



Article

Plasmid-Borne Biosynthetic Gene Clusters within a Permanently Stratified Marine Water Column

Paraskevi Mara ¹, David Geller-McGrath ², Elizabeth Suter ³, Gordon T. Taylor ⁴, Maria G. Pachiadaki ² and Virginia P. Edgcomb ^{1,*}

¹ Geology & Geophysics Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA; pmara@whoi.edu

² Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA; mcgrath1@mit.edu (D.G.-M.); mpachiadaki@whoi.edu (M.G.P.)

³ Biology, Chemistry and Environmental Science Department, Molloy University, New York, NY 11570, USA; esuter@molloy.edu

⁴ School of Marine, Atmospheric and Sustainability Sciences, Stony Brook University, New York, NY 11794, USA; gordon.taylor@stonybook.edu

* Correspondence: vedgcomb@whoi.edu

Abstract: Plasmids are mobile genetic elements known to carry secondary metabolic genes that affect the fitness and survival of microbes in the environment. Well-studied cases of plasmid-encoded secondary metabolic genes in marine habitats include toxin/antitoxin and antibiotic biosynthesis/resistance genes. Here, we examine metagenome-assembled genomes (MAGs) from the permanently-stratified water column of the Cariaco Basin for integrated plasmids that encode biosynthetic gene clusters of secondary metabolites (smBGCs). We identify 16 plasmid-borne smBGCs in MAGs associated primarily with Planctomycetota and Pseudomonadota that encode terpene-synthesizing genes, and genes for production of ribosomal and non-ribosomal peptides. These identified genes encode for secondary metabolites that are mainly antimicrobial agents, and hence, their uptake via plasmids may increase the competitive advantage of those host taxa that acquire them. The ecological and evolutionary significance of smBGCs carried by prokaryotes in oxygen-depleted water columns is yet to be fully elucidated.



Citation: Mara, P.; Geller-McGrath, D.; Suter, E.; Taylor, G.T.; Pachiadaki, M.G.; Edgcomb, V.P. Plasmid-Borne Biosynthetic Gene Clusters within a Permanently Stratified Marine Water Column. *Microorganisms* **2024**, *12*, 929. <https://doi.org/10.3390/microorganisms12050929>

Academic Editors: Joana Barcelos e Ramos and Susana Isabel Chaves Ribeiro

Received: 24 March 2024

Revised: 22 April 2024

Accepted: 29 April 2024

Published: 2 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Genes can be transferred between bacteria by three main mechanisms: transformation, where extracellular DNA is taken up by a recipient cell; transduction, where genetic material is transferred via bacteriophages; and conjugation, where genetic material is transferred between cells via a conjugal pilus. Plasmids are mobile elements that carry genes typically related to antibiotics or other secondary metabolites and they are exchanged between microorganisms during conjugation or can be acquired via transformation. Using marine sediments and mesocosm experiments, transfer of mobile DNA through plasmids was demonstrated using a naturally occurring plasmid and a mixed marine bacterial community [1,2]. While studies of terrestrial environments have revealed the importance of secondary metabolite biosynthetic gene clusters (smBGCs), we are just starting to learn about the roles they may play within marine prokaryotic communities, particularly in oxygen-depleted water columns (ODWCs: $O_2 \leq 20 \mu\text{mol kg}^{-1}$). As an example, siderophores that facilitate iron uptake and occur as smBGCs, were found encoded in plasmids of *Rhizobium* sp. that inhabit the rhizosphere of terrestrial plants [3].

Paoli et al. (2022) [4] investigated the biosynthetic potential of microorganisms in the open ocean by combining analyses of ~10,000 microbial genomes from cultured isolates

and single cell amplification with ~25,000 new draft genomes from >1000 seawater samples from diverse global epipelagic, mesopelagic, and bathypelagic zones. This analysis revealed ~40,000 putative and mostly new biosynthetic gene clusters, and highlighted that certain uncultured bacterial lineages have high abundances of smBGCs (e.g., '*Candidatus Eudoremicrobiaceae*'). In addition, metagenome surveys of cultured environmental bacterial isolates from marine sediments and the surface ocean showed that bacteria from these habitats commonly acquire and integrate mobile smBGCs into their genomes, a trait that can enhance their fitness and enable more cosmopolitan distributions (e.g., [5–8]).

While previous research indicates the importance of smBGCs under fully oxic conditions, the role and importance of smBGCs carried on plasmids under O₂ depleted/anoxic conditions in marine water columns has not been examined. The Cariaco Basin, Venezuela, is an end member in the continuum of ODWCs, with permanently anoxic (zero O₂), and euxinic (sulfidic) conditions below ~250 and ~275 m, respectively ([9]; Figure 1). This water column is characterized by steep oxygen and redox potential gradients (=redoxcline). Along the Cariaco redoxcline diverse particle-associated (PA: >2.7 µm) and free-living (FL: 0.22 to 2.7 µm) microbial communities exist that have varying energy and oxidant demands [10,11]. Previous analyses of metagenomic-assembled genomes (MAGs) and metatranscriptomes along the Cariaco's redoxcline provided evidence that FL and PA microbial communities encode biosynthetic gene clusters associated with various secondary metabolites [12].

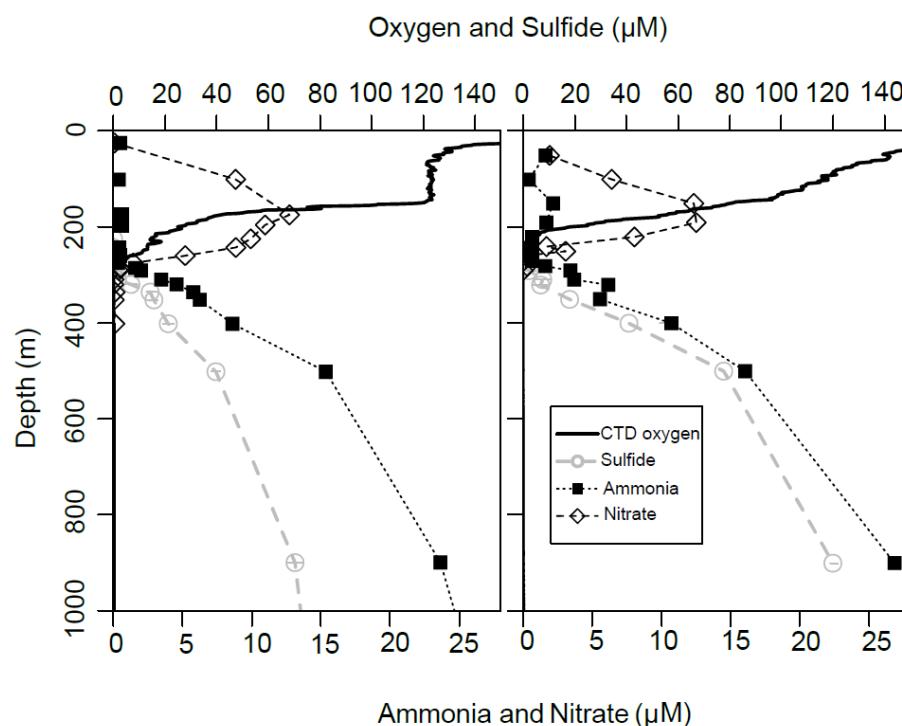


Figure 1. Depth profiles of water column biogeochemistry during May (left) and November (right) 2014.

The goal of this study was to learn whether smBGCs are carried on plasmids in low oxygen marine systems, what types of smBGCs are encoded, and by which taxa. To address this, we investigated 565 MAGs affiliated with particle-associated and free-living prokaryotic communities recovered from water samples along the Cariaco Basin redoxcline.

2. Materials and Methods

2.1. Sample Collection, Extraction and Sequencing of DNA and RNA Samples

Water samples for DNA and RNA were collected from 6 depths during two cruises in May and November 2014 in Cariaco Basin, Venezuela. Depths sampled during those cruises spanned the redoxcline of Cariaco Basin and included three sampling depths along the oxycline (O₂: ~80 µM to 0.4; Depths: 103, 196, 235 m in May, and 148, 200, 237 m

in November), two anoxic water depths at 295 m and 315 m in May, and 247 m and 267 m in November, and a euxinic water depth (zero O₂ and >80 µM H₂S) sampled from 900 m [10,12]. For DNA samples, 8–10 L water samples were gravity-filtered sequentially through EMD Millipore 2.7 µm glass fiber membranes 47 mm diameter (PA fraction), and then through 0.2 µm Sterivex filters (FL fraction), then preserved in lysis buffer. For RNA samples, 2–3 L were collected sequentially through EMD Millipore 2.7 µm glass fibre filters (for PA), then through 0.2 µm Millipore polysulfone membranes (for FL) using the Microbial Sampler- In Situ Incubation Device (MS-SID) and preserved in RNAlater [12–15]. All filters were stored frozen onboard at –20 °C and moved to –80 °C at the lab until further processing. DNA was extracted with cell lysis (lysozyme and proteinase K treatments) and phenol/chloroform as described in [10]. Extracted DNA was sent to Georgia Genomics and Bioinformatics Core for library preparation and PairedEnd 2 × 150 bp Illumina NextSeq sequencing and processed as described in [12]. Protocols for RNA extractions and sequencing cDNA libraries are reported in [12].

