bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

The Potential of CCA-associated Bacteria to Fight Antimicrobial-Resistant Pathogens: a Genomic Survey

Diego Lera-Lozano[1], Jordan S. Ruiz-Toquica[2,3], Matthew W. Holt[4], Elizabeth K. Jones[4], Kim B. Ritchie[4], Mónica Medina[1,2], Raúl A. González-Pech[1]

[1]Department of Biology, Penn State University, University Park, Pennsylvania, USA
[2]Faculty of Natural Sciences and Engineering, Universidad de Bogotá Jorge Tadeo Lozano, Bogota, Colombia
[3]Department of Biology and Marine Biology, University of North Carolina Wilmington, Wilmington, North Carolina, USA
[4]Department of Natural Sciences, University of South Carolina Beaufort, Beaufort, South Carolina, USA

ABSTRACT

The global rise of antimicrobial resistance has intensified efforts in bioprospecting, with researchers increasingly exploring unique marine environments for novel antimicrobials. In line with this trend, our study focused on bacteria isolated from the unique microbiome of crustose coralline algae (CCA), which has yet to be investigated for antimicrobial discovery. In the present work, bacteria were isolated from the CCA *Hydrolithon boergesenii*. After running antimicrobial assays against antibiotic-resistant human and marine pathogens, three isolates were selected for genome sequencing using the Oxford Nanopore technology. Genome mining of the high-quality assemblies revealed 100 putative Biosynthetic Gene Clusters (BGCs) across the three genomes. Further analysis uncovered BGCs potentially involved in the biosynthesis of novel antimicrobial compounds. Our study provides valuable resources for future research aimed at the discovery of novel antimicrobials, particularly in the face of the antibiotic-resistance global crisis and highlights the potential of specialized marine environments like CCA biofilms.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

INTRODUCTION

Antimicrobial resistance poses a critical threat to public health globally, with estimates suggesting that by 2050, it could claim up to 10 million lives annually and cause substantial economic losses, impacting healthcare and human development [1][2]. In response to this growing crisis, the urgent need for new antimicrobials, particularly novel structural classes that can circumvent existing resistance mechanisms, has become paramount [3]. Considering this, bioprospecting efforts have intensified in pursuit of new antimicrobial compounds with the potential to target antimicrobial-resistant pathogens.

Over the past century, scientists have identified natural products with a wide variety of use cases, including agricultural implementations (e.g., pesticides), nutraceuticals (e.g., dietary supplements), and notably, pharmaceuticals (e.g., antibiotics). Bioprospecting has been critical in these discoveries. In particular, bioactive secondary metabolites have been identified from bacteria living in environments with a high microbial diversity due to their broad metabolic spectra [4]. For example, the commonly used antibiotic vancomycin was originally recovered from a soil bacterium [5]. Similarly, the antiviral scytovirin, used to treat HIV, was derived from the freshwater cyanobacterium *Scytonema varium* [6]. Bacteria living in marine sediments contributed to the establishment of the antitumor antibiotics marinomycins [7].

Coral reefs are composed of several environmental niches that support a diverse array of both macro- and microorganisms. Recent research has revealed that these ecosystems harbor a rich microbial community compared to many other environments [8]. This microbial richness is characterized by substantial heterogeneity in community composition and metabolic potential across different microenvironments within a single coral reef [9]. Microbial communities associated with reefs have been identified as promising sources of antimicrobial compounds. For example, an initial screening of culturable bacterial strains from the elkhorn horal (*Acropora palmata*) revealed that 20% exhibited antibiotic activity [10]. Additionally, coral-associated bacteria, like *Pseudovibrio* sp. P12, have yielded compounds such as tropodithietic acid, which effectively inhibits various coral pathogens resistant to commercial antibiotics [11].

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

While reef organisms like sponges and soft corals have provided notable antibiotics including halichondrin B and eleutherobin, respectively [12][13], microorganisms associated with other reef organisms have been relatively underexplored for novel antimicrobials. A prime example are crustose coralline algae (CCA). CCAs host diverse bacterial communities with significant ecological and biotechnological potential [14][15][16]. These bacteria produce unique metabolites that contribute to host defense and health. While some CCA-associated bacteria like *Propionibacterium* and *Planomicrobium* show antimicrobial activity [17][18][19], research has primarily focused on their role in coral larval settlement [20][21] and antifouling properties [22]. The rich microbial diversity and metabolite production in CCA microenvironments offer promising opportunities for novel antimicrobial discovery.

Biosynthetic gene clusters (BGCs) play a crucial role in microbial adaptation and survival [23]. They regulate the production of secondary metabolites, i.e., compounds not directly involved in the growth or reproduction of microorganisms but essential for their environmental interactions, such as preventing overgrowth by other microbes [24]. Their presence might be especially important in environments with intricate microbial interactions, such as CCA biofilms. Because of the diverse roles of secondary metabolites, BGCs can be implicated in the synthesis of a wide array of bioactive compounds, with antimicrobials among the most prominent [24]. Consequently, the identification of BGCs in bacterial genomes is an integral part of the bioprospecting process, particularly when the product responsible for antimicrobial traits remains unknown [24].

In this study, CCA-associated bacteria were isolated and tested for antimicrobial properties against antibiotic-resistant and other human and marine pathogens. Selected isolates were subjected to whole-genome sequencing to explore features that might explain their antimicrobial phenotypes. Several putative BGCs were identified, some potentially involved in the production of novel antimicrobial compounds. This study sets the foundation for future research aimed at discovering new antimicrobials and highlights the need to target niche marine environments, such as CCA, for bioprospecting.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

MATERIALS & METHODS

Sample Collection, Bacteria Isolation, and Culture Conditions

The CCA *H. boergesenii* was isolated from coral rubble at 8 m at the Veradero Reef in the Cartagena Bay, Colombia, on March 7th, 2023. CCA fragments were stored in sterile 15 ml screw-capped tubes and swabbed within 3 h of collection using sterile cotton swabs, followed by culture onto marine agar (Difco) plates. Bacteria were subcultured to purification and stored in 96-well plates in 20% glycerol diluted in marine broth. CCA were taxonomically identified via microscopy.

Screening for Antimicrobial Activity

Each of the bacterial isolates were analyzed for antibacterial properties against several human and marine pathogens using an agar-overlay assay method. Culture libraries from 96-well plates were replica plated onto sterile marine agar and allowed to grow for 6 days. Cultures were then exposed to UV radiation for one hour followed by overlay with 0.8% soft agar containing one of the following pathogenetic test strains: *Escherichia coli* (EC), *Bacillus subtilis* (BS), Vancomycin resistant *Enterococcus* (VRE), *Enterococcus* (ENT), Methicillin-resistant *Staphylococcus aureus* (MRSA), Methicillin-sensitive *Staphylococcus aureus* (MSSA), *Vibrio vulnificus*, *Vibrio shiloi*, *Vibrio parahaemolyticus*, and *Serratia marcescens* PDL100. Zones of growth inhibition were measured in mm after 2 days. Test strains (shown in Table 1) were grown overnight at 37 °C (human pathogens) or 25 °C (marine pathogens) in 5 mL of strain-specific culture broth. Aliquots of each broth culture were inoculated into 0.8% agar containing marine broth, Luria broth or tryptic soy broth, depending on the test strain. CCA library plates were overlaid with approximately 10 mL of inoculated agar. Following overnight growth, plates were analyzed for antibacterial activity and zones of inhibition were measured in millimeters.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

Genomic DNA Extraction and Sequencing

Genomic DNA (gDNA) was extracted from 24 h-marine agar cultures using the DNeasy Powersoil® Pro Kit (Qiagen, USA) with modifications to the cell lysis steps. Briefly, two to three colonies of each isolate were picked and suspended in 500 µl into the spun PowerBead Pro tubes and mixed with 800 µL of the CD1 solution. Tubes were then placed in a bead-beater vortex and shaken as follows: 1) 5 min at 13,000 rpm, 2) a 2 min pause, 3) incubation for 1 min at 4°C, 4) 5 min at 13,000 rpm again, and 5) a final incubation at 4°C for 2 min. After this step, we followed the manufacturer's instructions. DNA was resuspended in 80 µL of elution buffer and quantified and quality-evaluated by electrophoresis in 0.8% agarose gels with 0.5 X TBE buffer (45 mM Tris-borate, 45 mM boric acid, 1 mM EDTA, pH 8.2) and using a Qubit fluorometer (Invitrogen™). The extracted DNA was further cleaned using a DNA Clean & Concentrator-5 kit (Zymo Research, Irvine, CA, USA) for high-purity absorbance ratios (260/280 = 1.7 and 260/230 = 2.2). DNA was then stored at -20°C until whole-genome sequencing on an Oxford Nanopore MinION Mk1B. Briefly, high-molecular-weight gDNA libraries were prepared using the Native Barcoding Kit 24 V14 (SQK-NBD114.24, Oxford, UK). Up to 400 ng of gDNA from each isolate were used to achieve equimolar concentration. Then, gDNA end-repair and ligation of native barcodes and adapters were done following manufacture's protocol instructions. The libraries were then pooled and loaded into the primed R10.4.1 flow cell. MinKNOW v23.07.15 was used to monitor sequencing and for demultiplexing. Base calling was conducted separately using Dorado v0.4.3 [25] implementing the high-accuracy model dna_r10.4.1_e8.2_400bps_sup@v4.2.0 (optimized for Duplex pairs) with default settings other than disabling CUDA acceleration.

Read Processing and Genome Assembly

Quality of raw reads was assessed using FastQC v0.12.0 [26]. Adapters were removed and reads were trimmed using chopper v0.7.0 [27], removing 100 bp from both 5′ and 3′ ends. Genome size and sequencing depth for each isolate were estimated using Jellyfish v2.3.0 [28] with k = 25 and GenomeScope 2.0 [29]. De novo genome assembly was done

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

using Flye v2.9.2 [30], followed by circularization of complete molecules using Circlator v1.5.5 [31]. Quality of the final assemblies was assessed using CheckM v1.2.2 [32]. To determine whether contigs were of chromosomal or plasmid origin, we aligned each contig against the NCBI nucleotide database using BLASTn v2.16.0 [33]. Taxonomic assignment was done using GTDB-tk v2.3.2 [34] and the GTDB database release 207 [35].

Gene Prediction and Functional Annotation

Genome annotation was done using prokka v1.14.5 [36] at default settings. Metabolic pathway analysis was done by first annotating predicted proteins with KEGG orthology using the blastKOALA web service [37] and then reconstructing metabolic pathways using the online KEGG mapper [38]. COGclassifier v1.0.5 [39] was used to assign predicted proteins into clusters of genes (COG) with annotated gene functions. COGs were also predicted from the genome of *Escherichia coli* strain K-12 substrain MG1655 (NCBI RefSeq Accession GCF_904425475) as a reference of a bacterium with little to none known secondary metabolite production [40]. Genes implicated in the production of tetrabromopyrrole (TBP)[41] were manually searched in the prokka annotations.

Biosynthetic Gene Clusters Identification and Network Analysis

Biosynthetic gene clusters (BGCs) were identified using antiSMASH v7.0.0 [42], DeepBGC v0.1.29 [43], and GECCO v0.9.10 [44]. Circular visualization of genome assemblies and features was done in R using shinyCircos v2.0 [45]. BiG-SCAPE v1.1.9 [46] and the MIBiG database release 2.1 inclusion [47] were used for gene similarity network analyses based on the BGCs identified by antiSMASH, DeepBGC, and GECCO. One run was done using all BGCs as input and a gene similarity cutoff of 0.7; networks resulting from this run were considered *conservative*. *Less-conservative* networks were obtained from a second run with the same input but a similarity cutoff of 0.3. Clusters resulting from the similarity network analyses were visualized using Cytoscape v3.7.0 [48].

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

RESULTS

CCA-Associated Bacteria Inhibit Growth of Human Pathogens

The 14 bacterial strains isolated from the CCA *Hydrolithon boergesenii* were found to produce several zones of inhibition against the test pathogens with the results listed in Table 2. Three isolates in particular displayed the capacity to inhibit three or more of the test strains and were selected for subsequent genomic characterization.

