Microbial Biotechnology

WILEY

MICROBIAL BIOTECHNOLOGY Open Access

RESEARCH ARTICLE OPEN ACCESS

Novel Gene Clusters for Secondary Metabolite Synthesis in Mesophotic Sponge-Associated Bacteria

Nuo Chen[1,2] | Liwei Liu[1] | Jingxuan Wang[1,2] | Deqiang Mao[1,2] | Hongmei Lu[1,2] | Tânia Keiko Shishido[3] | Shuai Zhi[4] | Hua Chen[5] | Shan He[1,6]

[1]Li Dak Sum Yip Yio Chin Kenneth Li Marine Biopharmaceutical Research Center, Health Science Center, Ningbo University, Ningbo, Zhejiang, China | [2]College of Food Science and Engineering, Ningbo University, Ningbo, Zhejiang, China | [3]Institute of Biotechnology, University of Helsinki, Helsinki, Finland | [4]School of Public Health, Ningbo University, Ningbo, Zhejiang, China | [5]Mingke Biotechnology Co., Ltd., Hangzhou, China | [6]Ningbo Institute of Marine Medicine, Peking University, Ningbo, Zhejiang, China

Correspondence: Liwei Liu (liuliwei@nbu.edu.cn) | Shan He (heshan@nbu.edu.cn)

Received: 24 April 2024 | **Revised:** 8 January 2025 | **Accepted:** 30 January 2025

Funding: This work was supported by grants from the National Natural Science Foundation of China (31600016 and 41776168), Startup Foundation of Ningbo University (422010882, 422110473, and 422207513), the Li Dak Sum Yip Yio Chin Kenneth Li Marine Biopharmaceutical Development Fund, the National 111 Project of China (D16013) and Ningbo Natural Science Foundation (2021Z04). T.K.S. is funded by Novo Nordisk Foundation (NNF22OC0080109).

Keywords: biosynthetic gene cluster | mesophotic coral ecosystem | metagenome-assembled genome | secondary metabolite | sponge

ABSTRACT

Mesophotic coral ecosystems (MCEs) host a diverse array of sponge species, which represent a promising source of bioactive compounds. Increasing evidence suggests that sponge-associated bacteria may be the primary producers of these compounds. However, cultivating these bacteria under laboratory conditions remains a significant challenge. To investigate the rich resource of bioactive compounds synthesised by mesophotic sponge-associated bacteria, we retrieved 429 metagenome-assembled genomes (MAGs) from 15 mesophotic sponges, revealing a strong correlation between bacterial diversity and sponge species. Furthermore, we identified 1637 secondary metabolite biosynthetic gene clusters (BGCs) within these MAGs. Among the identified BGCs, terpenes were the most abundant (495), followed by 369 polyketide synthases (PKSs), 293 ribosomally synthesised and post-translationally modified peptides (RiPPs) and 135 nonribosomal peptide synthetases (NRPSs). The BGCs were classified into 1086 gene cluster families (GCFs) based on sequence similarity. Notably, only five GCFs included experimentally validated reference BGCs from the Minimum Information about a Biosynthetic Gene cluster database (MIBiG). Additionally, an unusual abundance of BGCs was detected in *Entotheonella* sp. (s191209.Bin93) from the Tectomicrobia phylum. In contrast, members of Proteobacteria and Acidobacteriota harboured fewer BGCs (6–7 on average), yet their high abundance in MCE sponges suggests a potentially rich reservoir of BGCs. Analysis of the BGC distribution patterns revealed that a subset of BGCs, including terpene GCFs (FAM_00447 and FAM_01046), PKS GCF (FAM_00235), and RiPPs GCF (FAM_01143), were widespread across mesophotic sponges. Furthermore, 32 GCFs were consistently present in the same MAGs across different sponges, highlighting their potential key biological roles and capacity to yield novel bioactive compounds. This study not only underscores the untapped potential of mesophotic sponge-associated bacteria as a source of bioactive compounds but also provides valuable insights into the intricate interactions between sponges and their symbiotic microbial communities.

Nuo Chen and Liwei Liu contributed equally to this work.

1 | Introduction

Sponges (phylum Porifera) are renowned as a prolific resource of bioactive compounds, with approximately 200 new compounds, including nonribosomal peptide synthetase (NRPS), polyketide synthase (PKS) and terpenes, being reported each year (Laport et al. 2009; Varijakzhan et al. 2021). However, increasing evidence suggests that many of these bioactive compounds are not directly produced by the sponges themselves but by their symbiotic bacteria. Examples include mycalamide, pateamine, polytheonamide and peloruside, which have been linked to bacterial origin (Freeman et al. 2016; Rust et al. 2020; Storey et al. 2020). This realisation has intensified the focus on sponge-associated bacteria as potential sources for novel bioactive compounds. Sponges host a remarkably diverse microbial community, which can constitute up to 35% of their total biomass (Hentschel et al. 2012). Despite significant efforts to culture these bacteria in laboratory conditions, only a small fraction of these microbes have proven cultivable, thereby limiting their exploration (Dat et al. 2021).

Genome mining has emerged as a powerful strategy for the discovery of microbial secondary metabolites. By analysing genomic data, researchers can identify and predict biosynthetic gene clusters (BGCs) responsible for the production of novel compounds (Kautsar et al. 2018; Kalkreuter et al. 2020; Blin et al. 2021). Advances in sequencing technology and the associated reduction in costs have driven a rapid increase in available genome data, further enhancing the utility of genome mining for uncovering new bioactive compounds (Wei et al. 2021; Shaffer et al. 2022). As a culture-independent approach, metagenomics has proven particularly effective in exploring the genetic potential of environmental microbial communities. By combining metagenomic assembly, genome binning and BGC prediction, this method bypasses the cultivation barrier, enabling the identification of secondary metabolite resources directly from environmental microbiomes (Hugenholtz et al. 1998; Marchesi 2012; Ayling et al. 2020). This integrated approach has already demonstrated success in uncovering cryptic BGCs from Antarctic soil, marine waters and oral microbiomes (Aleti et al. 2019; Paoli et al. 2022; Waschulin et al. 2022).

Mesophotic coral ecosystems (MCEs) represent a distinct component of marine biodiversity. These ecosystems, defined as coral reef "twilight zones," are dominated by light-dependent corals, sponges, and algae, residing at depths of 30 to 150 m (Hinderstein et al. 2010). Accessing these ecosystems requires professional training, and their microbial diversity and biosynthetic potential for secondary metabolites remain largely unexplored due to the cultivation challenges associated with their bacterial communities (Olson and Kellogg 2010; Hentschel et al. 2012). In this study, we collected 15 sponge samples from MCEs using scuba diving and employed a culture-independent metagenomic strategy to investigate the biosynthetic potential of secondary metabolites from their associated bacteria. A total of 429 metagenome-assembled genomes (MAGs) were retrieved, from which 1637 BGCs were identified. Analysis of BGC distribution revealed an uneven pattern across MAGs and sponge species. Altogether, our study provides new insights into the untapped reservoir of secondary metabolites in mesophotic sponge-associated bacteria.

2 | Materials and Methods

2.1 | Sponge Collection and Identification

A total of 15 mesophotic sponge samples were collected from the Philippines using scuba diving (Table S1 and Figure S1). Samples were immediately stored at −80°C. Sponge-metagenomic DNA was extracted following a previously described protocol without modifications (Gurgui and Piel 2010). This DNA served as a template for polymerase chain reaction (PCR) amplification of the sponge 28S rRNA gene and the sponge mitochondrial 16S rRNA gene (Erpenbeck et al. 2005; Watkins and Beckenbach 1999). The PCR conditions were as follows: initial denaturation at 94°C for 5 min, followed by 30 cycles of 30 s at 94°C, 30 s at 48°C and 1 min at 72°C, with a final extension at 72°C for 5 min. Each 25 μL reaction contained 12.5 μL of 2× Wazyme Taq polymerase (Nanjing Vazyme Biotech), 9 μL of DNase-free water, 1 μL of each primer (10 μM) and 2 μL of DNA (10–20 ng). Specific primers for the sponge 28S rRNA gene were 28sCallyF (5′-TGCGACCCGAAAGATGGTGAACTA-3′) and 28sCallyR (5′-ACCAACACCTTTCCTGGTATCTGC-3′) (Gurgui and Piel 2010; López-Legentil et al. 2010). Primers for the mitochondrial 16S rRNA gene were 16S-F (5′-TCGACTGTTTACCAAAAACATAGC-3′) and 16S-R (5′-YRTAATTCAACATCGAGGTC-3′) (Watkins and Beckenbach 1999). PCR products were sequenced through an ABI 3730 platform (Genewiz, Suzhou, China), and the sequences were BLASTed against the rRNA/ITS database at the National Center for Biotechnology Information (NCBI) for taxonomic identification. Phylogenetic analysis of the 28S rRNA gene was performed to deduce the taxonomic classification of MCE sponges (Erpenbeck et al. 2005). The analysis of the sponge mitochondrial 16S rRNA gene was employed as a complementary approach to validate the taxonomic assignments.

2.2 | Sponges Metagenomic Sequencing

Crude sponge-metagenomic DNA was further treated with RNase (10 μg/mL, Sangon Biotech) at 37°C for 1 h, followed by treatment with Proteinase K (50 μL/mL, Beyotime Biotech) at 55°C for 2 h. DNA was subsequently purified with magnetic beads and subjected to high-throughput sequencing on the HiSeq PE150 platform (Illumina) with a sequencing depth of 30 Gbp. Sequencing services were provided by Genewiz (Suzhou, China).

2.3 | Quality Trimming Sequencing Data and Adapter Removal

Raw sequencing reads were processed using Trimmomatic (Bolger et al. 2014) to remove Illumina adapters and filter for quality. Parameters used were PE -threads 8 -phred33 ILLUMINACLIP:adapter.fa:2:30:10 LEADING:20 TRAILING:20 SLIDINGWINDOW:4:20 MINLEN:75.