2.2. MAG Co-Assembly, Binning, and Taxonomic Assignment

Reconstruction of MAGs was performed as described in [12]. The trimmed reads of metagenomic datasets from the anoxic and sulfidic depths (314 and 900 m in May, 267 and 900 m in November) and both PA and FL fractions were co-assembled into contigs using SPAdes 3.11.1 with default values and flag “–meta”. Assembled contigs were binned using MetaBAT 2.12.1 with default values. CheckM 1.0.1161 was used to estimate the completeness and contamination of the reconstructed genomes. Only MAGs with ≥75% complete and ≤5% contamination were used for the downstream analysis. The taxonomic placement of the MAGs was performed with GTDB-Tk85 2.1.1. The taxonomic identification of the recovered MAGs revealed the presence of three known lab contaminants, including *Burkholderia* contaminants that were removed from any downstream analysis.

2.3. Calculation of MAG Relative Abundances and Differential Abundance Analysis

A total of 565 MAGs were recovered from the oxycline, shallow anoxic, and euxinic layers of the Cariaco Basin. To derePLICATE the MAGs, the Anvi'o 4 workflow was implemented [12]. Reads from 48 metagenomic samples (two sample types: PA- and FL-fractions, two replicates per fraction from six depths, taken over two sampling time points = 48 metagenomes) were mapped to each MAG using the BWA 2.0 aligner via the CoverM 0.6.1 command line tool [12]. Differential abundance (PA-abundant, FL-abundant or no preference) was determined for individual MAGs exhibiting differences in abundance between sample type (PA, FL) and water layer (oxycline, shallow anoxic, euxinic). A count matrix was created with rows containing individual MAG read mapping counts and with metagenomic samples as columns. Count data for all MAGs were analyzed to calculate DESeq2 1.34.0 size factors for cross-sample count normalization. The differential abundance of MAGs between size fractions and water layers was modeled using the DESeq2 negative binomial model with the metadata variables of fraction size and water layer in which “count” was the dependent variable and “fraction” as well as “water layer” were independent variables. The significant differential abundances of MAGs (with an FDR-corrected *p* < 0.05) identified by comparing the PA and FL samples were grouped by water layer, and direct comparisons were made between normalized counts of significantly differentially abundant MAGs from the oxycline, shallow anoxic, and euxinic depths.

2.4. Scanning of MAGs for BGCs, Functionally Annotating Genes within Clusters, and Comparing Mined Clusters to the MiBIG Database BGCs

Genes encoded in the 565 MAGs were predicted using Prodigal 2.6.3 and the resulting genes of each MAG were individually scanned for smBGCs using antiSMASH 6.0 with default parameters as in [12]. The GUNC command line tool and manual curation were used to assess contigs from Cariaco MAGs for potential chimerism and/or contamination. Redundancy of the predicted biosynthetic cluster sequences recovered from the Cariaco

MAGs was assessed using the BiG-SCAPE 1.1.4 command line tool with default parameters. Gene clusters with a total length less than 10 kb were discarded from downstream analysis to minimize the inclusion of fragmented smBGCs in our data (see also [12]). Genes predicted using Prodigal were scanned using the InterProScan 5.81 and Prokka 1.14.6 command line tools with default parameters for functional annotations, as well as during the implementation of the antiSMASH 6.0 pipeline with the antiSMASH HMM databases as in [12]. We manually searched the resulting annotations for genes and domains that encoded a variety of functions, such protein domains involved in post-translational modifications. Results comparing mined Cariaco smBGCs to the MiBIG database BGCs were scraped using R from the output HTML files from scanning each MAG with antiSMASH 6.0.

2.5. Identification of Putative Plasmids Encoding BGCs in the MAGs, and Metatranscriptomic Read Mapping of RNA-Seq Data to the Plasmids

All Cariaco Basin MAGs identified to contain BGCs using antiSMASH 6.0 were examined for encoding putative plasmids using PlasClass 0.1.1 [16] and PlasFlow 1.1.0 [17] tools with default settings. These tools can recover linear plasmid sequences from assembled metagenomes and can classify contigs in a metagenomic assembly that are of plasmid origin. Of the positive plasmid classifications, the contigs that contained one or more smBGC, and had a size of ≥ 10 kilobases and were classified as a plasmid by both tools were retained for downstream analyses. Visualizations of representative plasmids were constructed using SnapGene Viewer 7.0.2. Metagenomic and metatranscriptomic reads were mapped to MAG contigs using coverM (<https://github.com/wwood/CoverM>, accessed on 15 November 2022), including only alignments incorporating at least 50% of the length of a read pair with at least 95% identity. Alignment counts were normalized using the reads per kilobase million (RPKM) method. Average contig expression of MAG contigs not containing rRNA genes was calculated and compared to the expression of putative plasmid contigs across all samples using the RPKM-normalized mapping results.

2.6. Assessment of Plasmid and BGC Sequence Similarity, and Taxonomic Classification of Plasmids

The sequence similarity of the plasmid contigs and the smBGCs was assessed using Sourdash 4.5.0 [18] and CD-HIT 4.8.1 [19]. Sourdash “scaled” sizes of 20, 100, and 1000 were used with the “sourdash dna sketch” command, followed by the “sourdash compare” and “sourdash plot” commands. CD-HIT sequence identity thresholds of 60, 70, 80, and 90 specified with the “-c” flag were used with the “cd-hit” command. Additionally, the Contig Annotation Tool (CAT; [20]) was used to classify each of the plasmid contigs using default parameters with the GTDB CAT database.

3. Results

We analyzed 565 MAGs recovered from the Cariaco Basin [12], with gene annotation tools (KofamScan, InterProScan, BLASTP, DRAM; [21–23]) and methods designed to identify secondary metabolites (antiSMASH 6.0; [24]). With these analyses we found that 239 of the 565 MAGs contained one or more contigs encoding a smBGC along with genes typically associated with gene mobility and plasmid conjugation (requiring e-values $< 1 \times 10^{-15}$). These 239 MAGs were primarily affiliated with Pseudomonadota, Planctomycetota, Thermodesulfobacteriota, and Omnitrophota. While these findings indicate the presence of plasmid-transferred smBGCs in these MAGs, for accuracy we re-examined all MAGs encoding smBGCs with bioinformatic tools designed to more reliably predict plasmid sequences in genomes and metagenomes. PlasClass 0.1.1 [16] and PlasFlow 1.1.0 [17] recover linear plasmid sequences and contigs of plasmid origin from genomes and metagenomes using additional features for plasmid prediction (e.g., %GC content, k-mer frequency). Using these pipelines we identified 38 integrated plasmids (episomes; [25]) with probability scores > 0.5 that carried smBGCs (hereafter “smBGC episomes”). These 38 episomes were in 28 MAGs affiliated mainly with Planctomycetota and Pseudomonadota (14/28 MAGs; Supplementary Table S1).

We selected 14 MAGs to discuss in this paper based on the following criteria: (a) must have been identified by both PlasClass and PlasFlow tools, (b) needed to contain ≥ 1 smBGC episome that was more than 10 kilobases in size, and (c) needed to encode one or more genes indicative of the capacity for gene transfer. These genes include plasmid conjugation (*tra* and *trb* operons), pilus biogenesis and assembly (e.g., Flp/Tda proteins), and enzymes associated with plasmid partitioning (Par), transposition, integration, recombination (e.g., transposases, integrases, recombinases), with e-values stronger than 1×10^{-15} . We acknowledge that our criteria may underestimate the actual numbers of smBGC episomes in the Cariaco Basin's database. Nonetheless, we choose this approach because existing literature on mobile smBGCs in ODWCs is rare (if any), and therefore, we want to describe only those MAGs that confidently encode plasmid-borne smBGCs.

The 14 MAGs contained 16 episomes from 11 to 197 kbp in size, and were affiliated with 2 archaeal phyla (Altiarchaeota, Halobacterota) and 4 bacterial phyla (Planctomycetota, Pseudomonadota, Verrucomicrobiota, Myxococcota). Two Planctomycetota MAGs (CarAnox_MAG_1628 and CarAnox_MAG_406) each had two contigs with smBGC episomes (Table 1). Eleven MAGs encoding smBGC episomes were recovered from anoxic and euxinic depths, while three MAGs (two affiliated with Pseudomonadota and one with Planctomycota) were recovered from all examined water regimes (oxycline, shallow anoxic, euxinic; Table 1). To reveal if the identified smBGC episomes are expressed, we mapped metatranscriptome reads generated from the same examined depths [12] to the 16 smBGCs. The episomes that received transcript mapping were from anoxic and euxinic depths (Table 1). Undetected expression of genes on episomes in MAGs from other depths should be interpreted with caution.