Genomic Characterization of CCA Bacterial Isolates

Taxonomic assignment resolved the bacterial isolates to the species level as listed in Table 3. Based on an Average Nucleotide Identity (ANI) cutoff of $\geq$95%, strain CCAH7 was identified as *Pseudovibrio denitrificans* (ANI = 95.29%), CCAH11 as *Pseudoalteromonas elyakovii* (ANI = 97.29%), and CCAH4 as *Rossellomorea marisflavi* (ANI = 98.92%). Genomic features of the three bacterial isolates are summarized in Table 3 and Fig. 1. The genomes were assembled into 3-4 contigs with N50 lengths ranging from 3.16 to 4.9 Mbp and high completeness (98.64%-100%). All assembled contigs in *P. denitrificans* CCAH7 consisted of entire circular molecules, as did one plasmid contig of *P. elyakovii CCAH11*; no circular molecules were recovered for *R. marisflavi* CCAH4 as listed in Table 3. The predicted protein-coding sequences (CDS) ranged between 4,435 and 5,331 per isolate.

COG classification assigned 80.5% of the CDS into functional categories (Fig. 2). The distributions of COG functional categories in chromosomes of *P. denitrificans* CCAH7 and *R. marisflavi* CCAH4 were similar to that of the BGC-deficient *Escherichia coli* K-12 substrain MG1655. Amino acid transport and metabolism, Transcription, and Carbohydrate transport and metabolism were the most represented categories in these three genomes, but Cell cycle control, cell division, chromosome partitioning was among the least abundant in *P. denitrificans* and *E. coli*, and among the most abundant in *R. marisflavi*. The distribution of COG functional categories in *P. elyakovii* was notably different, with Signal transduction mechanisms, Amino acid transport and metabolism, and Cell wall/membrane/envelope biogenesis being the most represented. COG

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

functional categories in plasmids of the three isolates were very uneven, potentially due to the highly variable molecule sizes (Supp. Fig. 1).

Potential Genomic Basis of Antimicrobial Activity

KEGG metabolic pathway analysis revealed that the genome of *P. elyakovii* contains genes pertaining to *biosynthesis of siderophore group nonribosomal peptides*; within this pathway, all genes involved in the production of myxochelin A and B were present (Supp. Fig. 2). The genome of *R. marisflavi* contained *carotenoid biosynthesis* genes that suggest production of lycopene by this bacterium (Supp. Fig. 3). Also, the genomes of all three isolates encoded *streptomycin synthesis* genes, though no isolate contained all genes required for actual streptomycin synthesis (Supp. Fig. 4).

The genomes of the three isolates were further investigated to assess their biosynthetic potential in regard to secondary metabolite production. In total, 130 putative BGCs were predicted by several tools (see Methods) with different types of biosynthetic classifications. BGCs identified by antiSMASH were classified into 14 biosynthetic cores (Fig. 3A)(Tables S1, S2, S3), whereas BGCs identified by GECCO (Table S4) and deepBGC (Table S5) were classified into 5 biosynthetic classes, a few pertaining to one or more classes (Fig. 3B, C). Some of the more specific cores found by antiSMASH corresponded to 4 of the classes found by GECCO and deepBGC (Fig. 3). Five BGCs had considerable gene similarity ($\geq$50%) to MiBiG BGCs implicated in the production of a range of metabolites, listed in Table 4.

The *conservative* similarity network analysis revealed that most BGCs are unique with no similarity to reference clusters and that there was substantial redundancy between predictions from the different tools (Fig. 4A). Unique BGCs in the similarity network (i.e., singletons) encompassed 100 BGCs. Redundancy between predictions was showcased by twelve pairings that grouped BGCs predicted by different tools but with the same genomic location. Further, three clusters (each containing three BGCs), grouped a MiBiG reference BGC, and an antiSMASH and a GECCO prediction, the two latter also occurred at the same genomic location. While 14 BGCs grouped with MIBiG clusters in the

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

less-conservative network, most BGCs were still classified as singletons, confirming the little representation of the identified BGCs in the reference database (Fig. 4B). Based on the network analyses and genomic locations of the predicted BGCs, we determined that a total of 115 non-redundant BGCs were predicted from these bacterial genomes: 55 from *P. denitrificans*, 33 from *P. elyakovii*, and 27 from *R. marisflavi*; no BGC was shared between isolate genomes.

DISCUSSION

The present study investigated culturable microorganisms from the CCA *Hydrolithon boergesenii*. as a first bioprospecting step. Three of them exhibited antagonistic activity against several human and marine pathogens, some of which were resistant to common antimicrobials in the market. To further assess the antimicrobial potential of these three isolates, their genomes were sequenced. Genes and BGCs potentially underpinning their antimicrobial phenotype were identified. Below, we discuss how these genes and BGCs might be implicated in the antimicrobial activity of these bacteria and in other aspects of their ecology.

Functional Roles of CCA Bacterial Isolates

The different gene functions and metabolic pathways found in each isolate (Fig. 2, Supp. Fig. 5) suggests that the isolates are occupying distinct niches in the CCA microenvironment. In previous work, *Pseudovibrio denitrificans* has been identified as an important stabilizing agent of reef ecosystems through its coral settlement induction properties [49][50], while simultaneously inhibiting larval settlement of other invertebrates, like barnacles and bryozoans [51]. Coral settlement and metamorphosis has been largely attributed to the secondary metabolite tetrabromopyrrole (TBP) [52]. Nevertheless, no genes (*bmp1-10* [41]) or BGCs involved in the generation of TBP could be found (Table 4). This observation indicates that: (i) this particular strain is not able of inducing coral larvae settlement, (ii) the biosynthetic pathways of TBP are not fully known yet, or (iii) other inducers yet to be fully characterized might be responsible for

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

this trait [53][54]. Future research on this *P. denitrificans* strain should therefore focus on verifying whether this bacterium can produce TBP and whether it can induce coral settlement, which could lead to the description of new coral larvae settlement inducers.

Various members of the genus *Pseudoalteromonas* have also been shown to induce coral settlement using TBP as a chemical cue [21]. However, similar to *P. denitirifcans*, no genes involved in TBP production were found in the *Pseudoalteromonas elyakovii* strain isolated in this study. Beyond this, *P. elyakovii* has been previously found in the mussel *Crenomytilus grayanus* and in the brown seaweed *Laminaria japonica* [55]; this is the first instance, to our knowledge, that this bacterium has been isolated from a member of the phylum Rhodophyta. Within *Laminaria*, *P. elyakovii* has been isolated from spot-wounded fronds, suggesting that *P. elyakovii* may be causing damage to this alga. Further, our analyses of metabolic pathways (Supp. Fig. 2) and BGCs (Table 4) suggest that this isolate may be actively acquiring iron from the environment and/or host through siderophores like myxochelin A and B, enterochelin, and pseudochelin A. Iron acquisition mechanisms drive potential virulence of well-established pathogens [56][57]. This bacterium might thus be an algal pathogen.

The genome of the other bacterium recovered here, *Rossellomorea marisflavi*, contained genes implicated in the resistance to arsenic (i.e., *arsB*, *arsC*, and *arsR*) and in the production of the antioxidant carotenoid lycopene (Table 4 and Supp. Fig. 3). A genomic characterization of another member of the same genus, *Rossellomorea* sp. y25, found similar metabolic capabilities, with added resistance to cadmium and zinc [58]. These genetic features were hypothesized to play a role in the adaptation to deep-sea environments, where this bacterium was isolated from. In *R. marisflavi*, resistance to heavy metals might be associated with the potential for bioaccumulation of trace elements of CCA [59]. Though we are not aware of any reports of arsenic in these algae, contamination with heavy metals has occurred several times in Cartagena Bay [60][61][62][63], where the CCA analyzed here was sampled. Lycopene production, on the other hand, might be related to its ability to prevent the lethal action of UV radiation on bacterial cells [64] given the high-light environments where CCA inhabit.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

Antimicrobial Activity is Likely Due to BGCs with Unknown Functions

Metabolic pathway exploration showed no prominent representation of genes involved in antibiotic biosynthesis in either of the three CCA isolates (Supp. Figs. 6-8). Nevertheless, metabolic complementation [65] among members of the CCA microbial consortium might be a mechanism leading to the biosynthesis of some antimicrobials. For instance, while none of the three isolates possessed the complete pathway for biosynthesis of streptomycin, all of them contained several of the implicated genes (Supp. Fig. 4). However, antibiotic biosynthetic pathways are not likely to explain the antimicrobial activity observed in the three isolates.

Antimicrobial activity is often a result of products of BGCs. The most frequent BGC classes identified in this study included non-ribosomal peptides (NRPs), ribosomally synthesized and post-translationally modified peptides (RiPPs), polyketides (PKS), saccharides, and terpenes [23]. These biosynthetic classes are commonly identified in bioprospecting studies of prokaryotes in marine environments [66]. Nevertheless, the number of BGCs recovered in the genomes of the three isolates studied here was unusually high. For example, BGCs predicted by antiSMASH ranged between five (in *R. flavimaris*) and 13 (in *P. elyakovii*), whereas a former large-scale study predicted 18,043 BGCs from 5,743 microbial genomes (i.e., on average 3.14 BGCs per genome) [67].

Only five BGCs could be annotated with a known function (Table 4), from which three were predicted to biosynthesize known secondary metabolites. One of the BGCs, identified in the genome of *P. denitrificans*, is involved in the production of pseudovibriamide A1-6 and B1-6, which are hypothesized to play a role in surface motility [68]. The other two BGCs with known function were found in the genome of *P. elyakovii*. One of them is implicated in the production of catecholate-type siderophores, like myxochelin A and B, and pseudochelin A [69]. While these siderophores are crucial for bacterial iron homeostasis [69], they also exhibit weak activity against Gram-positive bacteria [70]; yet *P. elyakovii* inhibited both Gram-positive and -negative pathogens in antagonistic assays. The other BGC is associated with the production of bromoalterochromide A. *Pseuedoalteromonas* bromoalterochromides often have

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

antibacterial activity [71], but no antimicrobial properties have been observed in bromoalterochromide A [72]. In addition, the carotenoid lycopene (Table 4 and Supp. Fig. 3) in *R. marisflavi* has been shown to have antimicrobial activity against Gram-negative bacteria, specifically against *Escherichia coli* [73] and members of the Chlamydiaceae family, such as *Chlamydia trachomatis* and *Chlamydia pneumoniae* [74]. While this carotenoid could explain some of the antimicrobial activity observed in the antagonistic assays, the *R. marisflavi* strain here did not inhibit the *E. coli* test strain, suggesting that *R. marisflavi* was not producing lycopene during the assays despite having the genomic capability. Therefore, the observed antimicrobial phenotypes are unlikely due to these BGCs.

Though biosynthetic pathways and known BGCs might underpin some of the antimicrobial activity revealed by the antagonistic assays, we hypothesize that this activity is most likely attributed to BGCs with unknown function. Given that most characterized biosynthetic pathways and BGCs have already been extensively explored for commercial antimicrobial production, these known compounds likely contribute to the existing resistance profiles of the tested pathogens. Hence, we believe that unknown metabolites are the ones to which the tested pathogens have not been commonly exposed to and are therefore susceptible to. BGCs can produce a wide range of metabolites [23], many of which remain to be explored; therefore, they could be responsible for this observed antimicrobial activity. Of special interest are the singletons, unique BGCs that appear only once in our dataset and show no significant similarity to known clusters. These singletons are particularly promising candidates for novel antimicrobial discovery as they represent previously uncharacterized biosynthetic targets.