2.4 | Metagenome Assembly

Filtered reads from all sponge samples were assembled using MEGAHIT (Li et al. 2016) with the --min-contig-len parameter

set to 500 bp. Low-coverage contigs were removed using Salmon (Patro et al. 2017; Xie et al. 2022). The final assembly data have been deposited in the NCBI database (Table S1).

2.5 | Sequence Estimated Coverage

Redundancy and coverage of reads were evaluated using Nonpareil v3.304 (Rodriguez et al. 2018) with the -T kmer and -X 1000000 parameters. Nonpareil curves were generated using the R package Nonpareil in RStudio (v4.0.3).

2.6 | MAG Binning and Classification

MAGs were binned using MetaBAT2 (Kang et al. 2019) with --abdFile, incorporating mean and variance information of base coverage depth. MAG completeness and contamination were evaluated using CheckM (Parks et al. 2015). Bins with ≥80% completeness and ≤10% contamination were retained. Clean reads were aligned to the MAGs using BWA-MEM (Li 2013). MAG-associated reads were reassembled with SPAdes (Nurk et al. 2013), retaining contigs >1000 bp. CheckM was used for re-evaluation, and average nucleotide identity (ANI ≥95%) was applied to remove redundant MAGs (Parks et al. 2015). Taxonomic classification of filtered MAGs was performed using the GTDB-Tk with default parameters (Chaumeil et al. 2022).

2.7 | AntiSMASH and BiG-SCAPE Analysis

BGCs were predicted using antiSMASH v6.0 (Blin et al. 2021) with the following parameters: --genefinding-tool prodigal --cb-knownclusters --tigrfam --cb-subclusters --asf --pfam2go --cc-mibig --rre. BiG-SCAPE v1.0.1 (Navarro-Muñoz et al. 2020) was used to cluster BGCs based on sequence similarity with the parameters --mix --mibig and a similarity cutoff of 0.5. Predicted BGCs were compared to reference pathways in the MIBiG database (Medema et al. 2015). Network visualisations were constructed in Cytoscape v3.10.1 (Shannon et al. 2003).

2.8 | Comparative Analysis of 429 Genomes

A highly resolved phylogenetic tree of 429 genomes was constructed based on a pre-defined core gene using the Universal Biased Clustered Genes (UBCG) pipeline (Na et al. 2018). The genome files were initially converted from the FASTA format to the BCG format. Then, the resulting BCG files containing UBCG gene sequences and metadata were aligned to construct the evolutionary tree, which was exported in the Newick (NWK) file format. The resulting phylogenetic tree was visualised and annotated using iTOL v5 (Letunic and Bork 2019; Na et al. 2018).

2.9 | Existence/Absence Matrix

The presence or absence of 41 gene cluster families (GCFs) across mesophotic sponge samples was visualised using iTOL v5. A script was developed to extract and integrate data on BGC occurrence. The output was visualised in iTOL v5 with MAG types labelled in different colours. The script is available at https://itol.embl.de/shared_projects.cgi.

3 | Results

3.1 | MAGs Retrieved From Mesophotic Sponges

To investigate the diversity of secondary metabolite resources in mesophotic sponge-associated bacteria, we collected 15 sponge samples from the mesophotic zones of Puerto Galera and Apo Reef via SCUBA diving (Table S1 and Figure S1). The sponge samples were preliminarily identified through phylogenetic analyses of 16S rRNA and 28S rRNA genes (Table S1). Subsequently, metagenomic sequencing was employed to explore the associated bacterial communities. Initial feasibility tests on samples '190707' and '191227' indicated that sequencing depths between 1 and 20 Gbp were required to recover 80% of the sequence diversity, as determined by nonpareil projection curves (Figure S2). Based on these findings, a sequencing depth of 30 Gbp per sample was applied to ensure comprehensive coverage.

Draft MAGs were constructed through de novo assembly and binning of contigs. To mitigate the impact of low genome quality—which often results in fragmented BGC predictions—MAGs with completeness >80% and contamination <10% were retained (Bowers et al. 2017). Of the 429 MAGs, 60.84% were classified as high-quality (completeness >90%, contamination <5%; Table S2). These MAGs exhibited a median GC content of 62.92% and a median contig N50 of 65,237 (Table S2). Concerning genome size, the MAGs range from 0.72 Mb to 8.00 Mb, including completeness values into the genome size estimations, with N50 values between 4986 and 3,337,178.

Using the GTDB-Tk pipeline (Chaumeil et al. 2022), the MAGs were classified into 27 bacterial phyla (Table S3). Sample-specific MAG counts varied from 3 to 59 (Figure 1, Table S1), with higher abundances observed in sponges *Melophlus*, *Rhabderemia*, *Terpios* and *Xestospongia* (average: 45 MAGs per sponge). Heatmap analysis (Figure 1) revealed that certain phyla, such as *Actinobacteriota*, *Desulfobacterota_D* and *Proteobacteria*, were widely distributed, while others showed strong host specificity. For instance, *Tectomicrobia* was enriched in *Theonella*, and *Nitrospinota* was prevalent in *Cinachyrella*. Notably, *Acidobacteriota* dominated in *Axinella*, *Rhabderemia* and *Terpios*, while *Chloroflexota* was abundant in *Melophlus*, *Terpios* and *Xestospongia*. This compositional specificity among different sponge genera offers valuable insights for the targeted exploration of bacteria resources within mesophotic sponge ecosystems. Although the bacteria composition of samples from the same sponge genus, such as *Axinella* and *Cinachyrella*, shows noticeable variation, we attribute this to the limited biomass of the samples and the low yield of high-quality DNA.

3.2 | Diversity and Novelty of BGCs in Assembled MAGs

The biosynthetic potential of sponge-associated bacteria was assessed using antiSMASH 6.0 (Blin et al. 2021), which

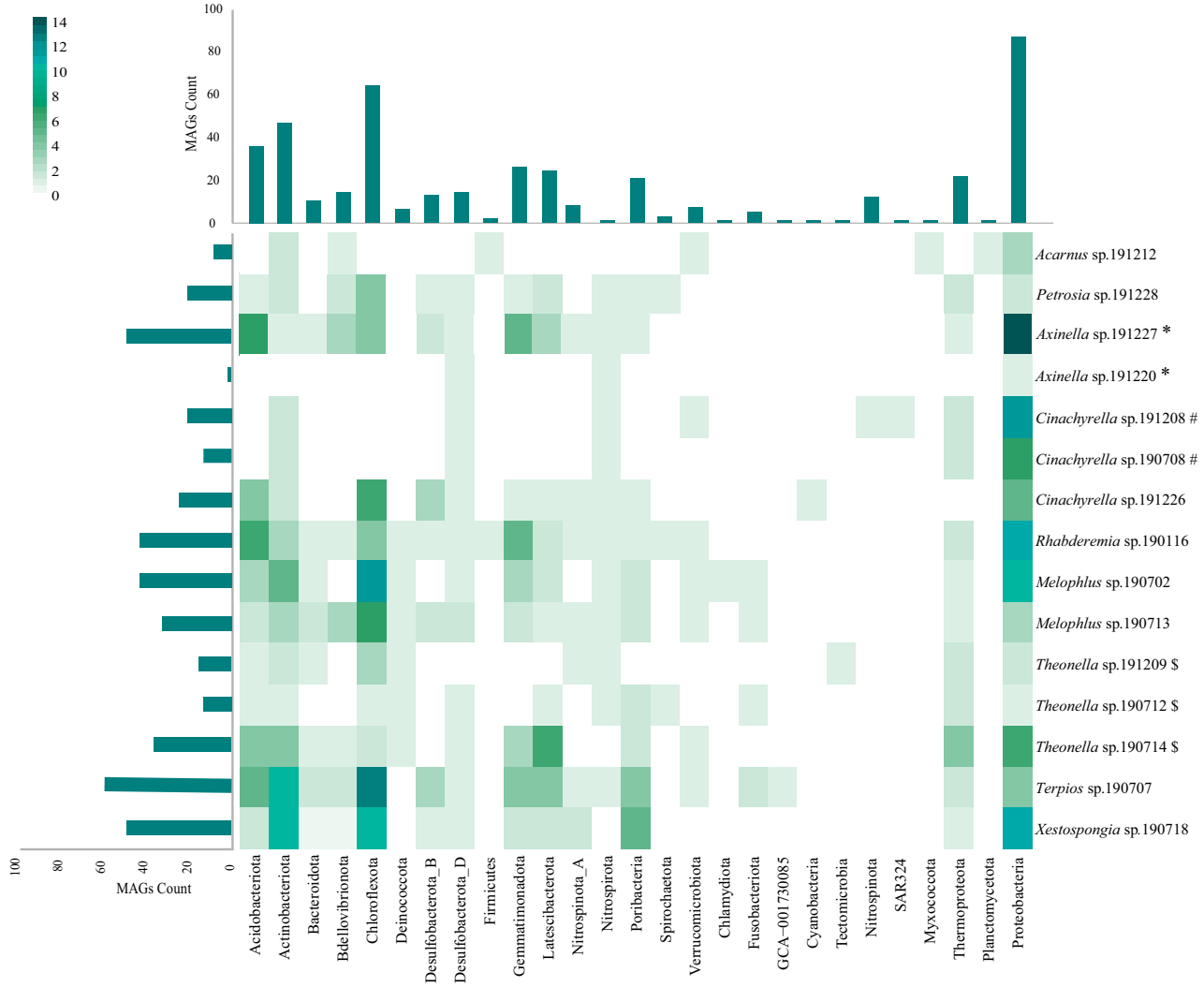

FIGURE 1 | The analysis of 429 metagenome-assembled genomes revealed a characteristic composition of bacteria in mesophotic sponges. Lineage statistics of metagenome assembly genomes (MAGs) from 15 mesophotic sponge samples (Left, bar graphs). The x-axis represents the taxonomic names of sponges, and the y-axis indicates the total number of MAGs in each sponge sample. Statistics of the number of MAGs classified at the phylum level (Above, bar graphs). The *x*-axis represents the taxonomic names of MAGs, and the *y*-axis indicates the total number of MAGs in each bacteria phylum. The MAGs (Below, heatmap) were classified at the phylum level using ANI analysis and assigned to their host sponges in the heat map. The *x*-axis indicates the taxonomic names of MAGs, and the *y*-axis indicates the sponge sample names. The depth of colour represents the number of MAGs, ranging from 0 to 14. The tool used for creating the heatmap is available at http://www.ehbio.com/ImageGP/index.php/Home/Index/PCAplot.html. The symbols *, # and $ indicate that the marked sponge samples are from the same species.

