kang, K. B.; Kessler, N.; Koester, I.; Korf, A.; Le Gouellec, A.; Ludwig, M.; Martin H, C.; McCall, L.-I.; McSayles, J.; Meyer, S. W.; Mohimani, H.; Morsy, M.; Moyne, O.; Neumann, S.; Neuweger, H.; Nguyen, N. H.; Nothias-Esposito, M.; Paolini, J.; Phelan, V. V.; Pluskal, T.; Quinn, R. A.; Rogers, S.; Shrestha, B.; Tripathi, A.; van der Hooft, J. J. J.; Vargas, F.; Weldon, K. C.; Witting, M.; Yang, H.; Zhang, Z.; Zubeil, F.; Kohlbacher, O.; Böcker, S.; Alexandrov, T.; Bandeira, N.; Wang, M.; Dorrestein, P. C. Feature-Based Molecular Networking in the GNPS Analysis Environment. *Nature Methods* **2020,** *17* (9), 905-908, DOI: 10.1038/s41592-020-0933-6.

68. Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V. SIRIUS 4: A Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information. *Nature Methods* **2019,** *16*, 299-302, DOI: 10.1038/s41592-019-0344-8.

69. Shannon, P.; Markie, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research* **2003,** *13*, 2498-2504, DOI: 10.1101/gr.1239303.

70. Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W. M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R. Proposed Minimum Reporting Standards for Chemical Analysis *Metabolomics* **2007,** *3* (3), 211-221, DOI: 10.1007/s11306-007-0082-2.

71. van der Hooft, J. J. J.; Wandy, J.; Young, F.; Padmanabhan, S.; Gerasimidis, K.; Burgess, K. E. V.; Barrett, M. P.; Rogers, S. Unsupervised Discovery and Comparison of Structural Families Across Multiple Samples in Untargeted Metabolomics. *Analytical Chemistry* **2017,** *89* (14), 7569-7577, DOI: 10.1021/acs.analchem.7b01391.

72. *PLS_Toolbox 8.8.1*, (2021). Eigenvector Research, Inc.: Manson, WA USA 98831, software available at http://www.eigenvector.com.

73. Khan, A.; Mathelier, A. Intervene: A Tool for Intersection and Visualization of Multiple Gene or Genomic Region Sets. *BMC Bioinformatics* **2017,** *18* (1), 287, DOI: 10.1186/s12859-017-1708-7.

74. van Santen, J. A.; Jacob, G.; Singh, A. L.; Aniebok, V.; Balunas, M. J.; Bunsko, D.; Neto, F. C.; Castaño-Espriu, L.; Chang, C.; Clark, T. N.; Cleary Little, J. L.; Delgadillo, D. A.; Dorrestein, P. C.; Duncan, K. R.; Egan, J. M.; Galey, M. M.; Haeckl, F. P. J.; Hua, A.; Hughes, A. H.; Iskakova, D.; Khadilkar, A.; Lee, J.-H.; Lee, S.; LeGrow, N.; Liu, D. Y.; Macho, J. M.; McCaughey, C. S.; Medema, M. H.; Neupane, R. P.; O'Donnell, T. J.; Paula, J. S.; Sanchez, L. M.; Shaikh, A. F.; Soldatou, S.; Terlouw, B. R.; Tran, T. A.; Valentine, M.; van der Hooft, J. J. J.; Vo, D. A.; Wang, M.; Wilson, D.; Zink, K. E.; Linington, R. G. The Natural Products Atlas: An Open Access Knowledge Base for Microbial Natural Products Discovery. *ACS Central Science* **2019,** *5* (11), 1824-1833, DOI: 10.1021/acscentsci.9b00806.

75. Blunt, J.; Munro, M. MarinLit: A Database of the Marine Natural Products Literature. http://pubs.rsc.org/marinlit/

76. Cao, L.; Guler, M.; Tagirdzhanov, A.; Lee, Y.-Y.; Gurevich, A.; Mohimani, H. MolDiscovery: Learning Mass Spectrometry Fragmentation of Small Molecules. *Nature Communications* **2021,** *12* (1), 3718, DOI: 10.1038/s41467-021-23986-0.