Future Directions

The advancement of bioprospecting efforts largely relies on deepening our understanding of fundamental microbial ecology. Certain microhabitats may be more prone to harboring novel antimicrobials, yet it remains difficult to identify which to target for bioprospecting. This knowledge gap extends beyond the ecology of microorganisms to the genomic basis of natural products, such as BGCs. Many BGCs remain

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

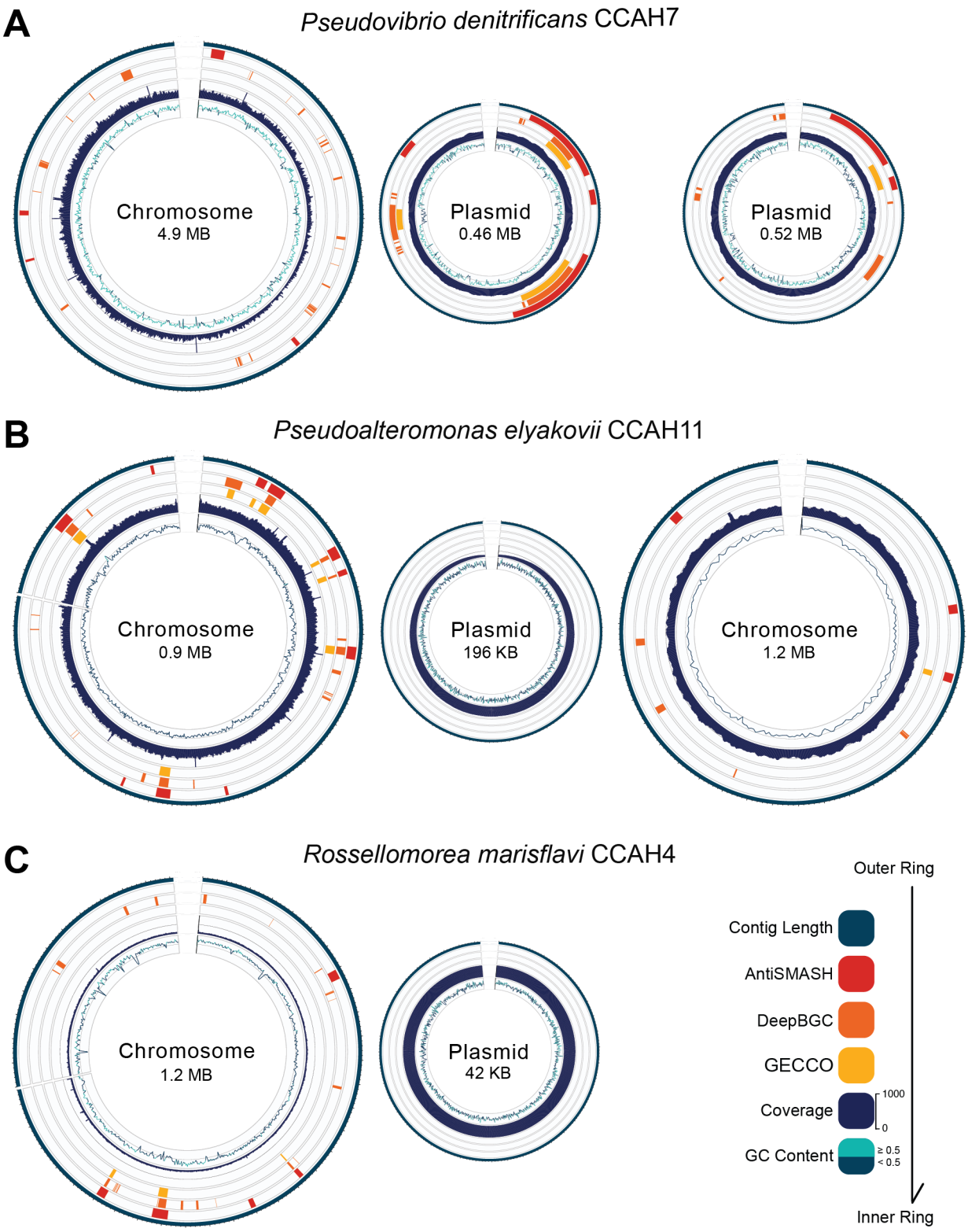
uncharacterized and their prediction from genomic data is based on similarity to other known BGCs, which hinders bioprospecting efforts as this study supports.

One priority moving forward should be the consistent curation and functional characterization of unknown BGCs. To further the breadth of BGCs we identify, effort should be placed in employing multiple complementary approaches rather than relying on a single tool. This strategy allows for more comprehensive identification of BGCs that might be missed when using a single method. Furthermore, new methods that expedite the bioprospecting pipeline are also warranted. In this study, we explored several methods to identify BGCs of interest and demonstrated how network analyses can effectively prioritize promising BGCs by comparing their similarity to previously characterized clusters. This approach enables more targeted investigation of potentially novel compounds while reducing the resources spent on redundant or less promising BGCs. Finally, though costly, the development and enrichment of BGC databases and their continuous curation will provide an essential foundation for more targeted and efficient antimicrobial discovery efforts.

In summary, our study highlights the potential of crustose coralline algae (CCA) as a valuable target for antimicrobial discovery. We highlight the wide range of metabolic potential of CCA-associated bacteria through metabolic pathway analysis and the identification of 97 singleton BGCs. These singletons represent promising targets for unlocking novel biosynthetic potential in CCA-associated bacteria.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

Figures



bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

FIG 1 Circular genome representations of *Pseudovibrio denitrificans* **(A)**, *Pseudoalteromonas elyakovii* **(B)**, and *Rossellomorea marisflavi* **(C)**, depicting the distribution of biosynthetic gene clusters (BGCs) across chromosomes and plasmids. Each representation displays, from outer to inner rings, contig length, BGCs identified by AntiSMASH, DeepBGC, and GECCO, sequencing coverage, and GC content, following the bottom right color code.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

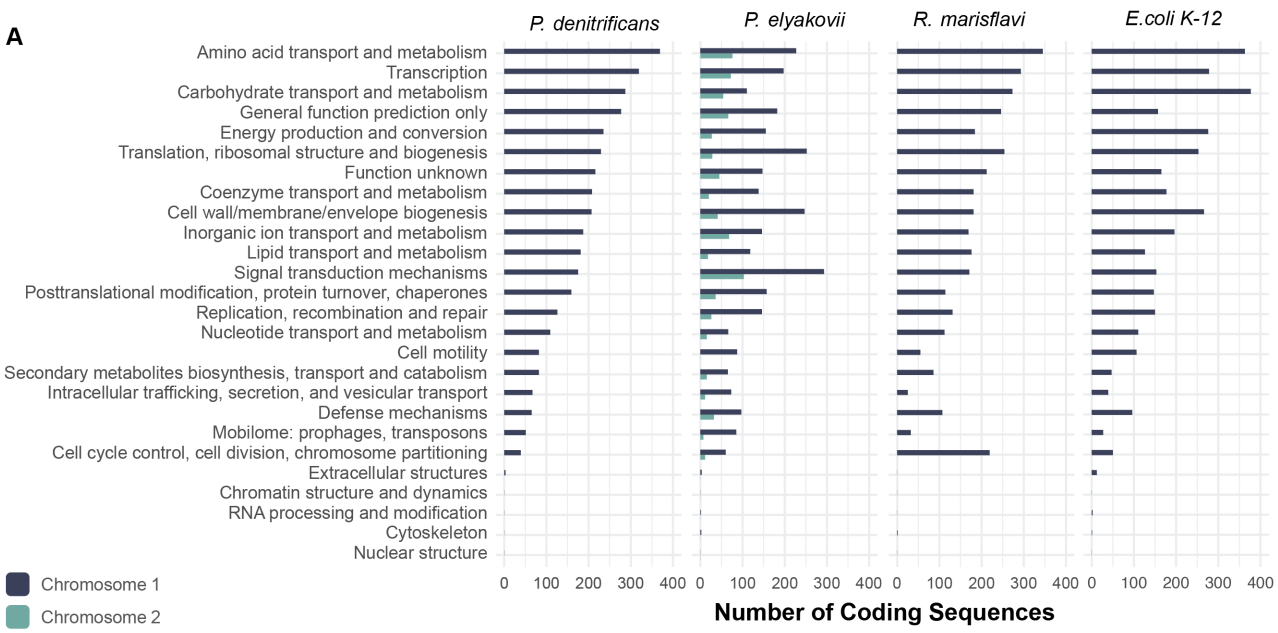


FIG 2 Representation of gene functions in genomes of the selected bacterial isolates and *Escherichia coli* strain K-12 substrain MG1655. Number of chromosomal protein-coding sequences (CDS) in functional categories determined by COGclassifier.

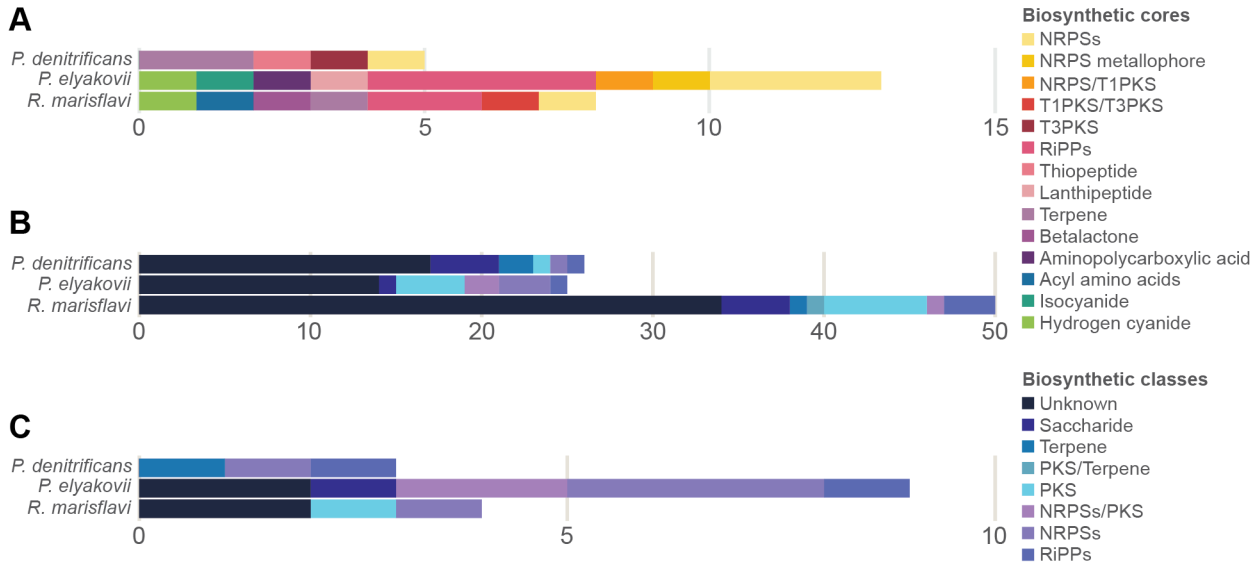


FIG 3 Counts of BGC cores and classes predicted for the three isolates using AntiSMASH **(A)**, DeepBGC **(B)**, and GECCO **(C)**.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

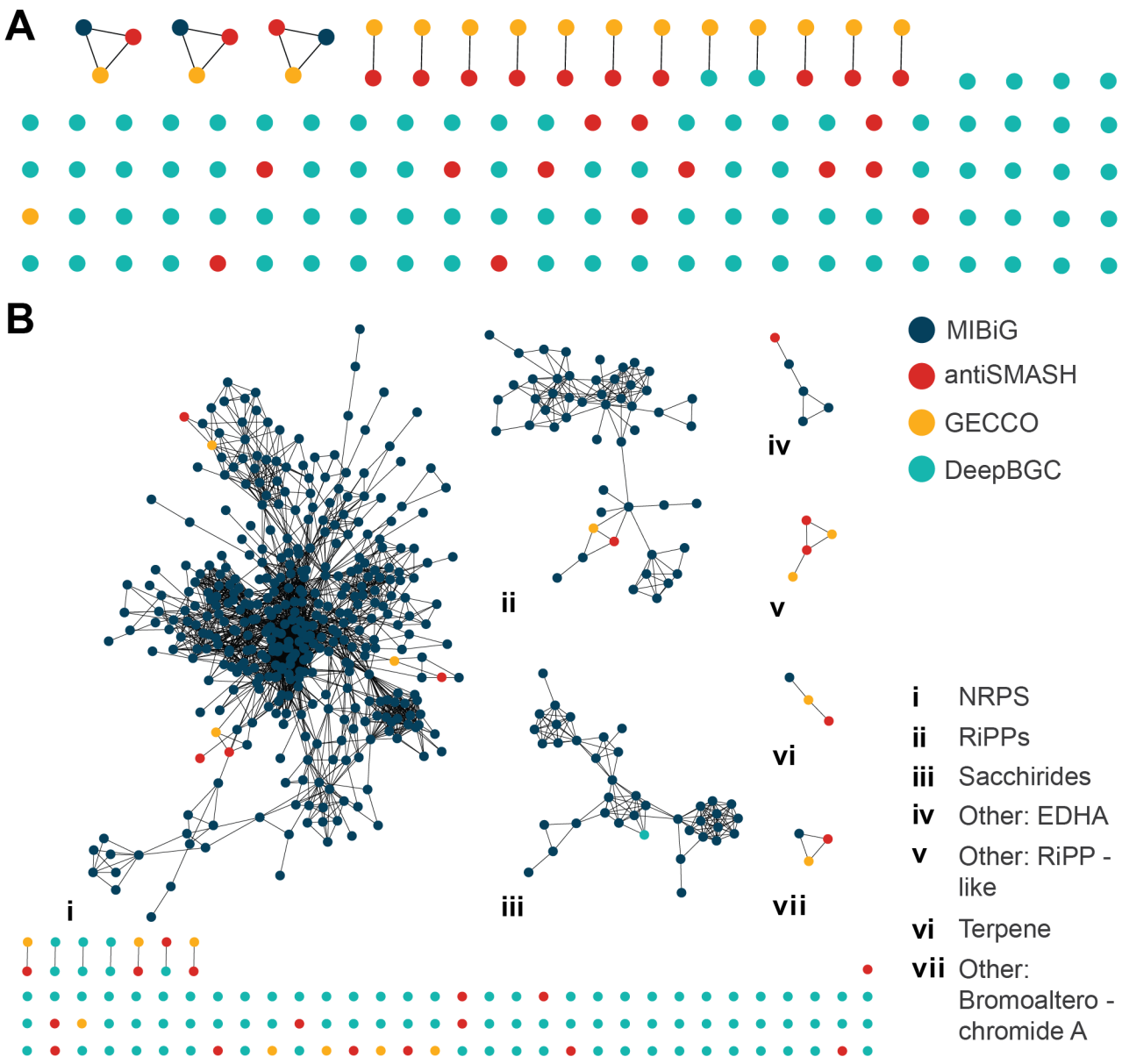


FIG 4 *Conservative* **(A)** and *less-conservative* **(B)** BGC similarity networks. Nodes represent individual BGCs, colored by the computational tool used for prediction (see bottom right legend). Edges indicate gene similarity ($\geq$ 70% for conservative and $\geq$30% for less-conservative) between BGCs. Clusters are labeled (i-vii) according to their predicted product class following the bottom right legend.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

353 **Tables**

TABLE 1 Human and marine pathogen test strains used in antagonistic assays, their antibiotic resistance profiles, and associated diseases.