identified 1637 BGCs across the 429 MAGs, spanning 35 categories, including NRPS, PKS, terpene, RiPPs and others (Table S3). Among these, terpene BGCs were the most abundant (30.2%), followed by PKS (22.1%), RiPPs (17.9%) and NRPS (8.3%), cumulatively comprising 78.5% of all detected BGCs (Figure S3 and Table S3). Phylogenetic analysis using the UBCG pipeline (Na et al. 2018) revealed that terpene and PKS BGCs were widely distributed across MAGs, whereas NRPS and RiPP clusters exhibited greater species specificity (Figure 2). NRPS clusters were predominantly found in Poribacteria, Latescibacterota and Acidobacteriota, while RiPPs were enriched in Nitrospirota, Acidobacteriota, Proteobacteria, Poribacteria and Gemmatimonadetes. Additionally, BGC types such as ectoine, arylpolyene, resorcinol and hserlactone were categorised as 'others' due to their sporadic distribution (Figure 2 and Table S3).

To enhance accuracy in defining BGC boundaries, BiG-SCAPE was employed to cluster BGCs into GCFs using pairwise distance metrics (Navarro-Muñoz et al. 2020). This approach identified 270 GCFs and 816 singletons, with only five GCFs matching experimentally validated BGCs in the MIBiG database. These included one desferrioxamine, one aerobactin and three eicosapentaenoic acid BGCs (Figure 3 and Table S3).

PKS and NRPS BGCs are widely regarded as critical targets for the discovery of clinically relevant drugs (Katz et al. 2016). In this study, we identified 369 PKS and 135 NRPS gene clusters (Table S4). Among the various bacterial groups analysed, Acidobacteriota exhibited the highest average number of PKS and NRPS gene clusters per MAG, with an average of three per MAG (Table S4). Of the identified PKS gene clusters, 40 were

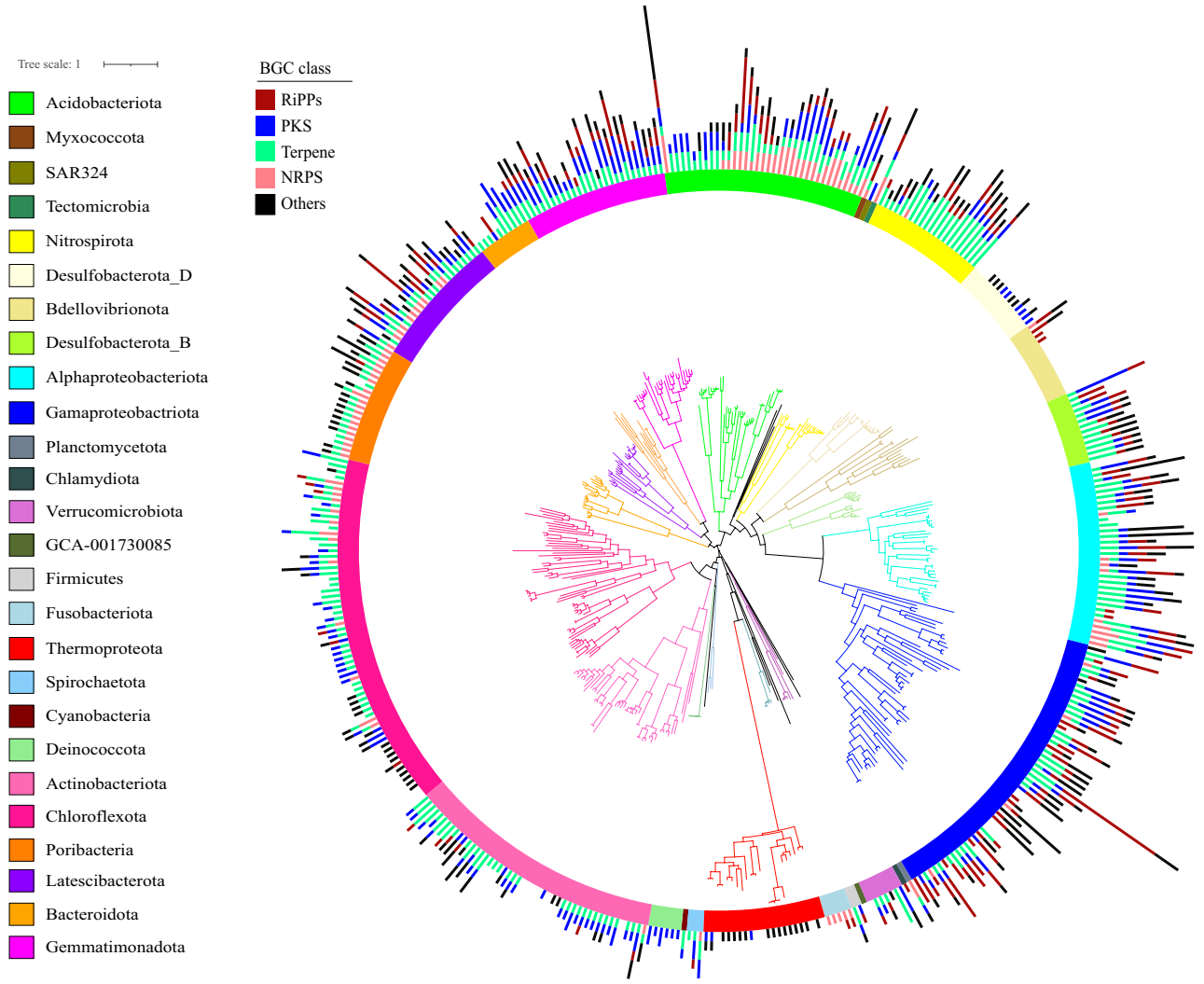

FIGURE 2 | Bioinformatics analysis of 429 metagenome-assembled genomes points toward a high biosynthetic genetic potential and distribution diversity. Layer 1: The phylum classification of MAGs is marked with different colours on the left. Layer 2: BGC content was classified into five BGC types in the top left corner, which are RiPPs, PKS, Terpene, NRPS and others (Table S3). All five groups were marked with different colours. The phylogenetic tree was constructed using UBCG with 92 genes.

categorised as type III PKS, while the remaining clusters were grouped as type I PKS or PKS-like (Table S3).

Five type I PKS GCFs displayed a conserved architecture characterised by an NRPS-like core gene that included domains for ketosynthase (KS), acyltransferase (AT), dehydrogenase (DH), c-methyltransferase (cMT), enoylreductase (ER), ketoreductase (KR) and acyl carrier protein (ACP), as well as a repeat of the KS-AT modules (Figure 4a). Flanking these core genes were additional genes, including abhydrolase, 4′-phosphopantetheinyl transferase, transporter and genes of unknown function. AntiSMASH analysis revealed similarities between these GCFs and the known BGC encoding strobilurin A (MiBIG accession no. BGC0001909.1). However, structural variations in the adjacent genes and the NRPS-like core domains strongly suggest the potential for cryptic bioactive compounds. Notably, GCFs FAM_00235 and FAM_00459 appeared to be vertically transferred, as they were exclusively derived from Actinobacteriota and Desulfobacterota_D, respectively. In contrast, the other three GCFs (FAM_00382, FAM_01276 and FAM_00977) exhibited horizontal gene transfer across diverse bacterial origins (Figure S4).

Additionally, two large trans-AT PKS BGCs were identified in a Verrucomicrobiota MAG (190,116 bin62, d_Bacteria; p_Verrucomicrobiota; c_Kiritimatiellae) (Figure 4b, Table S3). These two gene clusters contained 23 and 13 assembly modules, respectively, along with diverse modification domains, including KR, DH, ER, Ox, PS, MT, HMGS, ECH and NRPS modules. BGC I, approximately 106.7 kb in length and comprising 13 genes, encodes 23 PKS modules and various modification domains. Based on step-by-step domain analysis, we proposed that the biosynthesis of compound A begins with alkyl-CoA loading onto the first ACP domain, followed by 20 rounds of dicarbon unit extensions. The resulting linear intermediate is cyclised by the second TE domain at the end of the BGC I to form compound A. BGC II, approximately 58.9 kb in length and containing 13 genes, encodes 13 PKS modules and one NRPS module. This BGC is proposed to initiate with alkyl-CoA loading onto the first ACP domain, followed by 11 rounds of dicarbon unit extensions and one glycine incorporation by the NRPS module. The final linear-chain is cyclised into compound B by a thioesterase (TE) domain (Figure 4b).