We observed that archaeal MAGs encoding smBGC episomes were enriched either in the FL fraction (Altiarchaeota) or PA fraction (Halobacterota) along the Cariaco Basin's redoxcline (adjusted $p < 0.05$; Figure 2). For the bacterial MAGs encoding smBGC episomes we find that these bacterial phyla were either enriched in the PA fraction under all examined geochemical regimes (e.g., Verrucomicrobiota, Planctomycetota) or became enriched in the FL fraction only at certain depths (adjusted $p < 0.05$; Figure 2). As an example, all Pseudomonadota MAGs were equally abundant in the FL or PA fractions within the oxycline (adjusted $p > 0.05$), but the abundance of some Pseudomonadota MAGs was enriched in the FL fraction at 900 m (euxinic regime).

AntiSMASH analyses indicated that the 16 smBGC episomes were associated primarily with post-translationally modified peptides (RiPPs; 7/16), terpene production (3/16) and non-ribosomal peptides (NRPs; 3/16) (Table 1). Within the transferred smBGCs, other genes associated with basic metabolic and cellular functions were also detected, and included transporters and symporters of minerals and organic ions (e.g., Cu^{2+} , Na^+/H^+ , Ni, As, ABC and MFS transporters), genes involved in carbohydrate metabolism and cell wall/peptidoglycan synthesis (e.g., mannose-1-phosphate guanylyltransferase, phosphomannomutase, MurNAc-6-P etherase), and in vitamin and amino acid biosynthesis (cob(I)alamin adenosyltransferase, thiamine-monophosphate kinase, histidine ammonia-lyase, asparagine synthase). Additional genes within the smBGC clusters included genes related to plasmid conjugation and transfer (e.g., *tra* genes and genes associated with the type IV secretion system) and cellular movement (e.g., flagellum synthesis). Detailed information on the function and gene position of all identified genes carried on the smBGC episomes can be found in Supplementary Table S2.

Table 1. Description of MAGs encoding putative plasmids carrying smBGCs (smBGC episomes). The water column conditions (oxycline, anoxic, euxinic) from where these MAGs were recruited are noted inside parenthesis. Asterisk (*) indicates the archaeal MAG (CarAnox_MAG_18) and bacterial MAGs (CarAnox_MAG_1576 and CarAnox_MAG_770) used as representatives for the construction of the plasmid maps. NA: not available taxonomic assignment below class level, RPKM: Reads Per Kilobase per Million mapped reads. Repository containing all MAGs can be found at <https://osf.io/usm8r/> (accessed on 13 December 2023). The DNA coverage for MAGs at the three examined biogeochemical regimes can be found at <https://zenodo.org/records/10144545> (accessed on 22 July 2023). NA: Not Available.

MAG and Contig Encoding Putative Plasmids	MAG's Completeness/Contamination	Geochemical Regime	Transcript Read Recruitment that Plasmids Received [RPKM/Depth (m)]	smBGCs Encoded in the Plasmids	Secondary Metabolite	Phylum/Order/Family/Genus
CarAnox_MAG_1243_4	96.88/0.93	Anoxic, Euxinic	0	RiPP	RiPP	Altarchaeota/ SCGC-AAA252-I15/ SCGC-AAA252-I15/NA
* CarAnox_MAG_18_14	96.73/0.33	Euxinic	0.64/900	RiPP	RRE-containing	Halobacterota/ Syntrophoarchaeales/ Syntrophoarchaeaceae/ <i>Syntrophoarchaeum</i> sp.
CarAnox_MAG_1071_28	77/2.7	Anoxic, Euxinic	0	Terpene	Terpene	Verrucomicrobiota/ Kiritimatiellae/ YA12-FULL-48-11/NA
CarAnox_MAG_1075_66	91.22/2.48	Anoxic, Euxinic	0	Other	Resorcinol	Verrucomicrobiota/ Kiritimatiellae/ UBA8416/NA
CarAnox_MAG_828_1	96.28/2.39	Anoxic, Euxinic	0.03/267	RiPP hybrid	RiPP-like	Verrucomicrobiota/ Kiritimatiellae/ UBA8416/NA
CarAnox_MAG_1543_26	88.92/1.7	Euxinic	0	PKS hybrid	hglE-KS	Planctomycetota/ Phycisphaerales/ Phycisphaeraceae/NA
CarAnox_MAG_1628_60	95.33/2.4	Euxinic	0	Terpene	Terpene	Planctomycetota/ Pirellulales/ UBA1268/UBA1268 sp002694955

Table 1. *Cont.*

MAG and Contig Encoding Putative Plasmids	MAG's Completeness/Contamination	Geochemical Regime	Transcript Read Recruitment that Plasmids Received [RPKM/Depth (m)]	smBGCs Encoded in the Plasmids	Secondary Metabolite	Phylum/Order/Family/Genus
CarAnox_MAG_1628_81	95.33/2.4	Euxinic	0	NRPS	NRPS-like	Planctomycetota/ Pirellulales/ UBA1268/UBA1268 sp002694955
CarAnox_MAG_245_32	98.8/1.14	Euxinic	0.07/900	RiPP	RRE-containing	Planctomycetota/ Brocadiae/ SM23-32/NA
CarAnox_MAG_406_14	88.41/4.55	Anoxic, Euxinic	0	NRPS	NRPS-like	Planctomycetota/ Brocadiae/ NA/NA
CarAnox_MAG_406_39	88.41/4.55	Anoxic, euxinic	0	RiPP	RRE-containing	Planctomycetota/ Brocadiae/ NA/NA
CarAnox_MAG_879_3	86.59/3.98	Oxycline, Anoxic, Euxinic	0.05/314	Terpene	Terpene	Planctomycetota/ Phycisphaerales/ SM1A02/UBA12567
* CarAnox_MAG_1576_52	99.59/1.37	Oxycline, Anoxic, Euxinic	0.33/267	RiPP	RiPP	Pseudomonadota/ Sphingomonadales/ Sphingomonadaceae/ <i>Novosphingobium</i> sp.

Table 1. Cont.

MAG and Contig Encoding Putative Plasmids	MAG's Completeness/Contamination	Geochemical Regime	Transcript Read Recruitment that Plasmids Received [RPKM/Depth (m)]	smBGCs Encoded in the Plasmids	Secondary Metabolite	Phylum/Order/Family/Genus
CarAnox_MAG_164_53	88/2.1	Euxinic	0	RiPP	RiPP	Pseudomonadota/ Pseudomonales/ Pseudomonadaceae/ <i>Pseudomonas</i> sp.
CarAnox_MAG_782_37	88.99/0.8	Oxycline, Anoxic, Euxinic	0.09/148	RiPP	RiPP	Pseudomonadota/ Sphingomonadales/ Sphingomonadaceae/ <i>Erythrobacter</i> sp.
* CarAnox_MAG_770_69	83.87/4.52	Anoxic	0.06/200	NRPS	NRPS-like	Myxococcota/ Bradimonadia/ UBA7976/ UBA1532/NA

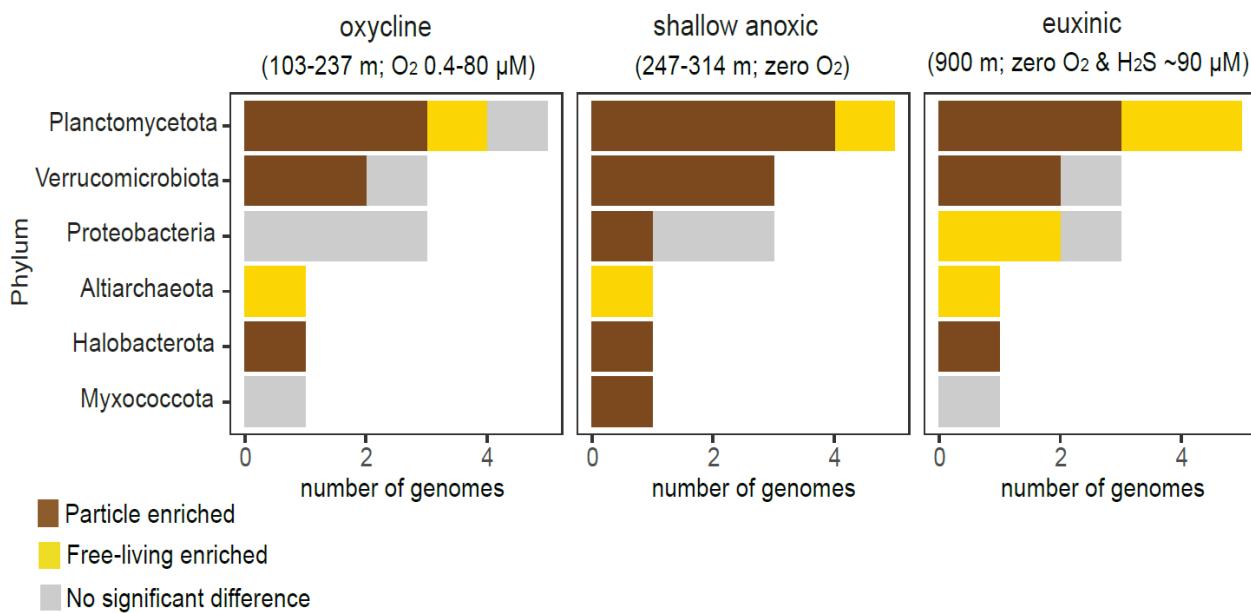


Figure 2. Preference of MAGs for particle-associated and free-living fractions along the oxycline (103–237 m), shallow anoxic (247–314 m) and euxinic regimes (900 m). “No significant difference” indicates adjusted *p* values > 0.05.