Test Pathogen	Origin	Resistance	Associated Disease
Bacillus subtilis	Human	Sensitive	Food spoilage, opportunistic infections
Escherichia coli	Human	Sensitive	Urinary tract infections, gastroenteritis, septicemia, meningitis
Serratia marcescens PDL100	Marine	Sensitive	Urinary tract infections, respiratory infections, septicemia, wound infections
Staphylococcus aureus	Human	Sensitive	Skin infections, food poisoning, pneumonia, septicemia, toxic shock syndrome, endocarditis
Staphylococcus aureus	Human	Methicillin	Skin infections, food poisoning, pneumonia, septicemia, toxic shock syndrome, endocarditis
Enterococcus faecalis	Human	Methicillin	Urinary tract infections, endocarditis, bacteremia, wound infections, intra-abdominal infections
Enterococcus faecalis	Human	Vancomycin	Urinary tract infections, endocarditis, bacteremia, wound infections, intra-abdominal infections
Vibrio vulnificus	Marine & Human	Sensitive	Severe wound infections, sepsis, foodborne illness, opportunistic marine pathogen
Vibrio shiloi	Marine	Sensitive	Causes coral bleaching in *Oculina patagonica* in the Mediterranean
Vibrio parahaemolyticus	Marine & Human	Sensitive	Foodborne illness, opportunistic marine pathogen

354

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

TABLE 2 Results of antagonistic assays displaying the inhibitory activity of isolates against various pathogens (in millimeters). An asterisk (*) indicates that the corresponding isolate was selected for sequencing. MRSA: Methicillin Resistant *Staphylococcus aureus*; MSSA: Methicillin Sensitive *Staphylococcus aureus*; BS: *Bacillus subtilis*; VV: *Vibrio vulnificus*; SM: *Serratia marcescens* PDL100; VRE: Vancomycin Resistant *Enterococcus faecalis*; EC: *Escherichia coli*.

Isolate	MRSA	MSSA	BS	VV	VRE	SM	EC
CCAF2	1	0	0	2	0	0	0
CCAF3	0	0	0	2	0	1	0
CCAF4	0	0	0	2	0	1	0
CCAF7	0	0	0	1	0	0	0
CCAF8	0	0	0	2	0	0	0
CCAG11	0	0	0	1	0	0	0
CCAH3	0	0	0	2	0	1	0
CCAH4*	0	1	0	1	0	1	0
CCAH6	0	0	0	2	0	0	0
CCAH7*	0	0	0	2	2	1	0
CCAH8	0	0	0	2	0	0	0
CCAH9	0	0	0	1	0	0	0
CCAH10	0	0	0	1	0	0	0
CCAH11*	7	7	2	2	1	0	1

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

TABLE 3 Summary of genome statistics for the sequenced bacterial isolates.

Metric	*P. denitrificans*	*P. elyakovii*	*R. marisflavi*
Strain ID	CCAH7	CCAH11	CCAH4
Estimated genome size (bp)	5,573,528	5,498,989	4,887,352
Total assembly length (bp)	5,892,140	5,691,751	4,416,806
Number of contigs	3	4	3
[Chromosomal/plasmid]	[1/2]	[3/1]	[2/3]
Number of circularized contigs	3	1	0
Contig N50 length (bp)	4,904,000	3,527,000	3,160,000
G+C content (%)	52.42	43.25	46.61
Number of genes	5331	4908	4435
Gene density (%)	86.80	88.10	86.30
Completeness (%)	100.00	100.00	98.68
Number of rRNAs	21	28	27
Number of tRNAs	77	106	100

TABLE 4 BGC hits with similarity to MiBiG BGCs and their respective products.

Isolate	MiBIG BGC Hits	Gene Similarity (%)	Product
P. denitrificans	BGC0002123	100	Pseudovibriamide A1-6, B1-6
P. elyakovii	BGC0002491	100	Myxochelin A, B, pseudochelin A
P. elyakovii	BGC0002345	66	Hydrogen cyanide
P. elyakovii	BGC0000314	100	Bromoalterochromide A
R. marisflavi	BGC0000645	50	Carotenoid

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

Acknowledgments

This project was supported by an NSF IOS grant 2227068. DLL thanks his Linda Letawa Schobert Enrichment Scholarship and Eberly College of Science Anita M. Collins Research Scholarship. RAGP was supported by an Eberly Postdoctoral Research Fellowship awarded by the Eberly College of Science at Penn State University.

Data availability

Raw sequencing data is available under the SRA accession xxxx and genome assembly and annotation under RefSeq accession xxxx.

Author contribution

MM and KR initially conceptualized the project. DLL and RAGP designed the analyses to be performed. DLL and JSRT performed the sequencing run. KR, MWH, and EKJ conducted the antagonistic assays. DLL and RAGP wrote the manuscript. All authors reviewed, commented on, and contributed to the final version of the manuscript.

Correspondence

The first author can be reached at dfl5440@psu.edu. The last author can be reached at rag5851@psu.edu.

References

[1] . 2020. World Health Organization https://www.who.int/publications/i/item/9789240000650.

[2] **O'Neill J**. may 2016. Tackling Drug-Resistant Infections Globally: Final Report and Recommendations. Government of the United Kingdom .

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

[3] **Bush K, Courvalin P, Dantas G, Davies J, Eisenstein B, Huovinen P, Jacoby GA, Kishony R, Kreiswirth BN, Kutter E, Lerner SA, Levy S, Lewis K, Lomovskaya O, Miller JH, Mobashery S, Piddock LJV, Projan S, Thomas CM, Tomasz A, Tulkens PM, Walsh TR, Watson JD, Witkowski J, Witte W, Wright G, Yeh P, Zgurskaya HI**. dec 2011. Tackling antibiotic resistance. *Nature Reviews Microbiology* 9 (12):894–896. doi: 10.1038/nrmicro2693.

[4] **Singh SB, Pelaez F**. 2008. *In* Petersen F, Amstutz R, (ed), Biodiversity, chemical diversity and drug discovery, vol 65 of Progress in Drug Research Birkhäuser Basel, Basel, p 141–174. doi:10.1007/978-3-7643-8117-2_4.

[5] **Levine DP**. Jan 2006. Vancomycin: A History. *Clinical Infectious Diseases* 42 (Supplement$_1$) : $S5\breve{}S12. doi$: 10.1086/491709.

[6] **Bokesch HR, O'Keefe BR, McKee TC, Pannell LK, Patterson GML, Gardella RS, Sowder RC, Turpin J, Watson K, Buckheit RW, Boyd MR**. Mar 2003. A Potent Novel Anti-HIV Protein from the Cultured Cyanobacterium Scytonema varium. *Biochemistry* 42 (9):2578–2584. doi:10.1021/bi0205698.

[7] **Kwon HC, Kauffman CA, Jensen PR, Fenical W**. Feb 2006. Marinomycins AD, Antitumor-Antibiotics of a New Structure Class from a Marine Actinomycete of the Recently Discovered Genus " Marinispora ". *Journal of the American Chemical Society* 128 (5):1622–1632. doi:10.1021/ja0558948.

[8] **Galand PE, Ruscheweyh HJ, Salazar G, Hochart C, Henry N, Hume BCC, Oliveira PH, Perdereau A, Labadie K, Belser C, Boissin E, Romac S, Poulain J, Bourdin G, Iwankow G, Moulin C, Armstrong EJ, Paz-García DA, Ziegler M, Agostini S, Banaigs B, Boss E, Bowler C, De Vargas C, Douville E, Flores M, Forcioli D, Furla P, Gilson E, Lombard F, Pesant S, Reynaud S, Thomas OP, Troublé R, Zoccola D, Voolstra CR, Thurber RV, Sunagawa S, Wincker P, Allemand D, Planes S**. Jun 2023. Diversity of the Pacific Ocean coral reef microbiome. *Nature Communications* 14 (1):3039. doi: 10.1038/s41467-023-38500-x.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

[9] **Tout J, Jeffries TC, Webster NS, Stocker R, Ralph PJ, Seymour JR**. Apr 2014. Variability in Microbial Community Composition and Function Between Different Niches Within a Coral Reef. *Microbial Ecology* 67 (3):540–552. doi:10.1007/s00248-013-0362-5.

[10] **Ritchie K**. Sep 2006. Regulation of microbial populations by coral surface mucus and mucus-associated bacteria. *Marine Ecology Progress Series* 322:1–14. doi: 10.3354/meps322001.

[11] **Raina JB, Tapiolas D, Motti CA, Foret S, Seemann T, Tebben J, Willis BL, Bourne DG**. Aug 2016. Isolation of an antimicrobial compound produced by bacteria associated with reef-building corals. *PeerJ* 4:e2275. doi:10.7717/peerj.2275.

[12] **Hirata Y, Uemura D**. Jan 1986. Halichondrins - antitumor polyether macrolides from a marine sponge. *Pure and Applied Chemistry* 58 (5):701–710. doi:10.1351/pac198658050701.

[13] **Long BH WACLCAJPLTFWFC Carboni JM**. 1998. Eleutherobin, a novel cytotoxic agent that induces tubulin polymerization, is similar to paclitaxel (Taxol). *Cancer Research* .

[14] **Abdulmohsin A Al-Sofyani AAAS**. Mar 2017. Antibiofilm and antioxidant activities of extracts of crustose coralline algaLithophyllum sp. from the central Red Sea, Saudi Arabia. *journal of king abdulaziz university marine science* 26 (2):35–43. doi:10.4197/Mar.26-2.4.

[15] **Quéré G, Intertaglia L, Payri C, Galand PE**. Aug 2019. Disease Specific Bacterial Communities in a Coralline Algae of the Northwestern Mediterranean Sea: A Combined Culture Dependent and -Independent Approach. *Frontiers in Microbiology* 10:1850. doi: 10.3389/fmicb.2019.01850.

[16] **Yang F, Xiao Z, Wei Z, Long L**. Jun 2021. Bacterial Communities Associated With Healthy and Bleached Crustose Coralline Alga Porolithon onkodes. *Frontiers in Microbiology* 12:646143. doi:10.3389/fmicb.2021.646143.

[17] **Sneed JM, Ritson-Williams R, Paul VJ**. Nov 2015. Crustose coralline algal species host distinct bacterial assemblages on their surfaces. *The ISME Journal* 9 (11):2527–2536. doi: 10.1038/ismej.2015.67.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

[18] **Skerratt J, Bowman J, Hallegraeff G, James S, Nichols P**. 2002. Algicidal bacteria associated with blooms of a toxic dinoflagellate in a temperate Australian estuary. *Marine Ecology Progress Series* 244:1–15. doi:10.3354/meps244001.

[19] **Nithyanand P, Manju S, Karutha Pandian S**. Jan 2011. Phylogenetic characterization of culturable actinomycetes associated with the mucus of the coral Acropora digitifera from Gulf of Mannar: Bioactive potential of coral-associated actinomycetes. *FEMS Microbiology Letters* 314 (2):112–118. doi:10.1111/j.1574-6968.2010.02149.x.

[20] **Negri A, Webster N, Hill R, Heyward A**. 2001. Metamorphosis of broadcast spawning corals in response to bacteria isolated from crustose algae. *Marine Ecology Progress Series* 223:121–131. doi:10.3354/meps223121.