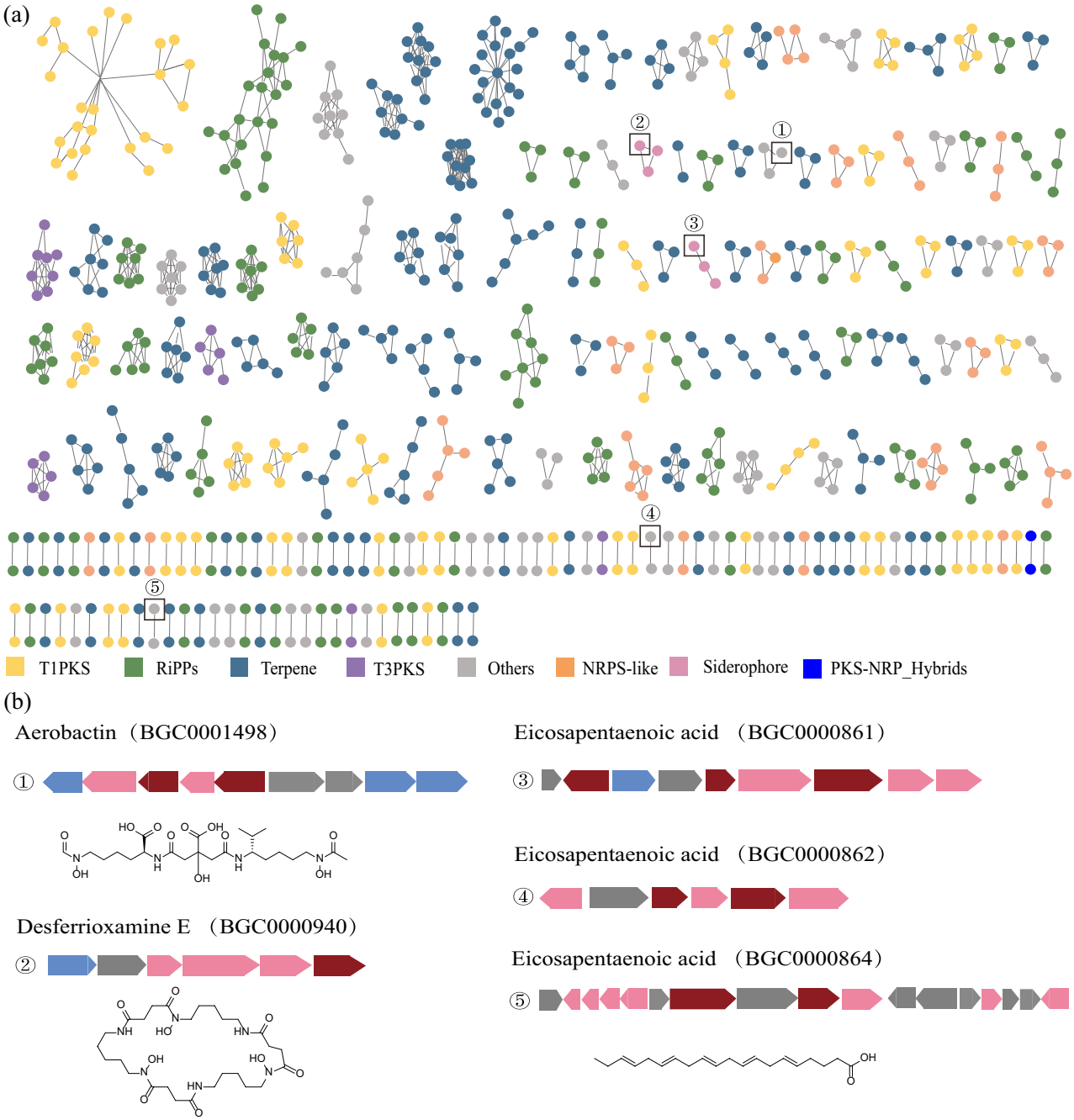

FIGURE 3 | BiG-SCAPE analysis of metagenome-assembled genomes of mesophotic sponges uncovers a giant unexplored genetic potential. (a) A global network of all retrieved gene cluster families was constructed using a cutoff of 0.5 in BiG-SCAPE. Different colours represent different GCF categories. The GCFs with known BGC were marked with a cube and a number. (b) Five GCFs are confirmed to own a known BGC, and five known gene clusters are depicted with their corresponding compounds.

Ribosomally synthesised and post-translationally modified peptides (RiPPs) represent an important class of natural products with significant antibiotic potential. Our study revealed extensive RiPP diversity in mesophotic sponge-associated bacteria, with BGCs clustered into 61 GCFs and 247 singletons (Figure 5a). These clusters encompass a wide range of RiPP classes, including proteusins, ranthipeptides, lanthipeptides, lassopeptides, linaridins, linear azole-containing peptides (LAPs), thiopeptides, thioamitides, cyanobactins and sactipeptides (Figure 5a). Proteusins were the most abundant RiPP type, with 16 GCFs comprising 107 BGCs. Notably, GCFs FAM_01149 and FAM_01143 clustered into a larger proteusin family, sharing conserved core genes (Figure 5a). Proteusin BGCs consistently contained two transporter genes, a YcaO cyclodehydratase, a precursor gene and several genes of unknown function (Figure 5b). Sequence alignment of proteusin precursor peptides revealed a conserved 'LCCC' motif across most identified BGCs, although the final structures exhibited notable diversity (Figure 5c, Figure S5). Furthermore, the presence of flanking genes encoding unknown compounds suggests untapped chemical diversity within these BGCs. Interestingly, RiPP clusters from other classes showed no significant similarity to known

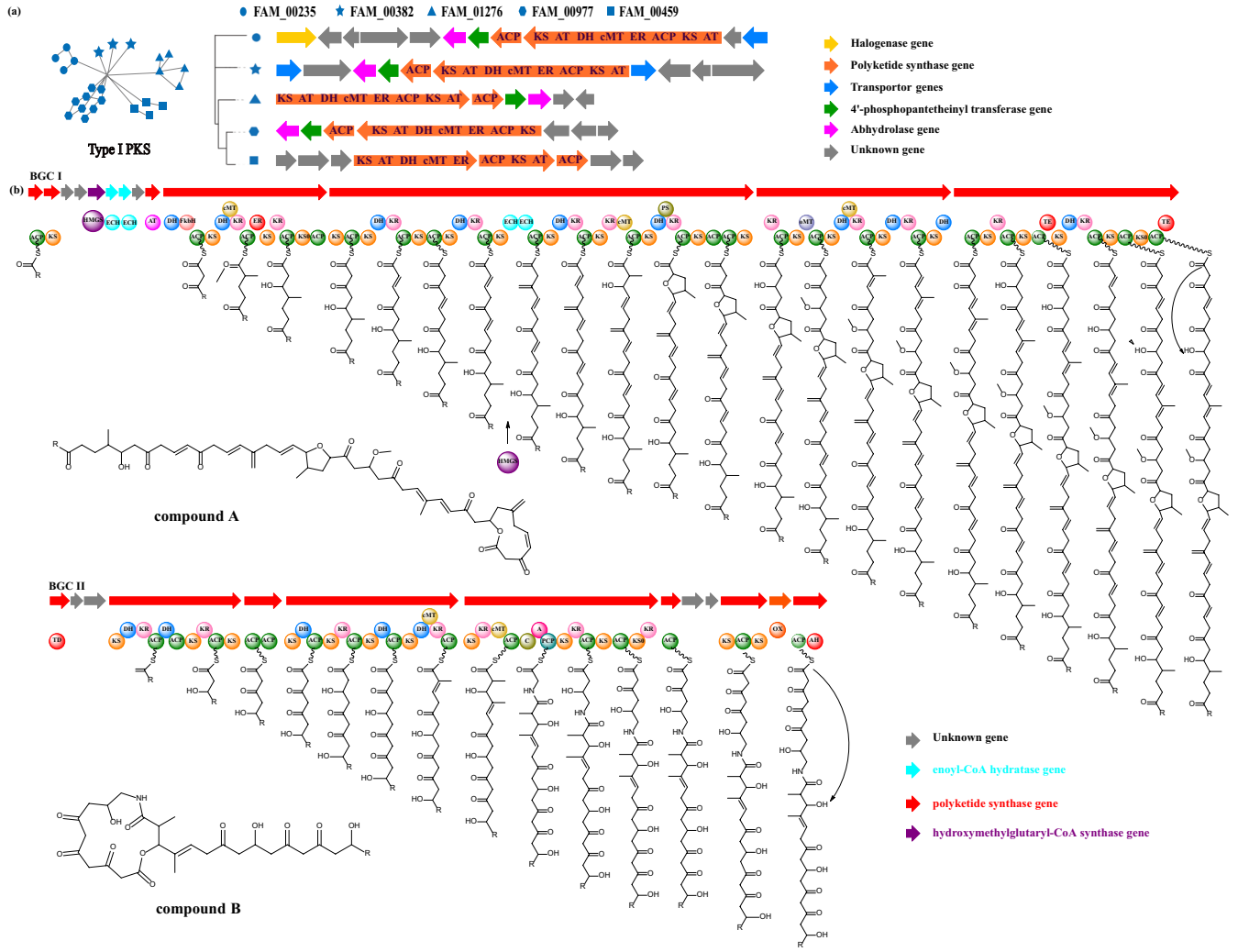

FIGURE 4 | The type I PKS BGCs from mesophotic sponge-associated bacteria exhibited great potential for cryptic polyketides. (a) The most abundant type I PKS gene cluster was identified to cross five GCFs and share similar core genes. Each GCF is represented by one representative gene cluster in the figure, and all gene clusters are shown in Figure S4. The domains in core genes and other biosynthetic enzymes were marked with different colours on the right. (b) Two trans-AT PKS gene clusters were found in MAG (190,116 bin62, d_Bacteria; p_Verrucomicrobiota; c_Kiritimatiellae), and their biosynthetic pathways are proposed. The enzymes that might be involved in the PKS biosynthesis are marked with different colours. Functional domains were indicated by bold letters: A, adenylation domain; ACP, acyl carrier protein; AH, abhydrolase; AT, malonyl CoA-acyl carrier protein transacylase; C, condensation domain; cMT, carboxyl methyltransferase; ECH, enoyl-CoA hydratase; ER, enoyl reductase; FkbH, FkbM family methyltransferase; HMGS, hydroxymethylglutaryl-CoA synthase; KR, ketoreductase; KS, ketosynthase; oMT, oxygenic methyltransferase; Ox: monooxygenase; PS, pyran synthase; PCP, thiolation structural domain (peptidyl carrier protein or acyl carrier protein); TD, thioester reductase domain-containing protein; TE, thioesterase domain.