To elaborate further into the gene content and gene position of the smBGC episomes, we selected one representative archaeal and two representative bacterial episomes, and constructed their maps (Figure 3). Specifically, the representative maps were constructed for (a) an archaeal MAG annotated to Syntrophoarchaeia encoding a RiPP recognition element (RRE) smBGC (CarAnox_MAG_18_14) (Figure 3a), (b) a bacterial MAG affiliated with Myxococcota encoding a RiPP smBGC (CarAnox_MAG_770_69) (Figure 3b), and (c) a bacterial MAG affiliated with Alphaproteobacteria encoding a NRP smBGC (CarAnox_MAG_1576_52) (Figure 3c). The Syntrophoarchaeia MAG was selected because it is one of two archaeal MAGs carrying smBGC episomes, and the one that received transcript mapping, suggesting the episome is expressed (Table 1). The MAG annotated to Alphaproteobacteria was selected because it encodes a suite of well-described genes typically involved in plasmid transfer (*tra* genes). Finally, the Myxococcota MAG was of interest because little is known about the ecological role and genomic structure of this taxon, particularly in ODWCs (e.g., [26–28]). It is postulated that Myxococcota produce a plethora of secondary metabolites that support their different ‘social’ traits in the marine environment (e.g., predation, sporulation, fruiting body formation; [29–31]). We also note that signatures of these selected representative taxa (Syntrophoarchaeia, Alphaproteobacteria, Myxococcota) are abundantly detected in the Cariaco Basin [10,12] and in similar, permanently anoxic pelagic systems (e.g., Black Sea; [32]). Aside from smBGCs and genes associated with plasmid conjugation/mobility, these maps also revealed episome-encoded genes for transporters, rRNA and tRNA modifications, cell signaling (e.g., kinases and proteins for cell division), transcription, nitrogen metabolism, and two-component regulatory systems involved in chemotaxis. In addition, they also contained diverse restriction enzymes sites, a common feature of prokaryotic restriction modification systems that recognize mobile genetic elements, with more restriction enzymes sites detected in the representative archaeal episome (Figure 3).

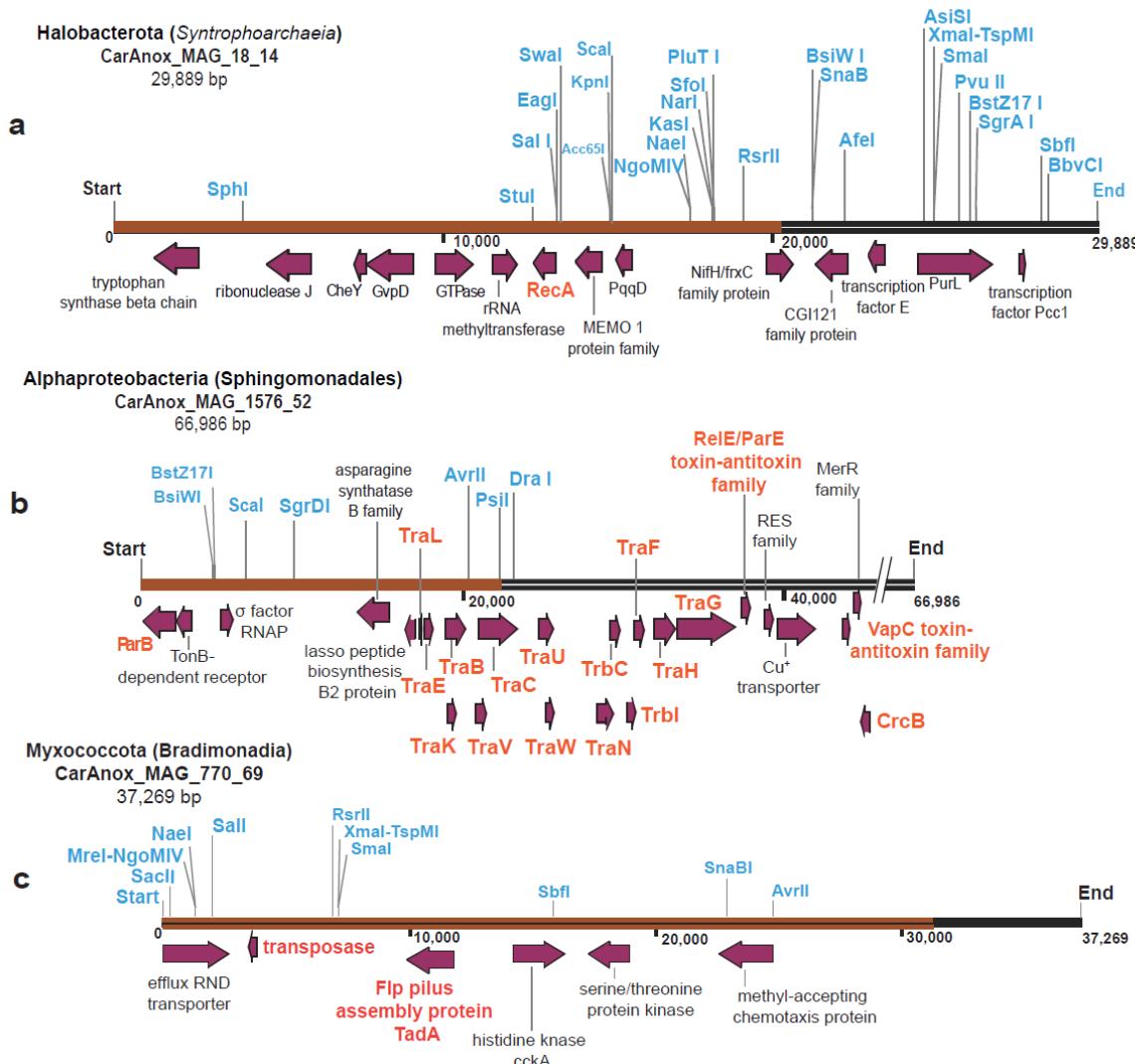


Figure 3. (a). Map of the representative archaeal episome identified in CarAnox_MAG_18 contig 14 that encoded RRE (RiPP) smBGC (brown color). (b). Map of the representative bacterial episome identified in CarAnox_MAG_1576 contig 52 that encoded RiPP smBGC (brown color). (c). Map of the representative bacterial episome identified in CarAnox_MAG_770 contig 69 that encoded NRP smBGC (brown color). Genes associated with plasmid conjugation, partitioning, recombination, pilus biogenesis, and toxin/antitoxin systems (indicators of plasmid mobility) are shown in orange. Enzyme restriction sites are shown in blue. The gene annotation for all genes encoded on the representative episomes shown in Figure 3a–c can be found in <https://zenodo.org/records/10144545> (accessed on 22 July 2023).

Sourmash [18] was used to examine sequence similarity for all 38 episomes carrying smBGCs (Supplementary Table S1) and indicated that all episomes contained distinct k -mer profiles. Clustering of our episome sequences using CD-HIT [19] at 70% sequence identity threshold did not yield any clusters. Altogether, this suggests that recovered episomes have low sequence similarity. To address the putative taxonomic origin of the identified mobile smBGCs we ran the Contig Annotation Tool (CAT; [20]). These results were largely consistent with taxonomic annotations of MAGs originally assigned by GTDB-Tk 2.1.1 and are presented in Supplementary Table S1.

4. Discussion

While >34,000 plasmids exist in the public plasmid database PLSDB [33] many plasmids in environmental samples remain undetected. This is because plasmid features

including structure, size and gene content do not allow bioinformatic/computational tools to easily distinguish plasmid-encoded DNA from chromosomal DNA fragments (e.g., [34–37]). Further, existing bioinformatic pipelines for plasmid identification are typically designed for complete or representative genomes, and therefore, when used to analyze MAGs reconstructed from complex environmental communities (e.g., Cariaco Basin), plasmid prediction and identification can be challenging. Finally, many of the existing plasmid prediction bioinformatic tools are trained against reference databases (PLSDB) that usually contain sequences from circular, not linear, plasmids, and originate from bacteria affiliated with human pathogens, terrestrial, and freshwater taxa [16,33]. As an example, Androsiuk et al. (2023) [37] examined the marine plasmidome of the Red Sea, and they were able to conclusively identify only 7 out of 362 putative plasmids from metagenomic data. The authors pointed out the shortcomings of existing tools and databases that lead to underestimation of the marine plasmidome [37].