77. Dührkop, K.; Shen, H.; Meusel, M.; Rousu, J.; S., B. Searching Molecular Structure Databases with Tandem Mass Spectra Using CSI:FingerID. *Proceedings of the National Academy of Sciences* **2015,** *112* (41), 12580-12585, DOI: 10.1073/pnas.1509788112.

78. Djoumbou, F. Y.; Eisner, R.; Knox, C.; Chepelev, L.; Hastings, J.; Owen, G.; Fahy, E.; Steinbeck, C.; Subramanian, S.; Bolton, E.; Greiner, R.; Wishart, D. S. ClassyFire: Automated Chemical Classification with a Comprehensive, Computable Taxonomy. *Journal of Cheminformatics* **2016,** *8* (1), 61, DOI: 10.1186/s13321-016-0174-y.

79. Dührkop, K.; Nothias, L.-F.; Fleischauer, M.; Reher, R.; Ludwig, M.; Hoffmann, M. A.; Petras, D.; Gerwick, W. H.; Rousu, J.; Dorrestein, P. C.; Böcker, S. Systematic Classification of Unknown Metabolites Using High-Resolution Fragmentation Mass Spectra. *Nature Biotechnology* **2020,** *39*, 462-471, DOI: 10.1038/s41587-020-0740-8.

80. Minoche, A. E.; Dohm, J. C.; Himmelbauer, H. Evaluation of Genomic High-Throughput Sequencing Data Generated on Illumina HiSeq and Genome Analyzer Systems. *Genome Biology* **2011,** *12* (11), R112, DOI: 10.1186/gb-2011-12-11-r112.

81. Eren, A. M.; Vineis, J. H.; Morrison, H. G.; Sogin, M. L. A Filtering Method to Generate High Quality Short Reads Using Illumina Paired-End Technology. *PLoS One* **2013,** *8* (6), e66643, DOI: 10.1371/journal.pone.0066643.

82. Martin, M. Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. *EMBnet.journal* **2011,** *17* (1), 3, DOI: 10.14806/ej.17.1.200.

83. Nurk, S.; Meleshko, D.; Korobeynikov, A.; Pevzner, P. A. metaSPAdes: A New Versatile Metagenomic Assembler. *Genome Research* **2017,** *27* (5), 824-834, DOI: 10.1101/gr.213959.116.

84. Eren, A. M.; Kiefl, E.; Shaiber, A.; Veseli, I.; Miller, S. E.; Schechter, M. S.; Fink, I.; Pan, J. N.; Yousef, M.; Fogarty, E. C.; Trigodet, F.; Watson, A. R.; Esen, Ö. C.; Moore, R. M.; Clayssen, Q.; Lee, M. D.; Kivenson, V.; Graham, E. D.; Merrill, B. D.; Karkman, A.; Blankenberg, D.; Eppley, J. M.; Sjödin, A.; Scott, J. J.; Vázquez-Campos, X.; McKay, L. J.; McDaniel, E. A.; Stevens, S. L. R.; Anderson, R. E.; Fuessel, J.; Fernandez-Guerra, A.; Maignien, L.; Delmont, T. O.; Willis, A. D. Community-Led, Integrated, Reproducible Multi-Omics with Anvi'o. *Nature Microbiology* **2021,** *6* (1), 3-6, DOI: 10.1038/s41564-020-00834-3.

85. Tamura, K.; Nei, M. Estimation of the Number of Nucleotide Substitutions in the Control Region of Mitochondrial DNA in Humans and Chimpanzees. *Molecular Biology and Evolution* **1993,** *10* (3), 512-526, DOI: 10.1093/oxfordjournals.molbev.a040023.

86. Felsenstein, J. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution* **1985,** *39* (4), 783-791, DOI: 10.1111/j.1558-5646.1985.tb00420.x.

87. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* **2018,** *35* (6), 1547-1549, DOI: 10.1093/molbev/msy096.

Graphical Abstract