BGCs, highlighting their novelty and potential for new bioactive compounds.

3.3 | Uncommon Abundant BGCs in Tectomicrobia

To investigate the abundance and diversity of BGCs across 27 bacterial phyla, we calculated the average number of BGCs per phylum (Figure 6a). While Proteobacteria and Chloroflexota were represented by the highest number of MAGs, their average BGC counts were relatively low, with 5 and 2 BGCs per genome, respectively. In contrast, the *Entotheonella* sp. genome (s191209.Bin93) from the Tectomicrobia phylum, which spans 6.4 Mb, encoded 11 BGCs with a cumulative length of 127 kb. These BGCs included one redox-cofactor, two RiPP-like clusters, two terpene clusters, one type III PKS, one ectoine, one RRE-containing cluster, one type I PKS and two NRPS BGCs. Remarkably, none of these BGCs showed similarity to known clusters in the MIBiG database, nor BGCs in previously characterised *Entotheonella* strains, such as *Entotheonella sp.* TSY1 and TSY2 (Table S5; Wilson et al. 2014). Among other phyla, Nitrospirota exhibited the highest average number of BGCs per genome, with approximately 8 BGCs per strain, followed by Acidobacteriota, Proteobacteria, Desulfobacterota_B and Gemmatimonadota, each with an average of 5–7 BGCs (Table S4). While Proteobacteria and Acidobacteriota harboured relatively few BGCs per genome, their abundance in the MAG dataset suggests that these phyla contribute significantly to the overall BGC pool in MCE sponge-associated bacteria.

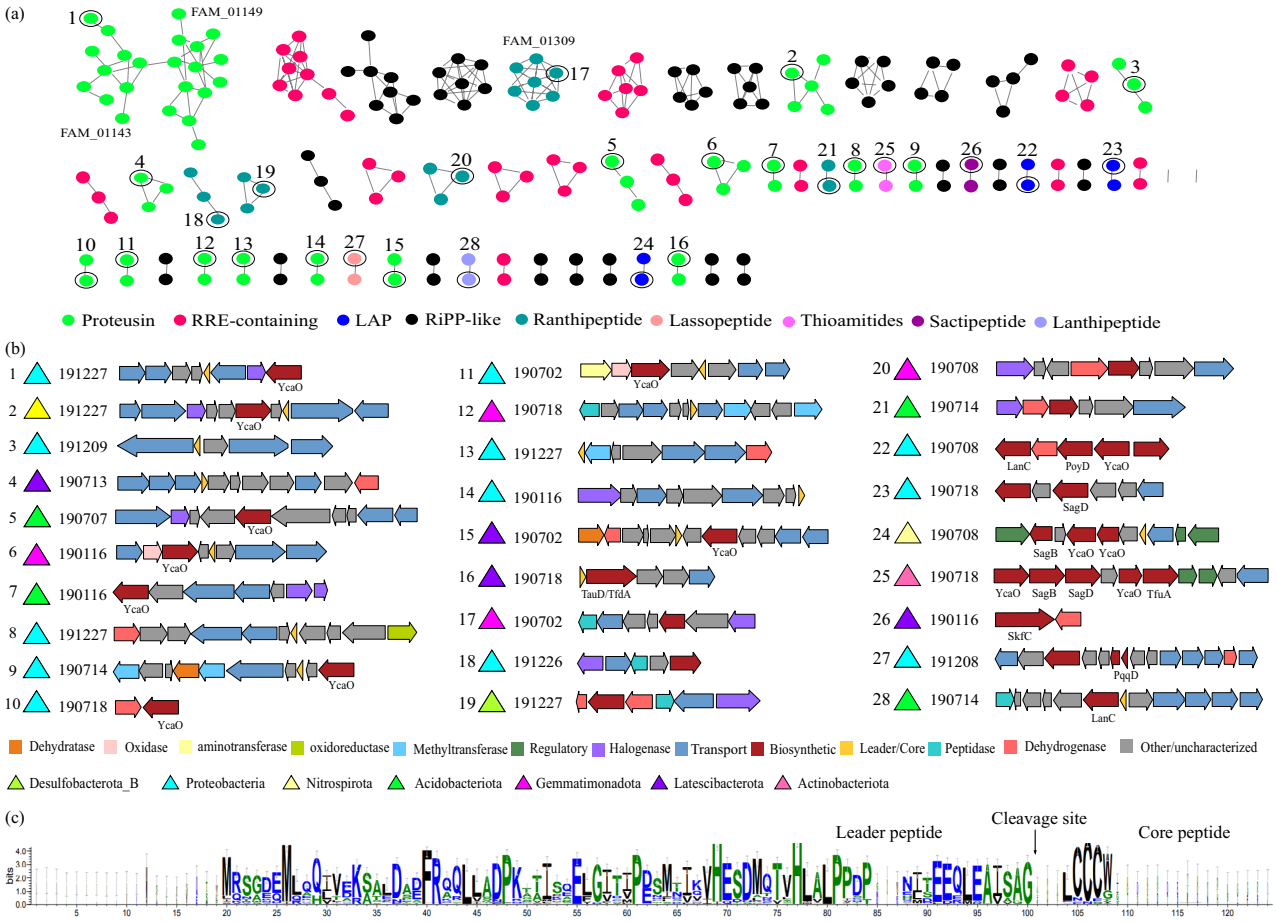

FIGURE 5 | The presence of RiPP gene clusters reflects the diversity of peptidic compounds biosynthesised by sponge-associated bacteria. (a) The similarity network of RiPPs was generated by BiG-SCAPE analysis. The different types of RiPPs BGCs and biosynthetic genes were marked with different colours. (b) The representative BGCs from each GCF were depicted, except REE and RiPP-like. The MAGs, biosynthetic genes and sponge hosts were marked with different colours and strain numbers. (c) Multiple sequence alignment of the proteusin precursor peptides. The cleavage sites were indicated by arrows. This figure was generated by the web-based software WebLogo (Crooks et al. 2004).

3.4 | A Small Subset of BGCs Is Widespread in Mesophotic Sponges

To evaluate the distribution of identified BGCs across mesophotic sponge samples, we grouped BGCs into GCFs, defined by shared core genes, to facilitate distribution analysis. For reliability, we focused on GCFs containing at least four BGCs, resulting in 41 GCFs (Table S6). We further analysed their presence across sponge samples and corresponding MAGs. Among these, 34 GCFs were identified in at least five sponge samples. Notably, four GCFs demonstrated particularly broad distributions, terpene GCFs FAM_00447 and FAM_01046, PKS GCF FAM_00235 and RiPPs GCF FAM_01143, which were detected in 9 sponges (5 species), 9 sponges (5 species), 7 sponges (5 species) and 11 sponges (7 species), respectively (Table S6 and Figure 6b). Of particular note, two terpene GCFs (FAM_00447 and FAM_01046) exhibited strong conservation but lacked similarity to any BGCs in the MIBiG database, highlighting their potential as novel natural product sources. In contrast, GCFs FAM_00235 and FAM_01143 demonstrated strong similarity to known biosynthetic clusters encoding strobilurin A (MIBiG accession no. BGC0001909.1) and proteusin (MIBiG accession no. BGC0000598), respectively. Interestingly, 32 GCFs were found consistently in the same MAGs across different sponges, suggesting their conservation within specific bacterial taxa. For instance, FAM_01046 and FAM_00447 were conserved in Nitrospirota and Latescibacterota, respectively. Conversely, 9 GCFs, such as FAM_01211, FAM_00235 and FAM_01143, were identified across different MAGs within single sponge samples, indicating potential horizontal gene transfer or lineage-specific diversification. This combination of conservation, novelty and functional similarity underscores the chemical and ecological diversity of BGCs in mesophotic sponge-associated bacteria.