[21] **Sneed JM, Sharp KH, Ritchie KB, Paul VJ**. Jul 2014. The chemical cue tetrabromopyrrole from a biofilm bacterium induces settlement of multiple Caribbean corals. *Proceedings of the Royal Society B: Biological Sciences* 281 (1786):20133086. doi:10.1098/rspb.2013.3086.

[22] **Gomez-Lemos LA, Diaz-Pulido G**. Jun 2017. Crustose coralline algae and associated microbial biofilms deter seaweed settlement on coral reefs. *Coral Reefs* 36 (2):453–462. doi:10.1007/s00338-017-1549-x.

[23] **Cimermancic P, Medema M, Claesen J, Kurita K, WielandBrown L, Mavrommatis K, Pati A, Godfrey P, Koehrsen M, Clardy J, Birren B, Takano E, Sali A, Linington R, Fischbach M**. Jul 2014. Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. *Cell* 158 (2):412–421. doi:10.1016/j.cell.2014.06.034.

[24] **Dilshad R, Jamil N, Batool R**. Jan 2022. Biosynthetic Gene Clusters in Bacteria: A Review: Bacterial Secondary metabolites producing genes. *Proceedings of the Pakistan Academy of Sciences: B. Life and Environmental Sciences* 58 (3):29–42. doi:10.53560/PPASB(58-3)665.

[25] . 2023. Oxford Nanopore Technologies .

[26] **Wingett SW, Andrews S**. Sep 2018. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research* 7:1338. doi:10.12688/f1000research.15931.2.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

[27] **De Coster W, Rademakers R**. May 2023. NanoPack2: population-scale evaluation of long-read sequencing data. *Bioinformatics* 39 (5):btad311. doi: 10.1093/bioinformatics/btad311.

[28] **Marçais G, Kingsford C**. Mar 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k -mers. *Bioinformatics* 27 (6):764–770. doi: 10.1093/bioinformatics/btr011.

[29] **Ranallo-Benavidez TR, Jaron KS, Schatz MC**. Mar 2020. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nature Communications* 11 (1):1432. doi:10.1038/s41467-020-14998-3.

[30] **Kolmogorov M, Yuan J, Lin Y, Pevzner PA**. May 2019. Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology* 37 (5):540–546. doi: 10.1038/s41587-019-0072-8.

[31] **Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR**. Dec 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biology* 16 (1):294. doi:10.1186/s13059-015-0849-0.

[32] **Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW**. Jul 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research* 25 (7):1043–1055. doi:10.1101/gr.186072.114.

[33] **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ**. Oct 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215 (3):403–410. doi: 10.1016/S0022-2836(05)80360-2.

[34] **Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH**. Nov 2022. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 38 (23):5315–5316. doi:10.1093/bioinformatics/btac672.

[35] **Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil PA, Hugenholtz P**. Jan 2022. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Research* 50 (D1):D785–D794. doi:10.1093/nar/gkab776.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

[36] **Seemann T**. Jul 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30 (14):2068–2069. doi:10.1093/bioinformatics/btu153.

[37] **Kanehisa M, Sato Y, Morishima K**. Feb 2016. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *Journal of Molecular Biology* 428 (4):726–731. doi:10.1016/j.jmb.2015.11.006.

[38] **Kanehisa M, Sato Y, Kawashima M**. Jan 2022. KEGG mapping tools for uncovering hidden features in biological data. *Protein Science* 31 (1):47–53. doi:10.1002/pro.4172.

[39] **Tatusov RL**. Jan 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research* 28 (1):33–36. doi:10.1093/nar/28.1.33.

[40] **Rousset F, Cabezas-Caballero J, Piastra-Facon F, Fernández-Rodríguez J, Clermont O, Denamur E, Rocha EPC, Bikard D**. Jan 2021. The impact of genetic diversity on gene essentiality within the Escherichia coli species. *Nature Microbiology* 6 (3):301–312. doi:10.1038/s41564-020-00839-y.

[41] **Alker AT, Delherbe N, Purdy TN, Moore BS, Shikuma NJ**. Nov 2020. Genetic examination of the marine bacterium Pseudoalteromonas luteoviolacea and effects of its metamorphosis-inducing factors. *Environmental Microbiology* 22 (11):4689–4701. doi:10.1111/1462-2920.15211.

[42] **Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, Fetter A, Terlouw BR, Metcalf WW, Helfrich EJN, van Wezel GP, Medema MH, Weber T**. Jul 2023. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Research* 51 (W1):W46–W50. doi:10.1093/nar/gkad344.

[43] **Hannigan GD, Prihoda D, Palicka A, Soukup J, Klempir O, Rampula L, Durcak J, Wurst M, Kotowski J, Chang D, Wang R, Piizzi G, Temesi G, Hazuda DJ, Woelk CH, Bitton DA**. Oct 2019. A deep learning genome-mining strategy for biosynthetic gene cluster prediction. *Nucleic Acids Research* 47 (18):e110–e110. doi:10.1093/nar/gkz654.

[44] **Carroll LM, Larralde M, Fleck JS, Ponnudurai R, Milanese A, Cappio E, Zeller G**. May 2021. Accurate de novo identification of biosynthetic gene clusters with GECCO. *bioRxiv* doi:10.1101/2021.05.03.442509.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

[45] **Yu Y, Ouyang Y, Yao W**. Apr 2018. shinyCircos: an R/Shiny application for interactive creation of Circos plot. *Bioinformatics* 34 (7):1229–1231. doi: 10.1093/bioinformatics/btx763.

[46] **Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, Parkinson EI, De Los Santos ELC, Yeong M, Cruz-Morales P, Abubucker S, Roeters A, Lokhorst W, Fernandez-Guerra A, Cappelini LTD, Goering AW, Thomson RJ, Metcalf WW, Kelleher NL, Barona-Gomez F, Medema MH**. Jan 2020. A computational framework to explore large-scale biosynthetic diversity. *Nature Chemical Biology* 16 (1):60–68. doi:10.1038/s41589-019-0400-9.

[47] **Kautsar SA, Blin K, Shaw S, Navarro-Muñoz JC, Terlouw BR, van der Hooft JJJ, van Santen JA, Tracanna V, Suarez Duran HG, Pascal Andreu V, Selem-Mojica N, Alanjary M, Robinson SL, Lund G, Epstein SC, Sisto AC, Charkoudian LK, Collemare J, Linington RG, Weber T, Medema MH**. Oct 2019. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Research* p gkz882. doi: 10.1093/nar/gkz882.

[48] **Su G, Morris JH, Demchak B, Bader GD**. Sep 2014. Biological Network Exploration with Cytoscape 3. *Current Protocols in Bioinformatics* 47 (1). doi:10.1002/0471250953.bi0813s47.

[49] **Petersen LE, Moeller M, Versluis D, Nietzer S, Kellermann MY, Schupp PJ**. Apr 2021. Mono- and multispecies biofilms from a crustose coralline alga induce settlement in the scleractinian coral Leptastrea purpurea. *Coral Reefs* 40 (2):381–394. doi: 10.1007/s00338-021-02062-5.

[50] **Turnlund AC**. Feb. 2024. Interactions between coral larval settlement, marine biofilms and crustose coralline algae microbiomes. Phd thesis. The University of Queensland. doi: 10.14264/30e05cf.

[51] **Wang KL, Xu Y, Lu L, Li Y, Han Z, Zhang J, Shao CL, Wang CY, Qian PY**. Oct 2015. Low-Toxicity Diindol-3-ylmethanes as Potent Antifouling Compounds. *Marine Biotechnology* 17 (5):624–632. doi:10.1007/s10126-015-9656-6.

[52] **Tebben J, Tapiolas DM, Motti CA, Abrego D, Negri AP, Blackall LL, Steinberg PD, Harder T**. Apr 2011. Induction of Larval Metamorphosis of the Coral Acropora millepora

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

by Tetrabromopyrrole Isolated from a Pseudoalteromonas Bacterium. *PLoS ONE* 6 (4):e19082. doi:10.1371/journal.pone.0019082.

[53] **Tebben J, Motti CA, Siboni N, Tapiolas DM, Negri AP, Schupp PJ, Kitamura M, Hatta M, Steinberg PD, Harder T**. Jun 2015. Chemical mediation of coral larval settlement by crustose coralline algae. *Scientific Reports* 5 (1):10803. doi:10.1038/srep10803.

[54] **Uemura D, Kita M, Arimoto H, Kitamura M**. May 2009. Recent aspects of chemical ecology: Natural toxins, coral communities, and symbiotic relationships. *Pure and Applied Chemistry* 81 (6):1093–1111. doi:10.1351/PAC-CON-08-08-12.

[55] **Sawabe T, Tanaka R, Iqbal MM, Tajima K, Ezura Y, Ivanova EP, Christen R**. Jan 2000. Assignment of Alteromonas elyakovii KMM 162T and five strains isolated from spot-wounded fronds of Laminaria japonica to Pseudoalteromonas elyakovii comb. nov. and the extended description of the species. *International Journal of Systematic and Evolutionary Microbiology* 50 (1):265–271. doi:10.1099/00207713-50-1-265.

[56] **Kalidasan V, Joseph N, Kumar S, Awang Hamat R, Neela VK**. Nov 2018. Iron and Virulence in Stenotrophomonas Maltophilia: All We Know So Far. *Frontiers in Cellular and Infection Microbiology* 8:401. doi:10.3389/fcimb.2018.00401.

[57] **Sheldon JR, Laakso HA, Heinrichs DE**. Mar 2016. Iron Acquisition Strategies of Bacterial Pathogens. *Microbiology Spectrum* 4 (2):4.2.05. doi: 10.1128/microbiolspec.VMBF-0010-2015.

[58] **Bai S, Huang Z, Li XG**. Jun 2024. Genome analysis of Rossellomorea sp. y25, a deep sea bacterium isolated from the sediments of South China Sea. *Marine Genomics* 75:101110. doi:10.1016/j.margen.2024.101110.

[59] **Barrat JA, Chauvaud L, Amice E, Grall J, Rouget ML, Bayon G, Germain Y**. May 2024. Trace elements in coralline algae as a new proxy for seawater chemistry and metal pollution. *Chemical Geology* 652:122026. doi:10.1016/j.chemgeo.2024.122026.

[60] **Caballero-Gallardo K, Alcala-Orozco M, Barraza-Quiroz D, De La Rosa J, Olivero-Verbel J**. Mar 2020. Environmental risks associated with trace elements in sediments from Cartagena Bay, an industrialized site at the Caribbean. *Chemosphere* 242:125173. doi:10.1016/j.chemosphere.2019.125173.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

[61] **Jaramillo-Colorado BE, Aga DS, Noguera-Oviedo K**. Oct 2016. Heavy metal contamination of estuarine sediments from Cartagena Bay, Colombia. *Toxicology Letters* 259:S170. doi:10.1016/j.toxlet.2016.07.405.

[62] **Hernando Campos N, Luis Marrugo-Negrete J**. Jun. 2023. *In* Mancuso M, H H Abbas M, Bottari T, A Abdelhafez A, (ed), Mercury in the Colombian Caribbean: The Bay of Cartagena, A Model in Resilience, vol 5 IntechOpen. doi:10.5772/intechopen.107240.

[63] **Duarte-Restrepo E, Noguera-Oviedo K, Butryn D, Wallace JS, Aga DS, Jaramillo-Colorado BE**. Mar 2021. Spatial distribution of pesticides, organochlorine compounds, PBDEs, and metals in surface marine sediments from Cartagena Bay, Colombia. *Environmental Science and Pollution Research* 28 (12):14632–14653. doi: 10.1007/s11356-020-11504-6.

[64] **Elamin O, Chraibi M, Ibnsouda SK, Iraqui MH**. Sep 2018. LYCOPENE PRODUCTION IN MYCOBACTERIUM SMEGMATIS BY EXPRESSION OF CRT GENES FROM MYCOBACTERIUM AURUM AND PROTECTIVE EFFECT OF LYCOPENE IN VIVO AND IN VITRO AGAINST UV RADIATION. *International Journal of Pharmacy and Pharmaceutical Sciences* 10 (9):49. doi:10.22159/ijpps.2018v10i9.25768.

[65] **Mori M, Ponce-de León M, Peretó J, Montero F**. Oct 2016. Metabolic Complementation in Bacterial Communities: Necessary Conditions and Optimality. *Frontiers in Microbiology* 7. doi:10.3389/fmicb.2016.01553.

[66] **Wei B, Hu GA, Zhou ZY, Yu WC, Du AQ, Yang CL, Yu YL, Chen JW, Zhang HW, Wu Q, Xuan Q, Xu XW, Wang H**. Jun 2023. Global analysis of the biosynthetic chemical space of marine prokaryotes. *Microbiome* 11 (1):144. doi:10.1186/s40168-023-01573-3.