Using samples collected from distinct geochemical regimes within the Cariaco Basin's water column, we investigated the presence of plasmid-borne smBGCs in MAGs affiliated with particle-associated (PA) and free-living (FL) prokaryotic communities. We identified 38 integrated plasmids in 28 MAGs (Supplementary Table S1). The small number of MAGs carrying integrated smBGCs is not a surprise considering the shortcomings of existing tools, as well as other possible ecological explanations discussed in detail below. Overall, the limited number of MAGs chosen for discussion in our Cariaco study precludes drawing conclusions about whether smBGC episomes are more commonly found in PA vs. FL prokaryotes of Cariaco Basin. DNA transfer via plasmids within PA communities is expected to be more efficient than those in the FL size fraction [38]. However, some debate exists because surface ionic charges on marine particles (as for sediments) can enhance DNA binding to particle surfaces. Consequently, transformation efficiency (uptake) of plasmid DNA may be reduced on particles [39,40]. Further, Niskin bottle-collected water samples primarily capture small, suspended particles and cells rather than rare, rapidly sinking particles [41]. Only by capturing a full representation of all sinking particles (such as collected by sediment traps) would it be possible to fully assess the role of smBGC episomes in the PA water fraction. Nonetheless, we postulate that plasmid-encoded smBGCs may present more selective advantages to particle-associated communities than to free-living counterparts for several reasons. Previous studies of Cariaco Basin showed that suspended particles are hotspots of microbial activity in the water column, and that *in situ* PA microbes depend on secondary metabolites to compete for nutrient resources [10,12]. Overall, particles can protect DNA from DNase degradation [42–44], facilitate bacterial colonization, and can sustain higher densities of active cells because they are nutrient-rich (e.g., [45]). Higher cell densities allow cells to interact more efficiently than those in the diffuse free-living fraction of the water column, likely enhancing transformation efficiency [46]. Among the prokaryotes that can acquire smBGCs through plasmids, those attached to particles in the Cariaco Basin may have a practical advantage for acquiring exogenous DNA for reasons discussed above, however, this process requires further examination.

Thirteen of the 16 identified smBGCs episomes encoded RiPPs, RiPP recognition elements (RREs), NRPs and terpene-synthesizing enzymes. RREs are conserved regions that direct the post-translational modifications of RiPPs [47]. RiPPs, NRPs and terpenes are known to have antimicrobial properties which may help microorganisms to outcompete their microbial "neighbors". Ocean deoxygenation can compress habitats for sensitive taxa and select against taxa that cannot adapt to declining oxygen [48,49]. Prokaryotes are known to retain plasmid-encoded genes in challenging environments even when these accessory genes are not immediately used [50,51]. Transfer and retention of plasmid-encoded smBGCs with antimicrobial effects could benefit those taxa in the Cariaco Basin that can acquire them, because they can increase the versatility of their hosts under dynamic and O₂-depleted/anoxic conditions [49,52,53].

Retention of plasmid-borne genes in host genomes is strongest when the encoded genes enhance host fitness, or when these genes are rare or absent from the host's

genome [54]. Letzel et al. (2017) [55] proposed that mobile transfer of smBGCs can enable a “plug and play” type of evolution that tests relative fitness of specialized metabolites produced under stressful conditions in marine bacteria. The smBGCs we identified in this study (e.g., RiPPs, terpene synthases) are known from the literature to have diverse functions. As an example, in addition to their antimicrobial effects, RiPPs may serve as quorum sensing signaling molecules, and terpenes may participate in cell fate decision (e.g., sporulation) [56–60]. Metagenomic studies suggest that acquisition of smBGCs with pleiotropic functions increases the secondary metabolic diversity and contributes to host fitness (e.g., [61]). Studies of cultured bacterial strains from ODWCs can shed additional light on the function(s) of these compounds.

As previously mentioned, identification of marine smBGC episomes is likely hindered by limitations of bioinformatic pipelines and databases, as well as by the difficulties of culturing many marine prokaryotic strains (e.g., [62–64]). Within Cariaco Basin’s redoxcline, the majority of recovered MAGs were associated with Omnitrophota (56), Chloroflexota (49), Myxococcota (33), Planctomycetota (65), Desulfobacterota (43) (now classified as Thermodesulfobacteriota; [65,66]) that collectively comprised ~55% of the assembled bacterial MAGs (463 bacterial and 102 archaeal MAGs, 565 MAGs in total; see Supplementary Data and Figure 1a from [12]). Some of these phyla (e.g., Omnitrophota) are difficult to cultivate, and hence, information on their overall genomic organization is limited. However, for other phyla (e.g., Thermodesulfobacteriota) more detailed information exists as they have also been cultured under laboratory conditions and are well-characterized. In the present study, our gene annotation and antiSMASH analyses revealed MAGs encoding at least one smBGC along with genes associated with plasmid transfer. While Omnitrophota MAGs are among the 239 MAGs, it is not yet clear whether this bacterial phylum can carry plasmid-transferred smBGCs. This is because we did not identify smBGC episomes within Omnitrophota MAGs through the conservative plasmid prediction pipelines that we applied (Supplementary Table S1). Similarly, based on our gene annotation and antiSMASH analyses, Thermodesulfobacteriota MAGs were identified to encode smBGC episomes along with genes involved in transfer of plasmid DNA. This agrees with previous literature that demonstrated that sulfate reducers carry plasmids [67] and are competent to exchange mobile DNA via conjugation with other sulfate reducers (e.g., *Desulfovibrio* sp.; [68,69] or even with other bacterial phyla (e.g., *E. coli* cells; [70,71] and references therein]). However, only one Thermodesulfobacteriota MAG (CarAnox_MAG_1179) was identified using the PlasFlow and PlasClass tools, which may also be due to the caveats discussed above.

While bioinformatic limitations exist, we offer a putative ecological explanation for the few MAGs identified to carry smBGC episomes in our datasets. In silico analyses of prokaryotic phyla that had at least 50 genomes deposited in GenBank [72] suggested that common taxa having high plasmid prevalence in their genomes include Pseudomonadota, Cyanobacteria, Actinobacteria, Firmicutes, and Verrucomicrobiota [72]. Overall, those bacteria with a high prevalence of plasmids are usually cosmopolitan taxa [72,73] that can be found ubiquitously in the water column (e.g., [74–77]). The presence of naturally competent bacteria is a prerequisite for exogenous DNA uptake, and in the marine environment bacteria known to be naturally competent include Alphaproteobacteria and Gammaproteobacteria [2,37,78–80]. In the Cariaco Basin MAG dataset that we examined here for smMBGCs, Alphaproteobacteria and Gammaproteobacteria comprised 44 out of the 463 reconstructed bacterial MAGs (9.5%), while MAGs associated with Cyanobacteria, Actinobacteria, Firmicutes and Verrucomicrobiota collectively represented 5% of the bacterial MAGs (see Supplementary Data and Figure 1a from [12]). Hence, phyla described by [72] to have a high plasmid prevalence represented only a small fraction of the bacterial MAGs recovered (14.5%; ≥75% completeness and <5% contamination). Additionally, ref. [72] suggests that lineages like Chloroflexota and Planctomycetota, constitute ~25% of our recovered MAGs, have low plasmid prevalence in their genomes. We also know from available genomic data for taxa with low plasmid prevalence (e.g., Planctomycetota), that it is mainly their terrestrial [81], and not marine representatives [82], that carry plasmids.

However, this perception might change in the future when culturing methods for Planctomycetota advance. Planctomycetota are currently difficult to culture because of their innate ability to resist many antibiotics, as well as their inability to utilize common growth media used for microbiology assays [64]. Finally, we need to underscore that exchange of genetic material within prokaryotic populations in Cariaco and other ODWCs occurs also through other mechanisms, like viruses and their auxiliary metabolic genes [83–85]. Viral transfer of genetic material may be a more common mechanism for transfer of mobile DNA in marine ecosystems, including ODWCs and Cariaco Basin. Collectively, these putative ecological explanations along with bioinformatic caveats discussed above may explain the recovery of relatively few MAGs encoding smBGC episomes in our dataset. While our data highlight the need for improved tools and databases for detecting marine plasmids, we note there is still much to learn about the balance between viral-mediated and plasmid-mediated transfer of genetic material in ODWCs.

5. Conclusions

We searched 565 published MAGs recovered from water samples collected within the Cariaco Basin's redoxcline for plasmid-encoded secondary metabolite biosynthetic gene clusters (smBGC episomes). We found 28 MAGs that contained 38 putative smBGC episomes, and we discuss 14 MAGs for which we have the strongest evidence. These 14 MAGs contain 16 smBGC episomes that encode genes for RiPPs, RiPP recognition elements (RREs), terpene synthesis and NRPs, and are affiliated with 2 archaeal phyla (Altiarchaeota, Halobacterota) and 4 bacterial phyla (Planctomycetota, Verrucomycota, Myxococcota and Pseudomonadota). These secondary metabolites may enhance the competitive advantage of free-living and particle-associated host taxa that are competent to acquire plasmid DNA in this geochemically stratified water column. However, until comparative data on smBGC episomes from other ODWCs is available, and bioinformatic tools and databases for detecting and identifying smBGC episomes improve, it is difficult to assess their evolutionary and ecological significance in the Cariaco Basin and other oxygen-depleted water column habitats.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/microorganisms12050929/s1>, Supplementary Table S1. Identified Cariaco MAGs containing plasmid-borne smBGCs using the two plasmid prediction tools (PlasFlow and PlasClass) (provided as excel file); Supplementary Table S2. Function and position of the encoded genes within, and outside the borders of the plasmid-borne smBGC clusters (provided as excel file). Annotations were performed using multiple databases which are noted in the separate tabs of the excel file.