4 | Discussion

Sponges-associated bacteria are a rich source of bioactive compounds (Freeman et al. 2016; Rust et al. 2020; Storey et al. 2020). In this study, we focused on mesophotic sponges as a novel target and employed a genome mining strategy to explore their biosynthetic potential for secondary metabolites, addressing the challenge of culturing these bacteria. This approach uncovered an exceptional secondary metabolite reservoir across various bacterial phyla and revealed the widespread distribution

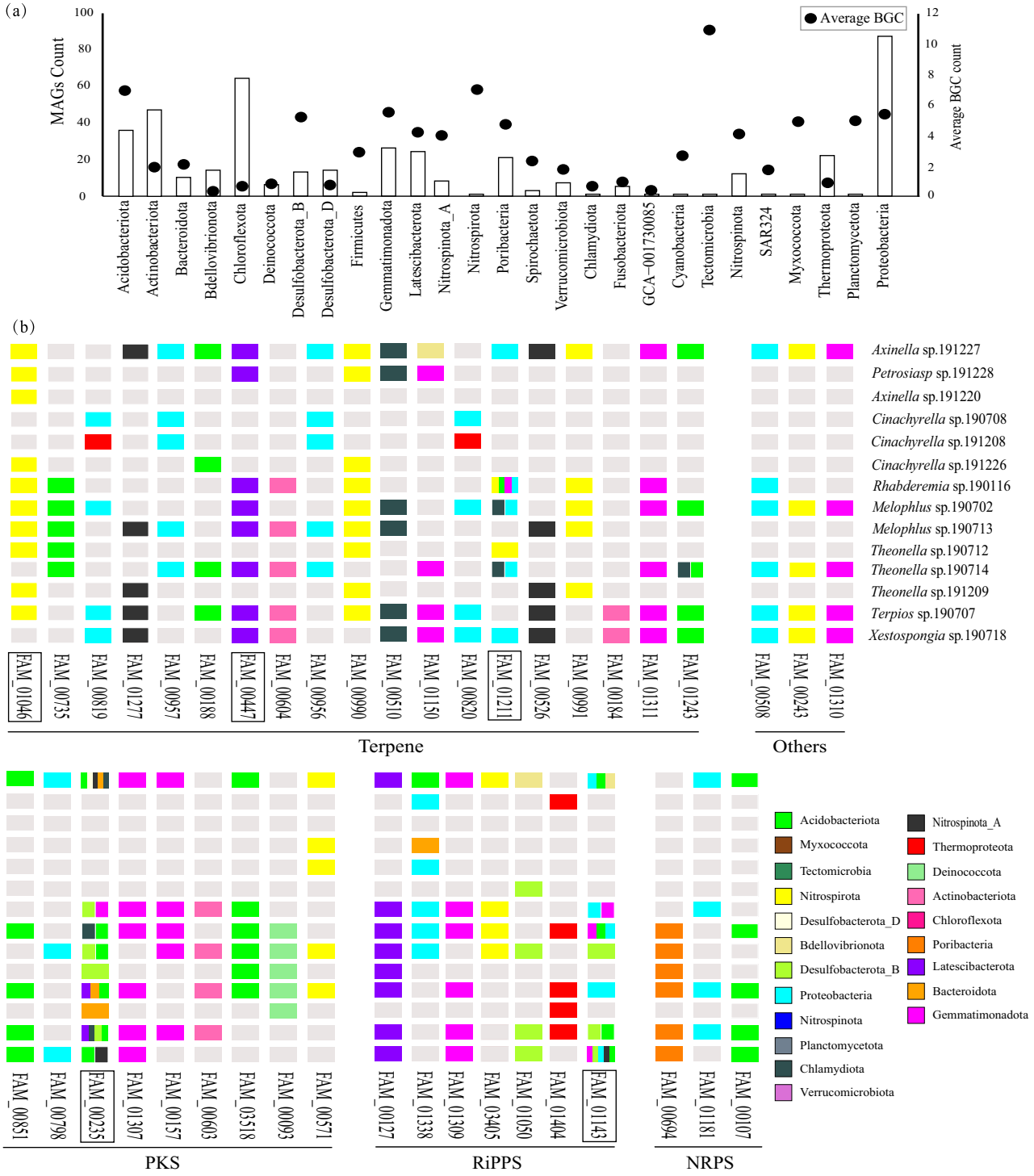

FIGURE 6 | The distribution pattern of biosynthetic gene clusters in sponges-associated bacteria. (a) Statistics of the average BGC of different bacterial phylum. The number of MAGs is indicated by the left y-axis (bar chart) and the average number of BGCs is indicated by the right y-axis (black dots). (b) The existence/absence matrix showed the distribution of BGCs from 41 GCFs among mesophotic sponge samples. The x-axis indicates the MAG names, and the y-axis indicates the sponge sample names. The taxonomy of the MAGs was marked by different colours in the bottom right corner. The five widely distributed GCFs are marked with squares.

of BGCs in mesophotic sponges. These findings offer valuable guidelines for discovering new bioactive compounds from mesophotic sponge-associated bacteria.

Sponge-associated bacteria have been estimated to constitute up to 35% of total sponge biomass (Vacelet and Donadey 1977; Hentschel et al. 2006; Dat et al. 2021; Freeman et al. 2021). However, isolating and culturing these bacteria under laboratory conditions remains a significant challenge, complicating the exploration of their chemical repertoire through traditional methods such as strain culturing, fermentation and compound isolation. To circumvent this limitation, we employed a

non-culture-based metagenomic approach to investigate the composition and distribution of symbiotic bacteria in mesophotic sponges. This analysis led to the identification of 429 MAGs (Table S3). Actinobacteriota, Desulfobacterota_D and Proteobacteria were found to be widely distributed, suggesting their role as core members of the MCE sponge microbiota. Distinct differences in microbial communities were also observed across sponge species; for instance, Chloroflexota was primarily associated with *Melophlus, Terpios and Xestospongia*, while Actinobacteriota was predominantly found in *Axinella, Rhabderemia and Terpios* (Figure 1). Given that all sponges were collected from the same region, these differences likely reflect host-specific and ecological factors rather than environmental influences, aligning with previous findings on the strong connection between sponge hosts and their microbiota (Reveillaud et al. 2014).

Sponge-associated bacteria are renowned for producing diverse bioactive compounds with potent bioactivities (Cheng et al. 2020; Li et al. 2023). Our analysis identified 1637 BGCs from the 429 MAGs, with the terpene family representing 30.2% of all identified BGCs, making it the most abundant type. Terpenes and terpenoids, known for their roles in signalling, defence, host–microbe interactions and stress responses (Avalos et al. 2022), are hypothesised to play essential roles in maintaining the symbiotic relationship between mesophotic sponges and their bacteria. Since mesophotic sponges mostly live in complex and variable marine environments, their associated bacteria are driven to synthesise diverse bioactive compounds to keep the symbiotic relationship healthy and safe. Thus, we hypothesise that the widespread presence of terpene BGCs in mesophotic sponge-associated bacteria suggests an as-yet-undiscovered function within the sponge-bacteria symbiotic system. This highlights their potential as a promising resource for discovering novel terpene-based bioactive compounds.

PKS and NRPS gene clusters also garnered significant attention due to their involvement in the biosynthesis of clinically important drugs such as erythromycin, tetracycline, penicillin, cephalosporins and vancomycin (Fischbach and Walsh 2006). They are composed of multiple catalytic domains, which coordinate to perform the roles, including activation, selection, condensation and transportation, as well as the modification of methylation, dehydration and reduction. Furthermore, type I PKS in bacteria usually exhibits a relatively collinear role between the PKS biosynthetic modules and the corresponding chemical structure catalysed by them, the new gene cluster usually means cryptic structures (Tao et al. 2023). Notably, we identified a giant PKS cluster group containing five GCFs that shared core genes but exhibited variations in adjacent genes, suggesting substantial chemical diversity in their products.

Additionally, we identified two novel trans-AT PKS gene clusters from a Verrucomicrobiota MAG (190,116 bin62, d_Bacteria; p_Verrucomicrobiota; c_Kiritimatiellae). This is only the second report of trans-AT PKS in Verrucomicrobiota, following the discovery of lasonolide A from *Candidatus Thermopylae lasonolidus* (Uppal et al. 2022). We proposed hypothetical chemical structures for two novel compounds (A-B) and their biosynthetic pathways, which indicate potential for novel chemical scaffolds and bioactivities. Verrucomicrobiota, with its relatively small average genome size of 2.95 Mb (Xue et al. 2024), appears to have evolved unique strategies to harbour large BGCs, providing competitive advantages in mesophotic sponge environments.

Ribosomally synthesised and RiPPs are well-known for their potent antibiotic activities and cryptic structures (Montalbán-López et al. 2021; Ongpipattanakul et al. 2022). However, within sponges and their associated bacteria, only a single type of RiPP, polytheonamides, has been identified (Freeman et al. 2016; Hamada et al. 2005). While Loureiro et al. (2022) investigated RiPP BGCs in three high-microbial-abundance (HMA) sponges from the Atlantic Ocean and Mediterranean Sea, revealing diversity among sponge-associated bacterial symbionts, the RiPP landscape in MCE sponge-associated bacteria is still unknown. In this study, we identified 61 GCFs (294 BGCs) along with 247 singletons from 429 MAGs, with proteusin comprising the majority and displaying substantial chemical diversity (Figure 5). Proteusins, a class of ribosomally synthesised bacterial peptides, are characterised by intricate structures arising from extensive post-translational modifications (PTMs) and exhibit potent antibiotic and cytotoxic activities (Haft et al. 2010; Hudson and Mitchell 2018). Known proteusin members include polytheonamide-type cytotoxins and landornamide A, both of which exhibit notable antiviral properties (Bösch et al. 2020; Hamada et al. 2005). The discovery of a large and diverse set of proteusin BGCs in MCE sponge-associated bacteria highlights the untapped potential of MCE sponges as a novel source for these compounds. Additionally, our findings also revealed a species-specific distribution of RiPPs at the MAG level, with higher prevalence in Nitrospirota, Acidobacteriota, Proteobacteria, Poribacteria and Gemmatimonadota. This uneven distribution suggests that RiPPs may play critical roles in the ecological fitness of their host bacteria, providing insights for targeted antibiotic discovery.

It has been shown that not all microorganisms exhibit a robust capacity to synthesise diverse secondary metabolites (Srinivasan et al. 2021; Wei et al. 2023). In this study, *Entotheonella sp.* (s191209.Bin93) from the Tectomicrobia phylum stands out with 11 BGCs, representing 2% of its genome. *Entotheonella* spp. are known symbionts of *Theonella swinhoei* and are prolific producers of diverse secondary metabolites, such as polytheonamides, keramamides, onnamides and cyclotheonellazoles (Wilson et al. 2014). These metabolites exhibit significant bioactivities, including cytotoxicity (onnamides against the P388 cell line) and protease inhibition (cyclotheonellazoles) (Matsunaga et al. 1992; Issac et al. 2017). Interestingly, the BGCs from *Entotheonella sp.* (s191209.Bin93) displayed distinct genetic profiles compared to those of *Entotheonella sp.* TSY1 and TSY2 (Table S5), suggesting evolutionary adaptations to the MCE environment. Besides, we found that Proteobacteria and Acidobacteriota, although possessing fewer BGCs per genome (6 and 7, respectively), represent rich sources of biosynthetic diversity due to their high MAG abundance in MCE sponges. Moreover, global analyses of marine prokaryotes and sponge-associated microbiota corroborate the exceptional chemical diversity harboured by these two phyla, marking them as promising targets for discovering bioactive compounds (Loureiro et al. 2022; Wei et al. 2023).