[67] **Wei B, Du A, Zhou Z, Lai C, Yu W, Yu J, Yu Y, Chen J, Zhang H, Xu X, Wang H**. Nov 2021. An atlas of bacterial secondary metabolite biosynthesis gene clusters. *Environmental Microbiology* 23 (11):6981–6992. doi:10.1111/1462-2920.15761.

[68] **Ióca LP, Dai Y, Kunakom S, Diaz-Espinosa J, Krunic A, Crnkovic CM, Orjala J, Sanchez LM, Ferreira AG, Berlinck RGS, Eustáquio AS**. Jul 2021. A Family of Nonribosomal Peptides Modulate Collective Behavior in Pseudovibrio Bacteria Isolated

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.26.625565; this version posted November 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

from Marine Sponges**. *Angewandte Chemie* 133 (29):16027–16034. doi: 10.1002/ange.202017320.

[69] **Korp J, Winand L, Sester A, Nett M**. Nov 2018. Engineering Pseudochelin Production in Myxococcus xanthus. *Applied and Environmental Microbiology* 84 (22):e01789–18. doi: 10.1128/AEM.01789-18.

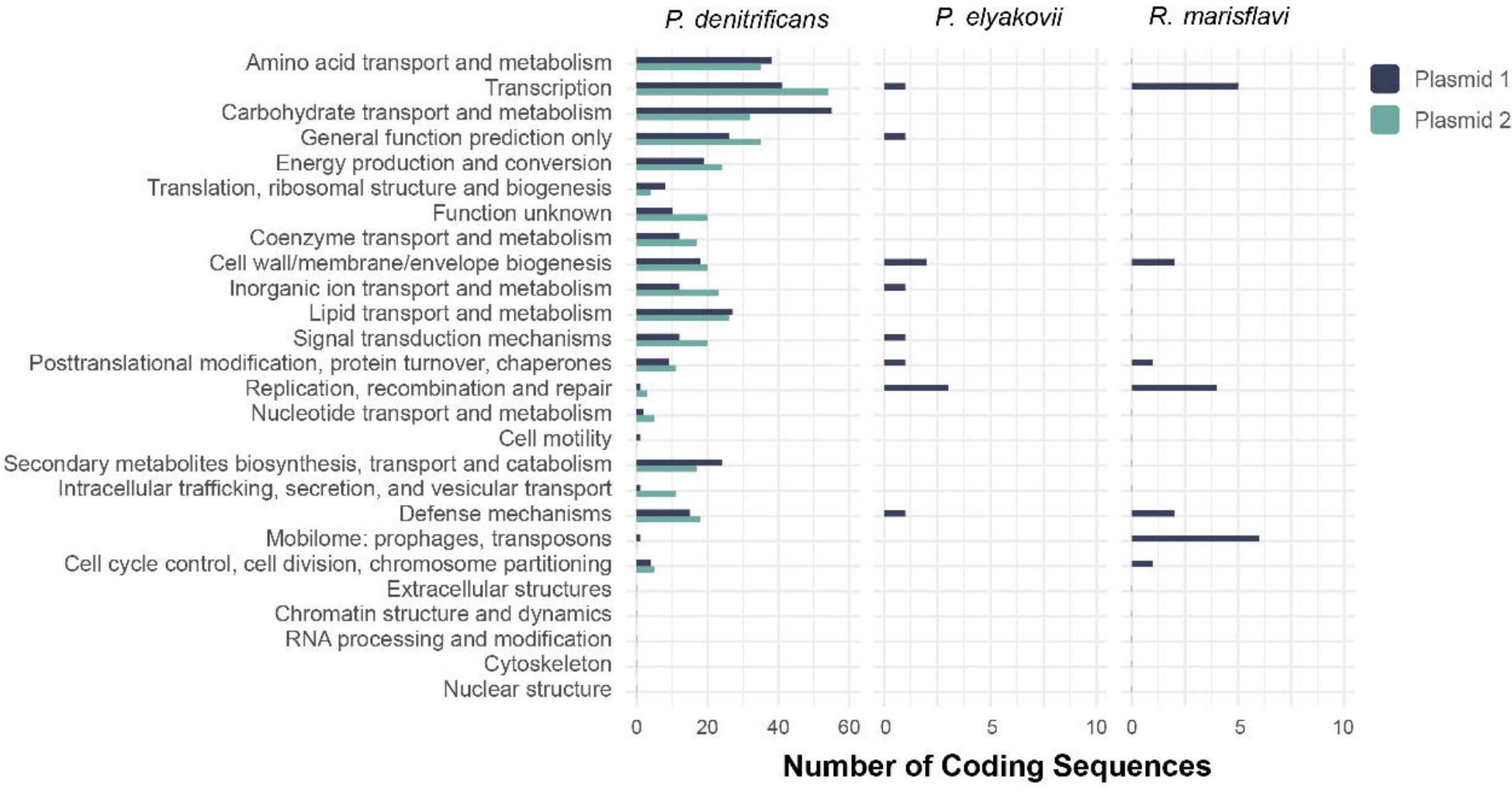
[70] **Kunze B, Bedorf N, Kohl W, Höfle G, Reichenbach H**. 1989. Myxochelin A, a new iron-chelating compound from Angiococcus disciformis (myxobacterales). Production, isolation, physico-chemical and biological properties. *The Journal of Antibiotics* 42 (1):14–17. doi:10.7164/antibiotics.42.14.

[71] **Vynne NG, Månsson M, Nielsen KF, Gram L**. Dec 2011. Bioactivity, Chemical Profiling, and 16S rRNA-Based Phylogeny of Pseudoalteromonas Strains Collected on a Global Research Cruise. *Marine Biotechnology* 13 (6):1062–1073. doi:10.1007/s10126-011-9369-4.

[72] **Speitling M, Smetanina OF, Kuznetsova TA, Laatsch H**. Jan 2007. Bromoalterochromides A and A, Unprecedented Chromopeptides from a Marine Pseudoalteromonas maricaloris Strain KMM 636T†. *The Journal of Antibiotics* 60 (1):36–42. doi:10.1038/ja.2007.5.

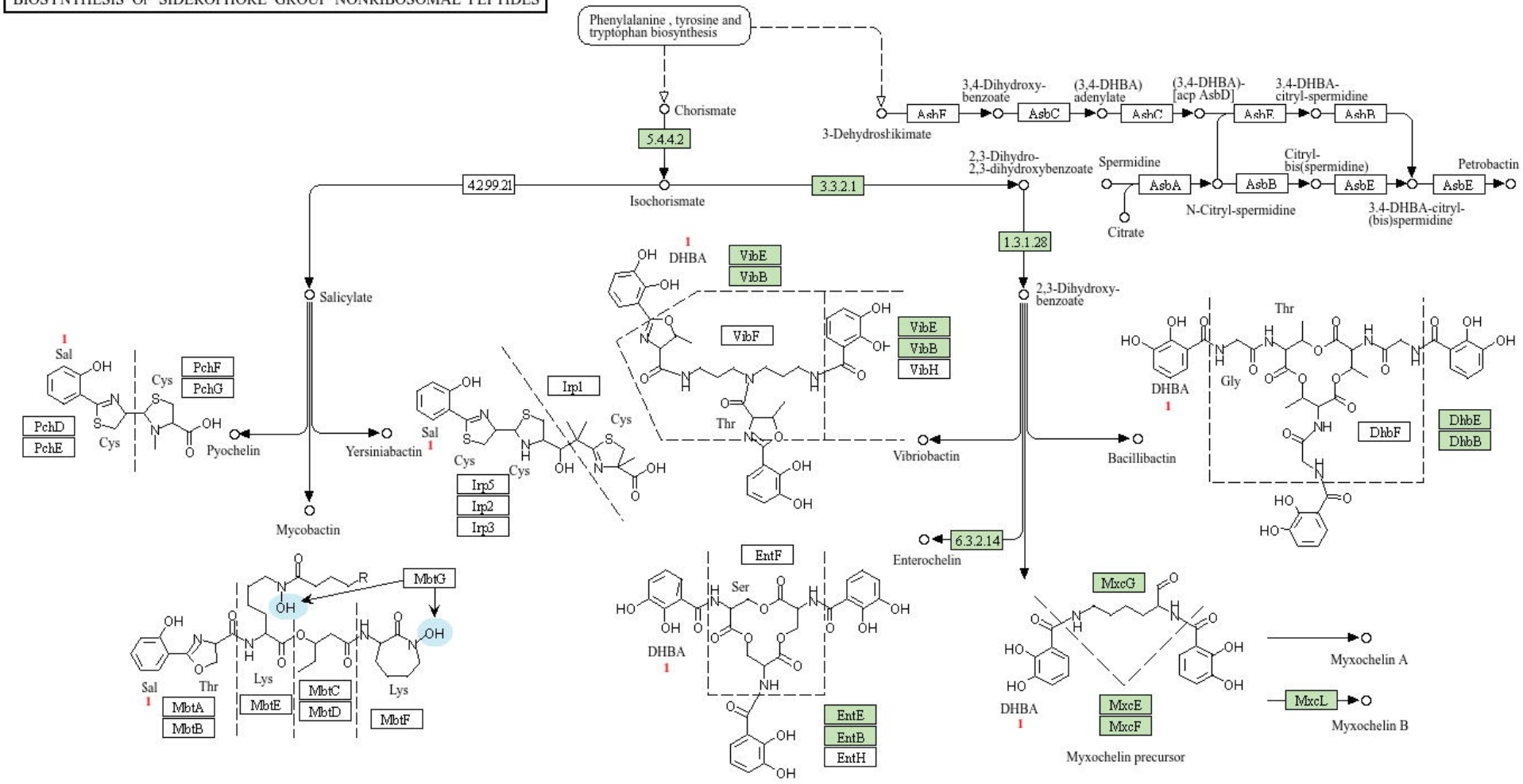
[73] **Lee W, Lee DG**. Sep 2014. Lycopene-induced hydroxyl radical causes oxidative DNA damage in Escherichia coli. *Journal of Microbiology and Biotechnology* 24 (9):1232–1237. doi: 10.4014/jmb.1406.06009.

[74] **Zigangirova NA, Morgunova EY, Fedina ED, Shevyagina NV, Borovaya TG, Zhukhovitsky VG, Kyle NH, Petyaev IM**. 2017. Lycopene Inhibits Propagation of Chlamydia Infection. *Scientifica* 2017:1–11. doi:10.1155/2017/1478625.


P. denitrificans
P. elyakovii
R. marisflavi
Plasmid 1
Plasmid 2
Amino acid transport and metabolism
Transcription
Carbohydrate transport and metabolism
General function prediction only
Energy production and conversion
Translation, ribosomal structure and biogenesis
Function unknown
Coenzyme transport and metabolism
Cell wall/membrane/envelope biogenesis
Inorganic ion transport and metabolism
Lipid transport and metabolism
Signal transduction mechanisms
Posttranslational modification, protein turnover, chaperones
Replication, recombination and repair
Nucleotide transport and metabolism
Cell motility
Secondary metabolites biosynthesis, transport and catabolism
Intracellular trafficking, secretion, and vesicular transport
Defense mechanisms
Mobilome: prophages, transposons
Cell cycle control, cell division, chromosome partitioning
Extracellular structures
Chromatin structure and dynamics
RNA processing and modification
Cytoskeleton
Nuclear structure
0
20
40
60
0
5
10
0
5
10
Number of Coding Sequences