Author Contributions: P.M. conceived of the analysis. M.G.P., V.P.E., E.S. and G.T.T. collected samples. M.G.P. extracted DNA and RNA from samples, assembled the metagenomes, and binned metagenome-assembled genomes. D.G.-M. performed genome-resolved metagenomics and metatranscriptomics, and ran bioinformatic analyses of plasmids. P.M. and V.P.E. wrote the paper, and all authors contributed to its final form. All authors have read and agreed to the published version of the manuscript.

Funding: The fieldwork that provided samples and data for this study was supported by National Science Foundation (NSF) grants (OCE-1336082 to V.E. and OCE-1335436 and OCE-1259110 to G.T.T., Stony Brook University).

Data Availability Statement: The metagenome and metatranscriptome data generated in this study have been deposited in the NCBI database under accession code PRJNA326482. The processed metagenome-assembled genomes (in FASTA format) and biosynthetic gene cluster files (in ZIP format) are available at OSF [<https://osf.io/usm8r/>] accessed on 13 December 2023]. The biogeochemistry data from the CARIACO Basin Time Series Station for May to November 2014 are available through the Biological and Chemical Oceanography Data Management Office (BCO-DMO) at the Woods Hole Oceanographic Institution [<https://www.bco-dmo.org/dataset/652313/data> accessed on 22

July 2016]. The scripts used for these analyses are available at the following GitHub repository: https://github.com/d-mcgrath/cariaco_plasmids accessed on 6 February 2023.

Acknowledgments: We thank the staff of Fundación La Salle de Ciencias Naturales (FLASA), EDI-MAR, Porlamar, Edo Nueva Esparta, Venezuela, and the crew of the R/V Hermano Ginés for their support during fieldwork for this study, especially Y. Astor and R. Varela. The fieldwork that provided samples and data for this study was supported by National Science Foundation (NSF) grants (OCE-1336082 to V.E. and OCE-1335436 and OCE-1259110 to G.T.T., Stony Brook University).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sandaa, R.A.; Enger, O. Transfer in marine sediments of the naturally occurring plasmid pRAS1 encoding multiple antibiotic resistance. *Appl. Environ. Microbiol.* **1994**, *60*, 4234–4238. [CrossRef] [PubMed]
2. Sobecky, P.A.; Mincer, T.J.; Chang, M.C.; Helinski, D.R. Plasmids isolated from marine sediment microbial communities contain replication and incompatibility regions unrelated to those of known plasmid groups. *Appl. Environ. Microbiol.* **1997**, *63*, 888–895. [CrossRef] [PubMed]
3. Loper, J.E.; Buyer, J.S. Siderophores in microbial interactions on plant surfaces. *MPMI* **1991**, *4*, 5–13. [CrossRef]
4. Paoli, L.; Ruscheweyh, H.-J.; Forneris, C.C.; Hubrich, F.; Kautsar, S.; Bhushan, A.; Lotti, A.; Clayssen, Q.; Salazar, G.; Milanese, A.; et al. Biosynthetic potential of the global ocean microbiome. *Nature* **2022**, *607*, 111–118. [CrossRef] [PubMed]
5. Penn, K.; Jenkins, C.; Nett, M.; Udwary, D.W.; Gontang, E.A.; McGlinchey, R.P. Genomic islands link secondary metabolism to functional adaptation in marine Actinobacteria. *ISME J.* **2009**, *3*, 1193–1203. [CrossRef] [PubMed]
6. Penn, K.; Wang, J.; Fernando, S.C.; Thompson, J.R. Secondary metabolite gene expression and interplay of bacterial functions in a tropical freshwater cyanobacterial bloom. *ISME J.* **2014**, *8*, 1866–1878. [CrossRef] [PubMed]
7. Ziemert, N.; Lechner, A.; Wietz, M.; Millán-Aguiñaga, N.; Chavarria, K.L.; Jensen, P.R. Diversity and evolution of secondary metabolism in the marine actinomycete genus *Salinispora*. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E1130–E1139. [CrossRef]
8. Popin, R.V.; Alvarenga, D.O.; Castelo-Branco, R.; Fewer, D.P.; Sivonen, K. Mining of cyanobacterial genomes indicates natural product biosynthetic gene clusters located in conjugative plasmids. *Front. Microbiol.* **2021**, *12*, 684565. [CrossRef] [PubMed]
9. Scranton, M.I.; Taylor, G.T.; Thunell, R.; Benitez-Nelson, C.R.; Muller-Karger, F.; Fanning, K.; Lorenzoni, L.; Montes, E.; Varela, R.; Astor, Y. Interannual and subdecadal variability in the nutrient geochemistry of the Cariaco Basin. *Oceanography* **2014**, *27*, 148–159. [CrossRef]
10. Suter, E.A.; Pachiadaki, M.; Taylor, G.T.; Astor, Y.; Edgcomb, V.P. Free-living chemoautotrophic and particle-attached heterotrophic prokaryotes dominate microbial assemblages along a pelagic redox gradient. *Environ. Microbiol.* **2018**, *20*, 693–712. [CrossRef] [PubMed]
11. Suter, E.A.; Pachiadaki, M.G.; Montes, E.; Edgcomb, V.P.; Scranton, M.I.; Taylor, C.D.; Taylor, G.T. Diverse nitrogen cycling pathways across a marine oxygen gradient indicate nitrogen loss coupled to chemoautotrophic activity. *Environ. Microbiol.* **2021**, *23*, 2747–2764. [CrossRef] [PubMed]
12. Geller-McGrath, D.; Mara, P.; Taylor, G.T.; Suter, E.; Edgcomb, V.P.; Pachiadaki, M.G. Diverse secondary metabolites are expressed in particle-associated and free-living microorganisms of the permanently anoxic Cariaco Basin. *Nat. Commun.* **2023**, *14*, 656. [CrossRef] [PubMed]
13. Taylor, C.D.; Doherty, K.W. Submersible incubation device (SID), autonomous instrumentation for the in situ measurement of primary production and other microbial rate processes. *Deep Sea Res. Part. A Ocean. Res. Pap.* **1990**, *37*, 343–358. [CrossRef]
14. Taylor, C.D.; Edgcomb, V.P.; Doherty, K.W.; Engstrom, I.; Shanahan, T.; Pachiadaki, M.G.; Molyneaux, S.; Honjo, S. Fixation filter, device for the rapid in situ preservation of particulate samples. *Deep Res. Part. I Ocean. Res. Pap.* **2015**, *96*, 69–79. [CrossRef]
15. Edgcomb, V.P.; Taylor, C.; Pachiadaki, M.G.; Honjo, S.; Engstrom, I.; Yakimov, M. Comparison of Niskin vs. in situ approaches for analysis of gene expression in deep Mediterranean Sea water samples. *Deep Res. Part. II* **2016**, *129*, 213–222. [CrossRef]
16. Pellow, D.; Mizrahi, I.; Shamir, R. PlasClass improves plasmid sequence classification. *PLoS Comput. Biol.* **2020**, *16*, e1007781. [CrossRef] [PubMed]
17. Krawczyk, P.S.; Lipinski, L.; Dziembowski, A. PlasFlow: Predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Res.* **2018**, *46*, e35. [CrossRef] [PubMed]
18. Brown, C.T.; Irber, L. sourmash: A library for MinHash sketching of, D.N.A. *J. Open Source Softw.* **2016**, *1*, 27. [CrossRef]
19. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* **2012**, *28*, 3150–3152. [CrossRef] [PubMed]
20. von Meijenfeldt, F.A.B.; Arkhipova, K.; Cambuy, D.D.; Coutinho, F.H.; Dutilh, B.E. Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT. *Genome Biol.* **2019**, *20*, 217. [CrossRef]
21. Aramaki, T.; Blanc-Mathieu, R.; Endo, H.; Ohkubo, K.; Kanehisa, M.; Goto, S.; Ogata, H. KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **2019**, *36*, 2251–2252. [CrossRef] [PubMed]
22. Shaffer, M.; Borton, M.A.; McGivern, B.B.; Zayed, A.A.; Leanti La Rosa, S.; Soden, L.M.; Liu, P.; Narrowe, A.B.; Rodríguez-Ramos, J.; Bolduc, B.; et al. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res.* **2020**, *48*, 8883–8900. [CrossRef] [PubMed]