Sponge-associated bacteria contribute significantly to the sponge holobiont, often through the production of secondary

metabolites that serve as chemical defences against infections and predation (Fan et al. 2012; Pita et al. 2018; Rubin-Blum et al. 2019). In this study, we observed that 32 GCFs were conserved within the same MAGs across different sponge species, suggesting a link between habitat specificity and GCF conservation (Nguyen et al. 2021). Notably, GCFs such as FAM_00235 and FAM_01143/01149, homologous to strobilurin A and proteusin biosynthetic pathways, respectively, underscore the potential of mesophotic sponge-associated bacteria to encode bioactive compounds with cytotoxic and antiviral properties (Hamada et al. 2005). This highlights their potential role in host defence mechanisms and their broader ecological importance. Additionally, nine GCFs exhibited distribution across different MAGs and sponge samples, reflecting the metabolic diversity and adaptability of these clusters within the sponge microbiome (Dharamshi et al. 2022). The identification of conserved and diverse BGCs emphasises their potential in mediating sponge-microbe interactions and underscores their promise for biotechnological applications. However, the functional validation of these predicted pathways requires further experimental efforts to fully elucidate their biosynthetic potential and ecological roles.

Author Contributions

Nuo Chen: investigation, data curation, visualization, formal analysis, writing – review and editing. **Liwei Liu:** writing – original draft, writing – review and editing, conceptualization, methodology, project administration, supervision, funding acquisition, data curation, investigation, visualization, formal analysis. **Jingxuan Wang:** writing – original draft, data curation, formal analysis, investigation, visualization. **Deqiang Mao:** visualization, investigation, data curation, formal analysis. **Hongmei Lu:** investigation, formal analysis, data curation, visualization. **Tânia Keiko Shishido:** writing – review and editing, writing – original draft. **Shuai Zhi:** writing – review and editing, writing – original draft. **Hua Chen:** investigation, data curation, formal analysis, visualization. **Shan He:** resources, funding acquisition.

Acknowledgements

This work was supported by grants from the National Natural Science Foundation of China (31600016 and 41776168), Startup Foundation of Ningbo University (422010882, 422110473, and 422207513), the Li Dak Sum Yip Yio Chin Kenneth Li Marine Biopharmaceutical Development Fund, the National 111 Project of China (D16013) and Ningbo Natural Science Foundation (2021Z04). T.K.S. is funded by Novo Nordisk Foundation (NNF22OC0080109).

Ethics Statement

No animals or humans were involved in this study.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All the sequencing data have been deposited in NCBI, with BioProject accession number PRJNA924988, PRJNA928209, PRJNA928219, PRJNA926003, PRJNA928229, PRJNA928500, PRJNA925676, PRJNA928504, PRJNA928571, PRJNA928839, PRJNA928840, PRJNA928841, PRJNA928854, PRJNA929091 and PRJNA929093 (15 metagenomic sequencing data for 15 different mesophotic sponge samples).

References

Avalos, M., P. Garbeva, L. Vader, G. P. van Wezel, J. S. Dickschat, and D. Ulanova. 2022. "Biosynthesis, Evolution and Ecology of Microbial Terpenoids." *Natural Product Reports* 39: 249–272.

Aleti, G., J. L. Baker, X. Tang, et al. 2019. "Identification of the Bacterial Biosynthetic Gene Clusters of the Oral Microbiome Illuminates the Unexplored Social Language of Bacteria During Health and Disease." *mBio* 10: e00321-19.

Ayling, M., M. D. Clark, and R. M. Leggett. 2020. "New Approaches for Metagenome Assembly With Short Reads." *Briefings in Bioinformatics* 21: 584–594.

Blin, K., S. Shaw, A. M. Kloosterman, et al. 2021. "AntiSMASH 6.0: Improving Cluster Detection and Comparison Capabilities." *Nucleic Acids Research* 49: 29–35.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data." *Bioinformatics* 30: 2114–2120.

Bösch, N. M., M. Borsa, U. Greczmiel, et al. 2020. "Landornamides: Antiviral Ornithine-Containing Ribosomal Peptides Discovered Through Genome Mining." *Angewandte Chemie* 59: 11763–11768.

Bowers, R. M., N. C. Kyrpides, R. Stepanauskas, et al. 2017. "Minimum Information About a Single Amplified Genome (MISAG) and a Metagenome-Assembled Genome (MIMAG) of Bacteria and Archaea." *Nature Biotechnology* 35: 725–731.

Chaumeil, P. A., A. J. Mussig, P. Hugenholtz, and D. H. Parks. 2022. "GTDB-Tk v2: Memory Friendly Classification With the Genome Taxonomy Database." *Bioinformatics* 38: 5315–5316.

Cheng, M. M., X. L. Tang, Y. T. Sun, et al. 2020. "Biological and Chemical Diversity of Marine Sponge-Derived Microorganisms Over the Last Two Decades From 1998 to 2017." *Molecules* 25: 853.

Crooks, G. E., G. Hon, J. M. Chandonia, and S. E. Brenner. 2004. "WebLogo: A Sequence Logo Generator." *Genome Research* 14: 1188–1190.

Dat, T. T. H., G. Steinert, N. T. K. Cuc, H. Smidt, and D. Sipkema. 2021. "Bacteria Cultivated From Sponges and Bacteria Not Yet Cultivated From Sponges – A Review." *Frontiers in Microbiology* 12: 737925.

Dharamshi, J. E., N. Gaarslev, K. Steffen, T. Martin, D. Sipkema, and T. J. G. Ettema. 2022. "Genomic Diversity and Biosynthetic Capabilities of Sponge-Associated Chlamydiae." *ISME Journal* 16: 2725–2740.

Erpenbeck, D., J. A. J. Breeuwer, and R. W. M. van Soest. 2005. "Implications From a 28S rRNA Gene Fragment for the Phylogenetic Relationships of Halichondrid Sponges (Porifera: Demospongiae)." *Journal of Zoological Systematics and Evolutionary Research* 43: 93–99.

Fan, L., D. Reynolds, M. Liu, et al. 2012. "Functional Equivalence and Evolutionary Convergence in Complex Communities of Microbial Sponge Symbionts." *Proceedings of the National Academy of Sciences of the United States of America* 109: 1878–1887.

Fischbach, M. A., and C. T. Walsh. 2006. "Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms." *Chemical Reviews* 106: 3468–3496.

Freeman, C. J., C. G. Easson, C. L. Fiore, and R. W. Thacker. 2021. "Sponge-Microbe Interactions on Coral Reefs: Multiple Evolutionary Solutions to a Complex Environment." *Frontiers in Marine Science* 8: 705053.

Freeman, M. F., A. L. Vagstad, and J. Piel. 2016. "Polytheonamide Biosynthesis Showcasing the Metabolic Potential of Sponge-Associated Uncultivated '*Entotheonella*' Bacteria." *Current Opinion in Chemical Biology* 31: 8–14.

Gurgui, C., and J. Piel. 2010. "Metagenomic Approaches to Identify and Isolate Bioactive Natural Products From Microbiota of Marine Sponges." *Methods in Molecular Biology* 668: 247–264.

Haft, D. H., M. K. Basu, and D. A. Mitchell. 2010. "Expansion of Ribosomally Produced Natural Products: A Nitrile Hydratase- and Nif11-Related Precursor Family." *BMC Biology* 8: 70.

Hamada, T., S. Matsunaga, G. Yano, and N. Fusetani. 2005. "Polytheonamides A and B, Highly Cytotoxic, Linear Polypeptides With Unprecedented Structural Features, From the Marine Sponge, *Theonella Swinhoei*." *Journal of the American Chemical Society* 127: 110–118.

Hentschel, U., J. Piel, S. M. Degnan, and M. W. Taylor. 2012. "Genomic Insights Into the Marine Sponge Microbiome." *Nature Reviews. Microbiology* 10: 641–654.

Hentschel, U., K. M. Usher, and M. W. Taylor. 2006. "Marine Sponges as Microbial Fermenters." *FEMS Microbiology Ecology* 55: 167–177.

Hinderstein, L. M., J. C. A. Marr, F. A. Martinez, et al. 2010. "Theme Section on "Mesophotic Coral Ecosystems: Characterization, Ecology, and Management"." *Coral Reefs* 29: 247–251.

Hugenholtz, P., B. M. Goebel, and N. R. Pace. 1998. "Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity." *Journal of Bacteriology* 180: 4765–4774.

Hudson, G., and D. A. Mitchell. 2018. "RiPP Antibiotics: Biosynthesis and Engineering Potential." *Current Opinion in Microbiology* 45: 61–69.

Issac, M., M. Aknin, A. Gauvin-Bialecki, et al. 2017. "Cyclotheonellazoles A-C, Potent Protease Inhibitors From the Marine Sponge *Theonella aff. swinhoei*." *Journal of Natural Products* 80: 1110–1116.

Kalkreuter, E., G. Pan, A. J. Cepeda, and B. Shen. 2020. "Targeting Bacterial Genomes for Natural Product Discovery." *Trends in Pharmacological Sciences* 41: 13–26.

Kang, D. D., F. Li, E. Kirton, et al. 2019. "MetaBAT 2: An Adaptive Binning Algorithm for Robust and Efficient Genome Reconstruction From Metagenome Assemblies." *PeerJ* 7: e7359.