CCAH11

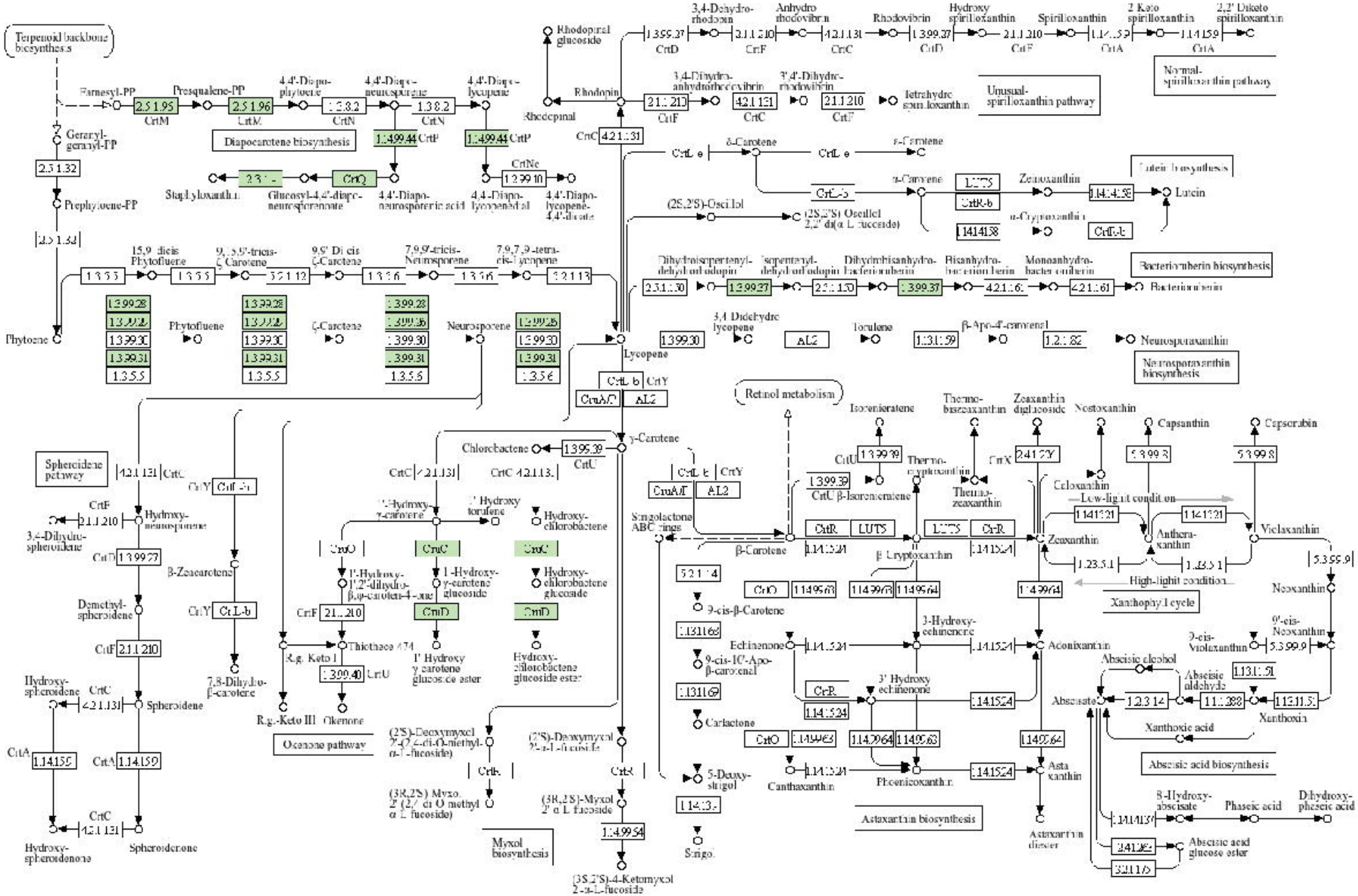
BIOSYNTHESIS OF SIDEROPHORE GROUP NONRIBOSOMAL PEPTIDES



01053 8/12/21
(c) Kanehisa Laboratories

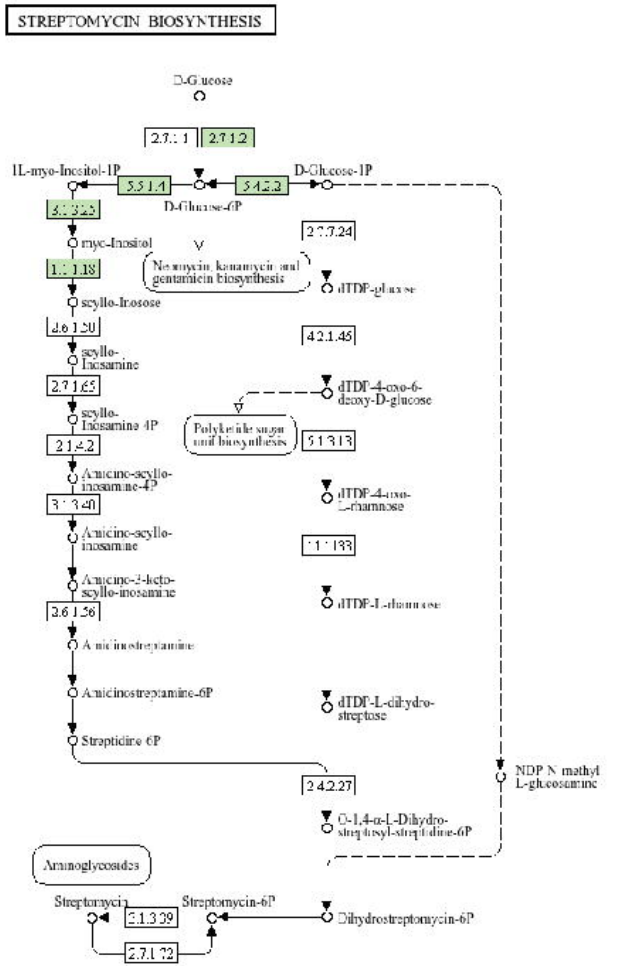
CCAH4

CAROTENOID BIOSYNTHESIS



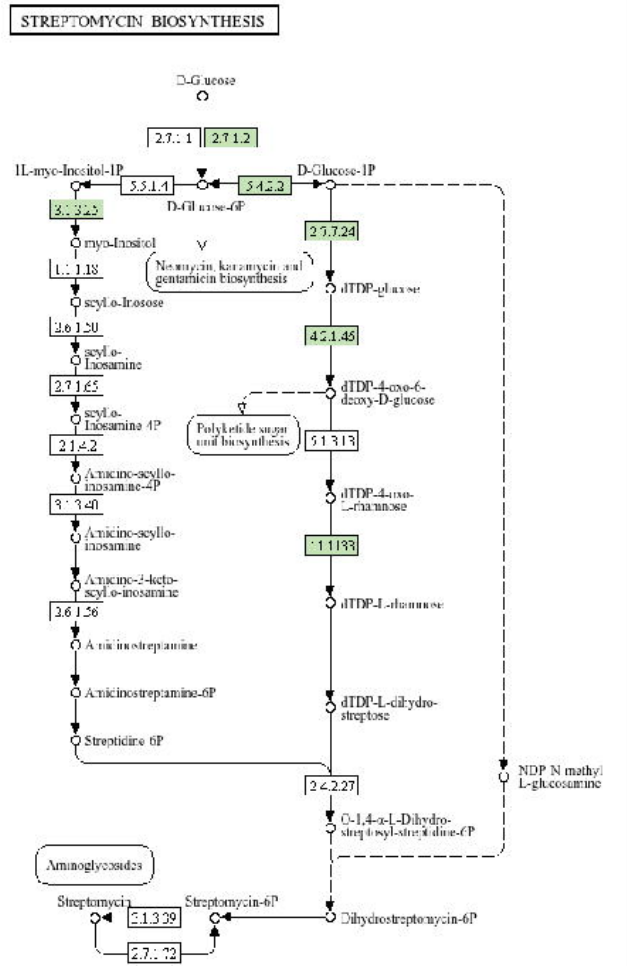
00900 5/26/22
(c) Kanehisa Laboratories

CCAH7



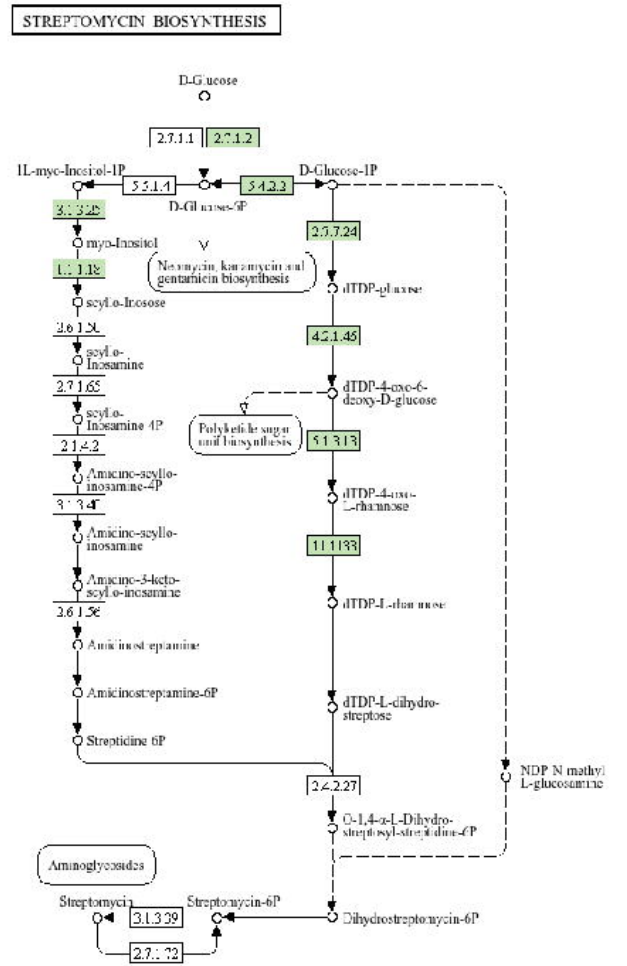
00521 8/10/16
(c) Kanehisa Laboratories

CCAH11

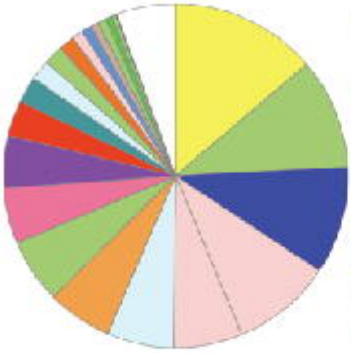


00521 8/10/16
(c) Kanehisa Laboratories

CCAH4

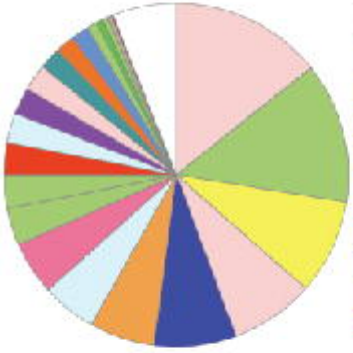


00521 8/10/16
(c) Kanehisa Laboratories



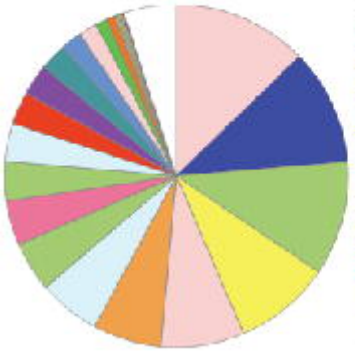
Psuedovibrio denitrificans CCAH7

- Environmental information processing
- Protein families: signaling and cellular processes
- Carbohydrate metabolism
- Protein families: genetic information processing
- Genetic information processing
- Unclassified: metabolism
- Amino acid metabolism
- Cellular processes
- Metabolism of cofactors and vitamins
- Energy metabolism
- Nucleotide metabolism



Psuedoalteromonas elyakovii CCAH11

- Protein families: genetic information processing
- Protein families: signaling and cellular processes
- Environmental information processing
- Genetic information processing
- Carbohydrate metabolism
- Amino acid metabolism
- Unclassified: metabolism
- Metabolism of cofactors and vitamins
- Cellular processes
- Unclassified: signaling and cellular processes
- Nucleotide metabolism