23. Jones, P.; Binns, D.; Chang, H.-Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **2014**, *30*, 1236–1240. [CrossRef] [PubMed]
24. Blin, K.; Kim, H.U.; Medema, M.H.; Weber, T. Recent development of antiSMASH and other computational approaches to mine secondary metabolite biosynthetic gene clusters. *Brief. Bioinform.* **2019**, *20*, 1103–1113. [CrossRef] [PubMed]
25. Campbell, A. Episomes. In *Bacteria, Bacteriophages, and Fungi*; King, R.C., Ed.; Springer: Boston, MA, USA, 1974.
26. Li, L.; Huang, D.; Hu, Y.; Rudling, N.M.; Canniffe, D.P.; Wang, F.; Wang, W. Globally distributed Myxococcota with photosynthesis gene clusters illuminate the origin and evolution of a potentially chimeric lifestyle. *Nat. Commun.* **2023**, *14*, 6450. [CrossRef] [PubMed]
27. Wang, C.; Xiao, Y.; Wang, Y.; Liu, Y.; Yao, Q.; Zhu, H. Comparative genomics and transcriptomics insight into myxobacterial metabolism potentials and multiple predatory strategies. *Front. Microbiol.* **2023**, *14*, 1146523. [CrossRef] [PubMed]
28. Acinas, S.G.; Sánchez, P.; Salazar, G.; Cornejo-Castillo, F.M.; Sebastián, M.; Logares, R.; Royo-Llonch, M.; Paoli, L.; Sunagawa, S.; Hingamp, P.; et al. Deep ocean metagenomes provide insight into the metabolic architecture of bathypelagic microbial communities. *Commun. Biol.* **2021**, *4*, 604. [CrossRef] [PubMed]
29. Thiery, S.; Kaimer, C. The predation strategy of *Myxococcus xanthus*. *Front. Microbiol.* **2020**, *11*, 2. [CrossRef] [PubMed]
30. Mohr, K.I. Diversity of Myxobacteria—We Only See the Tip of the Iceberg. *Microorganisms* **2018**, *6*, 84. [CrossRef] [PubMed]
31. Murphy, C.L.; Sheremet, A.; Dunfield, P.F.; Spear, J.R.; Stepanauskas, R.; Woyke, T.; Elshahed, M.S.; Youssef, N.H. Genomic Analysis of the Yet-Uncultured Binatota Reveals Broad Methylotrophic, Alkane-Degradation, and Pigment Production Capacities. *mBio* **2021**, *12*, e00985-21. [CrossRef]
32. Cabello-Yeves, P.J.; Callieri, C.; Picazo, A.; Mehrshad, M.; Haro-Moreno, J.M.; Roda-Garcia, J.J.; Dzhembekova, N.; Slabakova, V.; Slabakova, N.; Moncheva, S.; et al. The microbiome of the Black Sea water column analyzed by shotgun and genome centric metagenomics. *Env. Microbiome* **2021**, *16*, 5. [CrossRef] [PubMed]
33. Galata, V.; Fehlmann, T.; Backes, C.; Keller, A. PLSDB: A resource of complete bacterial plasmids. *Nucleic Acids Res.* **2018**, *47*, D195–D202. [CrossRef] [PubMed]
34. Antipov, D.; Hartwick, N.; Shen, M.; Raiko, M.; Lapidus, A.; Pevzner, P. plasmidSPAdes: Assembling plasmids from whole genome sequencing data. *Bioinformatics* **2016**, *32*, 3380–3387. [CrossRef] [PubMed]
35. Antipov, D.; Raiko, M.; Lapidus, A.; Pevzner, P.A. Plasmid detection and assembly in genomic and metagenomic data sets. *Genome Res.* **2019**, *29*, 961–968. [CrossRef] [PubMed]
36. Rozov, R.; Brown Kav, A.; Bogumil, D.; Shterzer, N.; Halperin, E.; Mizrahi, I.; Shamir, R. Recycler: An algorithm for detecting plasmids from de novo assembly graphs. *Bioinformatics* **2017**, *33*, 475–482. [CrossRef] [PubMed]
37. Androsiuk, L.; Shay, T.; Tal, S. Characterization of the Environmental Plasmidome of the Red Sea. *Microbiol. Spectr.* **2023**, *11*, e0040023. [CrossRef]
38. Hermansson, M.; Linberg, C. Gene transfer in the marine environment. *FEMS Microbiol. Ecol.* **1994**, *15*, 47–54. [CrossRef]
39. Herndl, G.J.; Reinharter, T. Microbial control of the dark end of the biological pump. *Nat. Geosci.* **2013**, *6*, 718–724. [CrossRef] [PubMed]
40. Nagler, M.; Insam, H.; Pietramellara, G.; Ascher-Jenull, J. Extracellular DNA in natural environments: Features, relevance and applications. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 6343–6356. [CrossRef] [PubMed] [PubMed Central]
41. Suter, E.A.; Scranton, M.I.; Chow, S.; Stinton, D.; Medina Faull, L.; Taylor, G.T. Niskin bottle sample collection aliases microbial community composition and biogeochemical interpretation. *Limnol. Oceanogr.* **2017**, *62*, 606–617. [CrossRef]
42. Lorenz, M.G.; Aardema, B.W.; Krumbein, W.E. Interaction of marine sediment with DNA and DNA availability to nucleases. *Mar. Biol.* **1981**, *64*, 225–230. [CrossRef]
43. Lorenz, M.G.; Wackernagel, W. Adsorption of DNA to sand and variable degradation rates of adsorbed DNA. *Appl. Environ. Microbiol.* **1987**, *53*, 2945–2952. [CrossRef] [PubMed]
44. Yin, X.; Stotzky, G. Gene transfer among bacteria in natural environments. *Adv. Appl. Microbiol.* **1997**, *45*, 153–212. [PubMed]
45. van Elsas, J.D.; Bailey, M.J. The ecology of transfer of mobile genetic elements. *FEMS Microbiol. Ecol.* **2002**, *42*, 187–197. [CrossRef] [PubMed]
46. Hasegawa, H.; Suzuki, E.; Saeda, S. Horizontal plasmid transfer by transformation in *Escherichia coli*: Environmental factors and possible mechanisms. *Front. Microbiol.* **2018**, *9*, 2365. [CrossRef] [PubMed]
47. Kloosterman, A.M.; Shelton, K.E.; van Wezel, G.P.; Medema, M.H.; Mitchell, D.A. RRE-Finder: A genome-mining tool for class-independent RiPP discovery. *mSystems* **2020**, *5*, e00267-20. [CrossRef] [PubMed]
48. Wishner, K.F.; Seibel, B.A.; Roman, C.; Deutsch, C.; Outram, D.; Shaw, C.T.; Birk, M.A.; Mislan, K.A.S.; Adams, T.J.; Moore, D.; et al. Ocean deoxygenation and zooplankton: Very small oxygen differences matter. *Sci. Adv.* **2018**, *4*, eaau5180. [CrossRef] [PubMed]
49. Li, L.; Dechesne, A.; Madsen, J.S.; Nesme, J.; Sørensen, S.J.; Smets, B.F. Plasmids persist in a microbial community by providing fitness benefit to multiple phylotypes. *ISME J.* **2020**, *14*, 1170–1181. [CrossRef] [PubMed]
50. Harrison, E.; Brockhurst, M.E. Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends Microbiol.* **2012**, *20*, 262–267. [CrossRef]
51. Harrison, E.; Dytham, C.; Hall, J.P.; Guymer, D.; Spiers, A.J.; Paterson, S.; Brockhurst, M.A. Rapid compensatory evolution promotes the survival of conjugative plasmids. *Mob. Genet. Elements.* **2016**, *6*, e1179074. [CrossRef] [PubMed]