Katz, M., B. M. Hover, and S. F. Brady. 2016. "Culture-Independent Discovery of Natural Products From Soil Metagenomes." *Journal of Industrial Microbiology & Biotechnology* 43: 129–141.

Kautsar, S. A., H. G. Suarez Duran, and M. H. Medema. 2018. "Genomic Identification and Analysis of Specialized Metabolite Biosynthetic Gene Clusters in Plants Using plantiSMASH." *Methods in Molecular Biology* 1795: 173–188.

Laport, M., O. Santos, and G. Muricy. 2009. "Marine Sponges: Potential Sources of New Antimicrobial Drugs." *Current Pharmaceutical Biotechnology* 10: 86–105.

Letunic, I., and P. Bork. 2019. "Interactive Tree of Life (iTOL) v4: Recent Updates and New Developments." *Nucleic Acids Research* 47: 256–259.

Li, D., R. Luo, C. M. Liu, et al. 2016. "MEGAHIT v1.0: A Fast and Scalable Metagenome Assembler Driven by Advanced Methodologies and Community Practices." *Methods* 102: 3–11.

Li, H. 2013. "Aligning Sequence Reads, Clone Sequences and Assembly Contigs With BWA-MEM." arXiv: Genomics 1303.

Li, P. S., H. M. Lu, Y. Z. Zhang, et al. 2023. "The Natural Products Discovered in Marine Sponge-Associated Microorganisms: Structures, Activities, and Mining Strategy." *Frontiers in Marine Science* 10: 1191858.

López-Legentil, S., P. M. Erwin, T. P. Henkel, T. L. Loh, and J. R. Pawlik. 2010. "Phenotypic Plasticity in the Caribbean Sponge *Callyspongia vaginalis* (Porifera: Haplosclerida)." *Scientia Marina* 74: 445–453.

Loureiro, C., A. Galani, A. Gavriilidou, et al. 2022. "Comparative Metagenomic Analysis of Biosynthetic Diversity Across Sponge Microbiomes Highlights Metabolic Novelty, Conservation, and Diversification." *mSystems* 7: e0035722.

Marchesi, J. R. 2012. "Metagenomics: Current Innovations and Future Trends." *Future Microbiology* 7: 813–814.

Matsunaga, S., N. Fusetani, and Y. Nakao. 1992. "Eight New Cytotoxic Metabolites Closely Related to Onnamide A From Two Marine Sponges of the Genus *Theonella*." *Tetrahedron* 48: 8369–8376.

Medema, M. H., R. Kottmann, P. Yilmaz, et al. 2015. "Minimum Information About a Biosynthetic Gene Cluster." *Nature Chemical Biology* 11: 625–631.

Montalbán-López, M., T. A. Scott, S. Ramesh, et al. 2021. "New Developments in RiPP Discovery, Enzymology and Engineering." *Natural Product Reports* 38: 130–239.

Na, S. I., Y. O. Kim, S. H. Yoon, S. M. Ha, I. Baek, and J. Chun. 2018. "UBCG: Up-To-Date Bacterial Core Gene Set and Pipeline for Phylogenomic Tree Reconstruction." *Journal of Microbiology* 56: 280–285.

Navarro-Muñoz, J. C., N. Selem-Mojica, M. W. Mullowney, et al. 2020. "A Computational Framework to Explore Large-Scale Biosynthetic Diversity." *Nature Chemical Biology* 16: 60–68.

Nguyen, N. A., Z. Lin, I. Mohanty, N. Garg, E. W. Schmidt, and V. Agarwal. 2021. "An Obligate Peptidyl Brominase Underlies the Discovery of Highly Distributed Biosynthetic Gene Clusters in Marine Sponge Microbiomes." *Journal of the American Chemical Society* 143: 10221–10231.

Nurk, S., A. Bankevich, D. Antipov, et al. 2013. "Assembling Genomes and Mini-Metagenomes From Highly Chimeric Reads." *Computational Molecular Biology* 7821: 158–170.

Olson, J. B., and C. A. Kellogg. 2010. "Microbial Ecology of Corals, Sponges, and Algae in Mesophotic Coral Environments." *FEMS Microbiology Ecology* 73: 17–30.

Ongpipattanakul, C., E. K. Desormeaux, A. DiCaprio, W. A. van der Donk, D. A. Mitchell, and S. K. Nair. 2022. "Mechanism of Action of Ribosomally Synthesized and Post-Translationally Modified Peptides." *Chemical Reviews* 122: 14722–14814.

Paoli, L., H. J. Ruscheweyh, C. C. Forneris, et al. 2022. "Biosynthetic Potential of the Global Ocean Microbiome." *Nature* 607: 111–118.

Parks, D. H., M. Imelfort, C. T. Skennerton, P. Hugenholtz, and G. W. Tyson. 2015. "CheckM: Assessing the Quality of Microbial Genomes Recovered From Isolates, Single Cells, and Metagenomes." *Genome Research* 25: 1043–1055.

Patro, R., G. Duggal, M. I. Love, R. A. Irizarry, and C. Kingsford. 2017. "Salmon Provides Fast and Bias-Aware Quantification of Transcript Expression." *Nature Methods* 14: 417–419.

Pita, L., L. Rix, B. M. Slaby, A. Franke, and U. Hentschel. 2018. "The Sponge Holobiont in a Changing Ocean: From Microbes to Ecosystems." *Microbiome* 6: 46.

Reveillaud, J., L. Maignien, A. M. Eren, et al. 2014. "Host-Specificity Among Abundant and Rare Taxa in the Sponge Microbiome." *ISME Journal* 8: 1198–1209.

Rodriguez, R. L., S. Gunturu, J. M. Tiedje, J. R. Cole, and K. T. Konstantinidis. 2018. "Nonpareil 3: Fast Estimation of Metagenomic Coverage and Sequence Diversity." *mSystems* 3: e00039-18.

Rubin-Blum, M., C. P. Antony, L. Sayavedra, et al. 2019. "Fueled by Methane: Deep-Sea Sponges From Asphalt Seeps Gain Their Nutrition From Methane-Oxidizing Symbionts." *ISME Journal* 13: 1209–1225.

Rust, M., E. J. N. Helfrich, M. F. Freeman, et al. 2020. "A Multiproducer Microbiome Generates Chemical Diversity in the Marine Sponge." *Proceedings of the National Academy of Sciences of the United States of America* 117: 9508–9518.

Shaffer, J. P., L. F. Nothias, L. R. Thompson, et al. 2022. "Standardized Multi-Omics of Earth's Microbiomes Reveals Microbial and Metabolite Diversity." *Nature Microbiology* 7: 2128–2150.

Shannon, P., A. Markiel, O. Ozier, et al. 2003. "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks." *Genome Research* 13: 2498–2504.

Srinivasan, R., A. Kannappan, C. Shi, and X. Lin. 2021. "Marine Bacterial Secondary Metabolites: A Treasure House for Structurally Unique and Effective Antimicrobial Compounds." *Marine Drugs* 19: 530.

Storey, M. A., S. K. Andreassend, J. Bracegirdle, et al. 2020. "Metagenomic Exploration of the Marine Sponge Uncovers Multiple Polyketide-Producing Bacterial Symbionts." *mBio* 11: e02997-19.

Tao, X. B., S. LaFrance, Y. Xing, et al. 2023. "ClusterCAD 2.0: An Updated Computational Platform for Chimeric Type I Polyketide Synthase and Nonribosomal Peptide Synthetase Design." *Nucleic Acids Research* 51: 532–538.

Uppal, S., J. L. Metz, R. K. M. Xavier, et al. 2022. "Uncovering Lasonolide a Biosynthesis Using Genome-Resolved Metagenomics." *mBio* 13: e0152422.

Vacelet, J., and C. Donadey. 1977. "Electron Microscope Study of the Association Between Some Sponges and Bacteria." *Journal of Experimental Marine Biology and Ecology* 30: 301–314.

Varijakzhan, D., J. Y. Loh, W. S. Yap, et al. 2021. "Bioactive Compounds From Marine Sponges: Fundamentals and Applications." *Marine Drugs* 19: 246.

Waschulin, V., C. Borsetto, R. James, et al. 2022. "Biosynthetic Potential of Uncultured Antarctic Soil Bacteria Revealed Through Long-Read Metagenomic Sequencing." *ISME Journal* 16: 101–111.

Watkins, R. F., and A. T. Beckenbach. 1999. "Partial Sequence of a Sponge Mitochondrial Genome Reveals Sequence Similarity to Cnidaria in Cytochrome Oxidase Subunit II and the Large Ribosomal RNA Subunit." *Journal of Molecular Evolution* 48: 542–554.

Wei, B., A. Q. Du, Z. Y. Zhou, et al. 2021. "An Atlas of Bacterial Secondary Metabolite Biosynthesis Gene Clusters." *Environmental Microbiology* 23: 6981–6992.

Wei, B., G. A. Hu, Z. Y. Zhou, et al. 2023. "Global Analysis of the Biosynthetic Chemical Space of Marine Prokaryotes." *Microbiome* 11: 144.

Wilson, M. C., T. Mori, C. Rückert, et al. 2014. "An Environmental Bacterial Taxon With a Large and Distinct Metabolic Repertoire." *Nature* 506: 58–62.

Xie, J., J. Wang, F. Zhao, et al. 2022. "Metagenomic Analysis of Gut Microbiome in Gout Patients With Different Chinese Traditional Medicine Treatments." *Evidence-Based Complementary and Alternative Medicine* 2022: 6466149.

Xue, D., L. Peng, J. X. Wang, et al. 2024. "Genome Mining Analysis Uncovers the Previously Unknown Biosynthetic Capacity for Secondary Metabolites in Verrucomicrobia." *Marine Biotechnology* 26: 1324–1335.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.