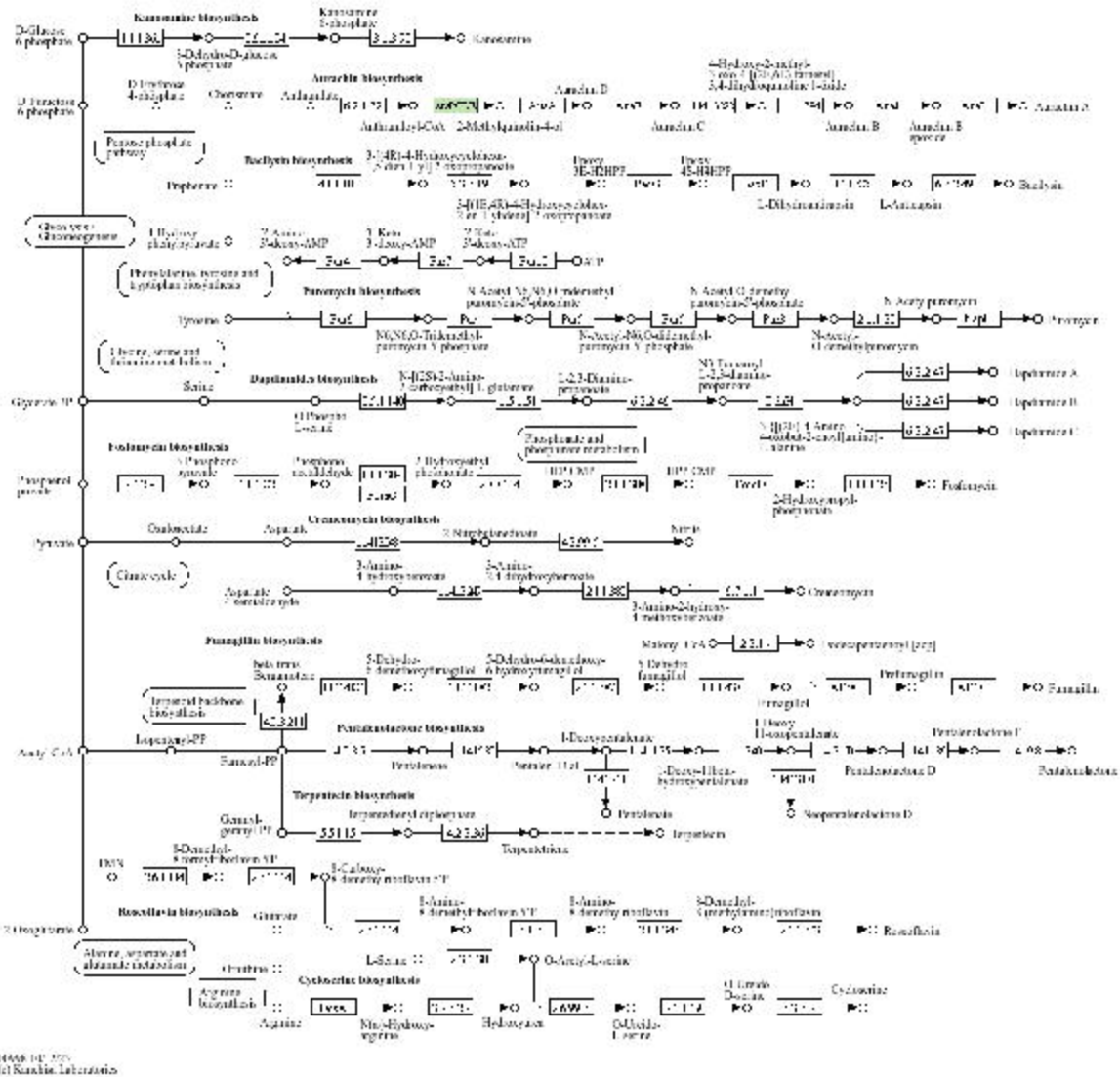


Rosellomorea marisflavi CCAH4

- Protein families: genetic information processing
- Carbohydrate metabolism
- Protein families: signaling and cellular processes
- Environmental information processing
- Genetic information processing
- Amino acid metabolism
- Unclassified: metabolism
- Unclassified: signaling and cellular processes
- Metabolism of cofactors and vitamins
- Cellular processes
- Protein families: metabolism

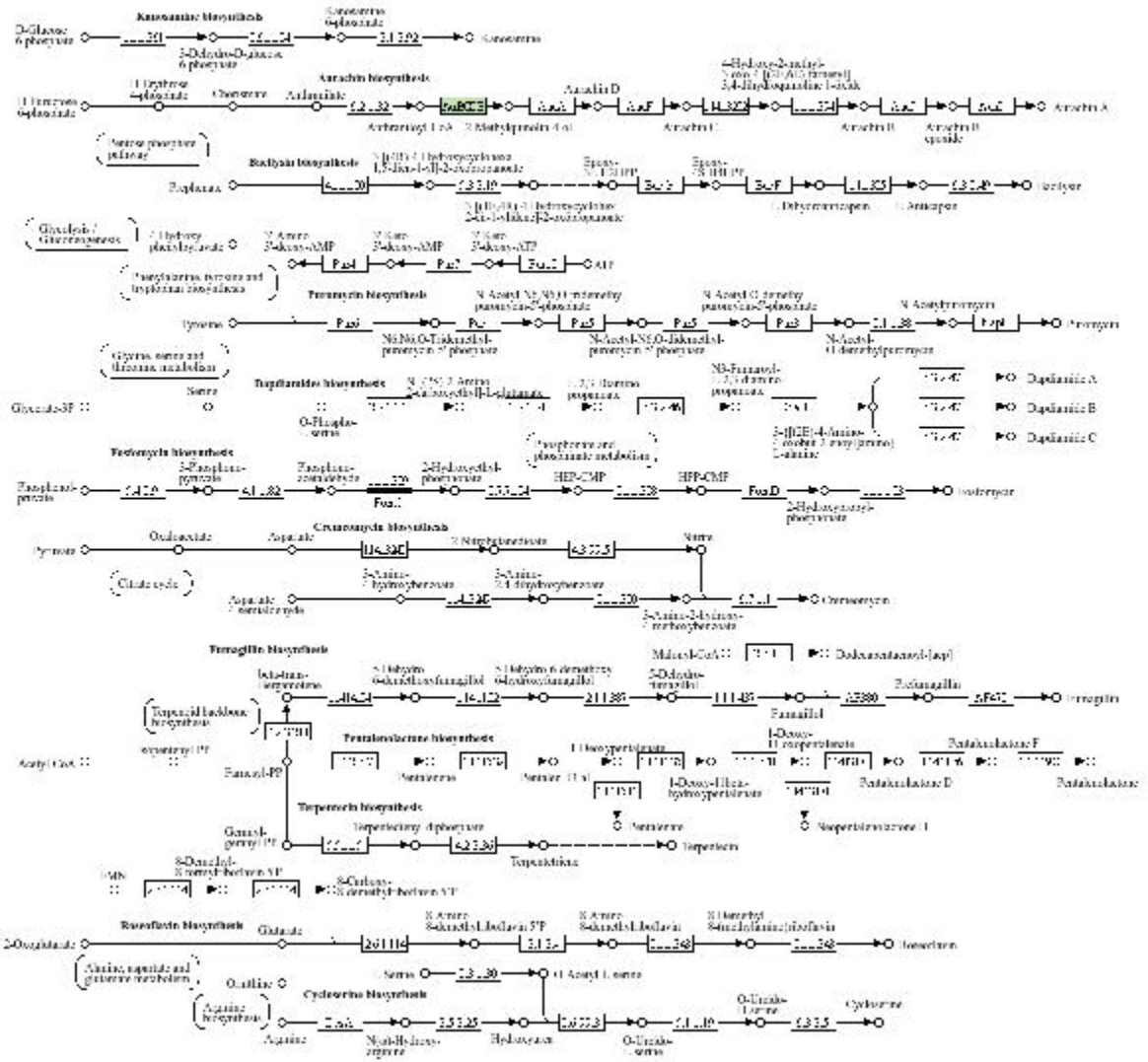
Pseudovibrio denitrificans CCAH7

BIOSYNTHESIS OF VARIOUS ANTIBIOTICS


Kanosamine biosynthesis
Kanosamine 3-phosphate
3-Deoxy-D-glucose 6-phosphate
Kanosamine
Aureothin biosynthesis
D-erythrose 4-phosphate
Chorismate
Anthranilate
Aureothin
Aureothin-CoA
2-Methylquinolin-4-ol
Aureothin D
Aureothin C
Aureothin B
Aureothin A
spacer
Pentose phosphate pathway
Tachykin biosynthesis
Tachykin
Diketide
Triketide
Tetraketide
Pentaketide
Hexaketide
Heptaketide
Octaketide
Nonaketide
Decaketide
Undecaketide
Dodecaketide
Tridecaketide
Tetradecaketide
Pentadecaketide
Hexadecaketide
Heptadecaketide
Octadecaketide
Nonadecaketide
Eicosaketide
Heneicosaketide
Docosaketide
Tricosaketide
Tetracosaketide
Pentacosaketide
Hexacosaketide
Heptacosaketide
Octacosaketide
Nonacosaketide
Triacontaketide
Hentriacontaketide
Dotriacontaketide
Tritriacontaketide
Tetratriacontaketide
Pentatriacontaketide
Hexatriacontaketide
Heptatriacontaketide
Octatriacontaketide
Nonatriacontaketide
Tetracontaketide
Hentetracontaketide
Dotetracontaketide
Tritetracontaketide
Tetratetracontaketide
Pentatetracontaketide
Hexatetracontaketide
Heptatetracontaketide
Octatetracontaketide
Nonatetracontaketide
Pentac

Pseudoalteromonas elyakovii CCAH11

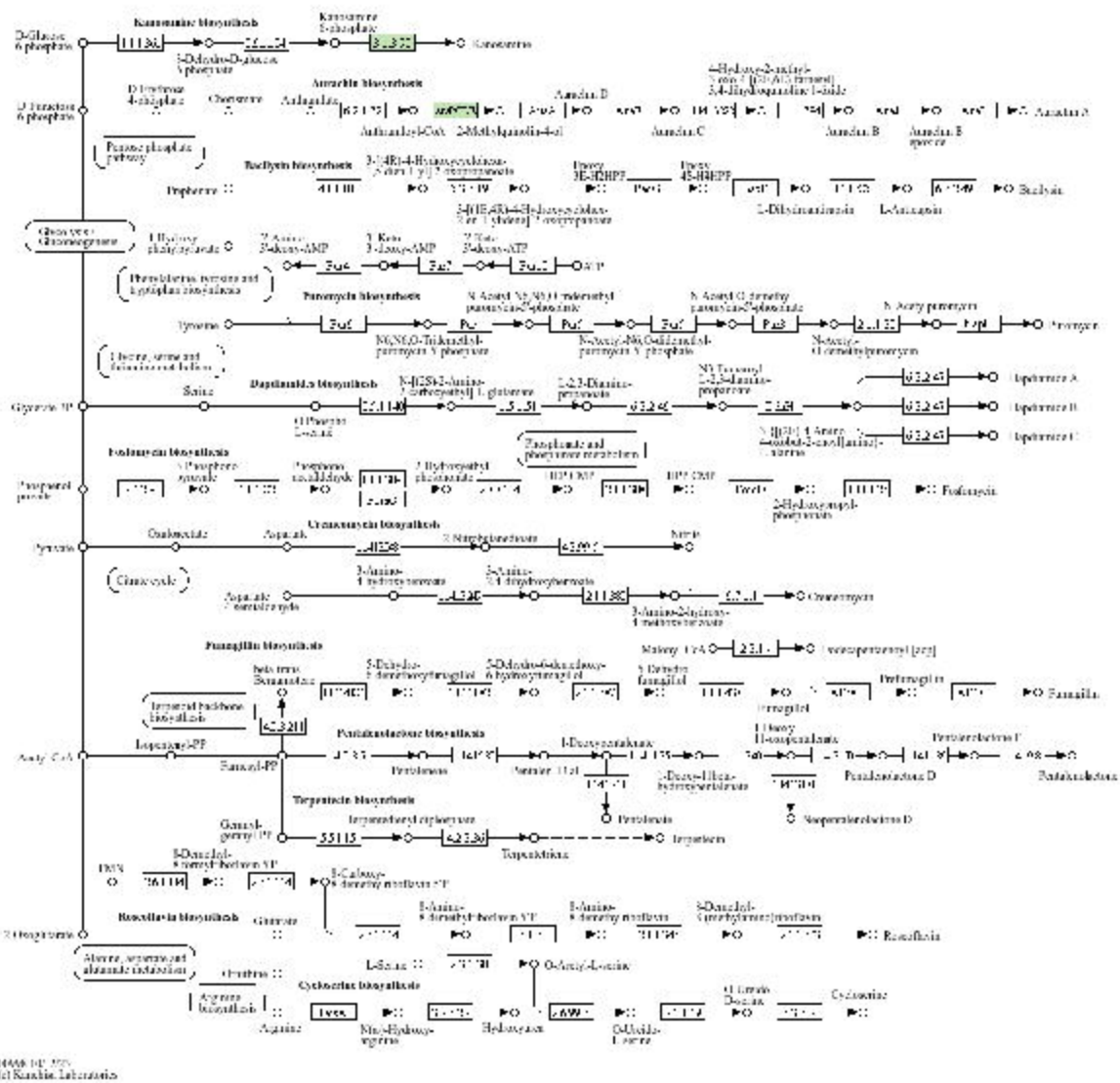
BIOSYNTHESIS OF VARIOUS ANTIBIOTICS



(c) Kanehisa Laboratories

Rosellomorea marisflavi CCAH4

BIOSYNTHESIS OF VARIOUS ANTIBIOTICS



(c) Kanehisa Laboratories