52. Jaén, K.E.; Sigala, J.C.; Olivares-Hernández, R.; Niehaus, K.; Lara, A.R. Heterogeneous oxygen availability affects the titer and topology but not the fidelity of plasmid DNA produced by *Escherichia coli*. *BMC Biotechnol.* **2017**, *17*, 60. [[CrossRef](#)] [[PubMed](#)]
53. Lara, A.R.; Jaén, K.E.; Folarin, O.; Keshavarz-Moore, E.; Büchs, J. Effect of the oxygen transfer rate on oxygen-limited production of plasmid DNA by *Escherichia coli*. *Biochem. Eng. J.* **2019**, *150*, 107303. [[CrossRef](#)]
54. Lehtinen, S.; Huismans, J.S.; Bonhoeffer, S. Evolutionary mechanisms that determine which bacterial genes are carried on plasmids. *Evol. Lett.* **2021**, *5*, 290–301. [[CrossRef](#)] [[PubMed](#)]
55. Letzel, A.-C.; Li, J.; Amos, G.C.A.; Millán-Aguiñaga, N.; Ginigini, J.; Abdelmohsen, U.R.; Gaudêncio, S.P.; Ziemert, N.; Moore, B.S.; Jensen, P.R. Genomic insights into specialized metabolism in the marine actinomycete *Salinispora*. *Environ. Microbiol.* **2017**, *19*, 3660–3673. [[CrossRef](#)] [[PubMed](#)]
56. Chevrette, M.G.; Thomas, C.S.; Hurley, A.; Rosario-Meléndez, N.; Sankaran, K.; Tu, Y.; Hall, A.; Magesh, S.; Handelsman, J. Microbiome composition modulates secondary metabolism in a multispecies bacterial community. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2212930119. [[CrossRef](#)] [[PubMed](#)]
57. Chen, Y.; Yang, Y.; Ji, X.; Zhao, R.; Li, G.; Gu, Y.; Shi, A.; Jiang, W.; Zhang, Q. The SCIFF-derived ranthipeptides participate in quorum sensing in solventogenic clostridia. *Biotechnol. J.* **2020**, *15*, e2000136. [[CrossRef](#)] [[PubMed](#)]
58. Montalbán-López, M.; Scott, T.A.; Ramesh, S.; Rahman, I.R.; van Heel, A.J.; Viel, J.H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; et al. New developments in RiPP discovery, enzymology and engineering. *Nat. Prod. Rep.* **2021**, *38*, 130–239. [[CrossRef](#)] [[PubMed](#)]
59. Kontnik, R.; Bosak, T.; Butcher, R.A.; Brocks, J.J.; Losick, R.; Clardy, J.; Pearson, A. Sporulenes, heptaprenyl metabolites from *Bacillus subtilis* spores. *Org. Lett.* **2008**, *10*, 3551–3554. [[CrossRef](#)] [[PubMed](#)]
60. Netotea, S.; Bertani, I.; Steindler, L.; Kerényi, A.; Venturi, V.; Pongor, S. A simple model for the early events of quorum sensing in *Pseudomonas aeruginosa*: Modeling bacterial swarming as the movement of an “activation zone”. *Biol. Direct.* **2009**, *4*, 6. [[CrossRef](#)]
61. Seshadri, R.; Roux, S.; Huber, K.J.; Wu, D.; Yu, S.; Udwary, D.; Call, L.; Nayfach, S.; Hahnke, R.L.; Pukall, R.; et al. Expanding the genomic encyclopedia of Actinobacteria with 824 isolate reference genomes. *Cell Genom.* **2022**, *2*, 100213. [[CrossRef](#)] [[PubMed](#)]
62. Wang, F.; Li, M.; Huang, L.; Zhang, X.H. Cultivation of uncultured marine microorganisms. *Mar. Life Sci. Technol.* **2021**, *3*, 117–120. [[CrossRef](#)] [[PubMed](#)]
63. Joint, I.; Mühlung, M.; Querellou, J. Culturing marine bacteria—An essential prerequisite for biodiscovery. *Microb. Biotechnol.* **2010**, *3*, 564–575. [[CrossRef](#)] [[PubMed](#)]
64. Kallscheuer, N.; Jogler, C. The bacterial phylum Planctomycetes as novel source for bioactive small molecules. *Biotechnol. Adv.* **2021**, *53*, 107818. [[CrossRef](#)] [[PubMed](#)]
65. Oren, A.; Garrity, G.M. Valid publication of the names of forty-two phyla of prokaryotes. *Int. J. Syst. Evol. Microbiol.* **2021**, *71*, 005056. [[CrossRef](#)] [[PubMed](#)]
66. Göker, M.; Oren, A. Valid publication of names of two domains and seven kingdoms of prokaryotes. *Int. J. Syst. Evol. Microbiol.* **2024**, *74*, 006242. [[CrossRef](#)] [[PubMed](#)]
67. Strittmatter, A.W.; Liesegang, H.; Rabus, R.; Decker, I.; Amann, J.; Andres, S.; Henne, A.; Fricke, W.F.; Martinez-Arias, R.; Bartels, D.; et al. Genome sequence of *Desulfobacterium autotrophicum* HRM2, a marine sulfate reducer oxidizing organic carbon completely to carbon dioxide. *Environ. Microbiol.* **2009**, *11*, 1038–1055. [[CrossRef](#)] [[PubMed](#)]
68. Postgate, J.R.; Kent, H.M.; Robson, R.L. DNA from diazotrophic *Desulfovibrio* strains is homologous to *Klebsiella pneumoniae* structural nif DNA and can be chromosomal or plasmid-borne. *FEMS Microbiol. Lett.* **1986**, *33*, 159–163. [[CrossRef](#)]
69. Argyle, J.L.; Rapp-Giles, B.J.; Wall, J.D. Plasmid transfer by conjugation in *Desulfovibrio desulfuricans*. *FEMS Microbiol. Lett.* **1992**, *94*, 255–262. [[CrossRef](#)]
70. Powell, B.; Mergeay, M.; Christofi, N. Transfer of broad host-range plasmids to sulphate-reducing bacteria. *FEMS Microbiol. Lett.* **1989**, *59*, 269–274. [[CrossRef](#)]
71. Rabus, R.; Hansen, T.A.; Widdel, F. Dissimilatory Sulfate- and sulfur-reducing prokaryotes. In *The Prokaryotes*; Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2013.
72. Rodríguez-Beltrán, J.; DelaFuente, J.; León-Sampedro, R.; MacLean, R.C.; San Millán, Á. Beyond horizontal gene transfer: The role of plasmids in bacterial evolution. *Nat. Rev. Microbiol.* **2021**, *19*, 347–359. [[CrossRef](#)] [[PubMed](#)]
73. Saati-Santamaría, Z. Global map of specialized metabolites encoded in prokaryotic plasmids. *Microbiol. Spectr.* **2023**, *11*, e0152323. [[CrossRef](#)] [[PubMed](#)]
74. Freitas, S.; Hatosy, S.; Fuhrman, J.A.; Huse, S.M.; Welch, D.B.; Sogin, M.L.; Martiny, A.C. Global distribution and diversity of marine Verrucomicrobia. *ISME J.* **2012**, *6*, 1499–1505. [[CrossRef](#)] [[PubMed](#)]
75. Flombaum, P.; Gallegos, J.L.; Gordillo, R.A.; Rincón, J.; Zabala, L.L.; Jiao, N.; Karl, D.M.; Li, W.K.; Lomas, M.W.; Veneziano, D.; et al. Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 9824–9829. [[CrossRef](#)] [[PubMed](#)]
76. Solano, G.; Rojas-Jiménez, K.; Jaspars, M.; Tamayo-Castillo, G. Study of the diversity of culturable actinomycetes in the North Pacific and Caribbean coasts of Costa Rica. *Antonie Leeuwen.* **2009**, *96*, 71–78. [[CrossRef](#)] [[PubMed](#)]
77. Zhou, Z.; Tran, P.Q.; Kieft, K.; Anantharaman, K. Genome diversification in globally distributed novel marine Proteobacteria is linked to environmental adaptation. *ISME J.* **2020**, *14*, 2060–2077. [[CrossRef](#)] [[PubMed](#)]

78. Dahlberg, C.; Bergström, M.; Andreasen, M.; Christensen, B.B.; Molin, S.; Hermansson, M. Interspecies bacterial conjugation by plasmids from marine environments visualized by gfp expression. *Mol. Biol. Evol.* **1998**, *15*, 385–390. [[CrossRef](#)]
79. Petersen, J.; Vollmers, J.; Ringel, V.; Brinkmann, H.; Ellebrandt-Sperling, C.; Spröer, C.; Howat, A.M.; Murrell, J.C.; Kaster, A.K. A marine plasmid hitchhiking vast phylogenetic and geographic distances. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 20568–20573. [[CrossRef](#)] [[PubMed](#)]
80. Cuadrat, R.R.C.; Sorokina, M.; Andrade, B.G.; Goris, T.; Dávila, A.M.R. Global ocean resistome revealed: Exploring antibiotic resistance gene abundance and distribution in TARA Oceans samples. *Gigascience* **2020**, *9*,giaa046. [[CrossRef](#)] [[PubMed](#)]
81. Ivanova, A.A.; Naumoff, D.G.; Miroshnikov, K.K.; Liesack, W.; Dedysh, S.N. Comparative Genomics of Four Isosphaeraceae Planctomycetes: A common pool of plasmids and glycoside hydrolase genes shared by *Paludisphaera borealis* PX4T, *Isosphaera pallida* IS1BT, *Singulisphaera acidiphila* DSM 18658T, and strain SH-PL62. *Front Microbiol.* **2017**, *8*, 412. [[CrossRef](#)]
82. Salbreiter, M.; Waqqas, M.; Jogler, M.; Kallscheuer, N.; Wiegand, S.; Peeters, S.H.; Heuer, A.; Jetten, M.S.M.; Boedeker, C.; Rast, P.; et al. Three Planctomycetes isolated from biotic surfaces in the Mediterranean Sea and the Pacific Ocean constitute the novel species *Symmachiella dynata* gen. nov., sp. nov. and *Symmachiella macrocystis* sp. nov. *Antonie Leeuwen.* **2020**, *113*, 1965–1977. [[CrossRef](#)]
83. Mara, P.; Vik, D.; Pachiadaki, M.G.; Suter, E.A.; Poulos, B.; Taylor, G.T.; Sullivan, M.B.; Edgcomb, V.P. Viral elements and their potential influence on microbial processes along the permanently stratified Cariaco Basin redoxcline. *ISME J.* **2020**, *14*, 3079–3092. [[CrossRef](#)] [[PubMed](#)]
84. Jurgensen, S.K.; Roux, S.; Schwenck, S.M.; Stewart, F.J.; Sullivan, M.B.; Brum, J.R. Viral community analysis in a marine oxygen minimum zone indicates increased potential for viral manipulation of microbial physiological state. *ISME J.* **2022**, *16*, 972–982. [[CrossRef](#)] [[PubMed](#)]
85. Vik, D.; Gazitúa, M.C.; Sun, C.L.; Zayed, A.A.; Aldunate, M.; Mulholland, M.R.; Ulloa, O.; Sullivan, M.B. Genome-resolved viral ecology in a marine oxygen minimum zone. *Environ. Microbiol.* **2021**, *23*, 2858–2874. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.