Pan-Genome Provides Insights into *Vibrio* Evolution and Adaptation to Deep-Sea Hydrothermal Vents

Emanuele Bosi [1,2,*], Elisa Taviani[1,2], Alessia Avesani[1], Lapo Doni[1,2], Manon Auguste[1,2], Caterina Oliveri[1], Martina Leonessi[1,2], Jaime Martinez-Urtaza[3], Costantino Vetriani[4,5], Luigi Vezzulli[1,2]

[1]Department of Earth, Environmental and Life Sciences (DISTAV), University of Genoa, Genoa 16132, Italy

[2]National Biodiversity Future Center, Palermo, Italy

[3]Facultat de Biociéncies, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona (UAB), Bellaterra, Barcelona 08193, Spain

[4]Department of Marine and Coastal Sciences, Rutgers University, New Brunswick, NJ 08901, USA

[5]Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ 08901, USA

*Corresponding author: Email: emanuele.bosi@unige.it

Accepted: June 01, 2024

Abstract

This study delves into the genomic features of 10 *Vibrio* strains collected from deep-sea hydrothermal vents in the Pacific Ocean, providing insights into their evolutionary history and ecological adaptations. Through sequencing and pan-genome analysis involving 141 *Vibrio* species, we found that deep-sea strains exhibit larger genomes with unique gene distributions, suggesting adaptation to the vent environment. The phylogenomic reconstruction of the investigated isolates revealed the presence of 2 main clades: The first is monophyletic, consisting exclusively of *Vibrio alginolyticus*, while the second forms a monophyletic clade comprising both *Vibrio antiquarius* and *Vibrio diabolicus* species, which were previously isolated from deep-sea vents. All strains carry virulence and antibiotic resistance genes related to those found in human pathogenic *Vibrio* species which may play a wider ecological role other than host infection in these environments. In addition, functional genomic analysis identified genes potentially related to deep-sea survival and stress response, alongside candidate genes encoding for novel antimicrobial agents. Ultimately, the pan-genome we generated represents a valuable resource for future studies investigating the taxonomy, evolution, and ecology of *Vibrio* species.

Key words: *Vibrio*, deep-sea hydrothermal vents, pan-genome, antibiotic resistance genes, virulence factors, comparative genomics.

Significance

Deep-sea hydrothermal vents are remote ocean environments harboring rich bacterial communities which may show peculiar environmental adaptations and evolutionary history. Using pan-genome analysis this study examines 10 Vibrio strains isolated from deep-sea vents in the Pacific Ocean, revealing their phylogenetic placement and ecological adaptations to these extreme environments. Notably, these isolates share virulence and antibiotic resistance genes with coastal Vibrio pathogenic for humans and aquatic animals. By shedding light on deep-sea *Vibrio* strains' genetic and ecological features, our work provides new knowledge relevant to understanding the ecology, evolution, and roots of antimicrobial resistance in marine microbial pathogens.

Introduction

Despite its harsh conditions, i.e. lack of light, low temperatures, and elevated pressure, the deep-sea environment represents one of the largest Earth's biomes (Bar-On et al. 2018), harboring a vast number of microbes whose metabolic capabilities play a significant role in biogeochemical cycles (Rousk and Bengtson 2014). Among the different deep-sea habitats, hydrothermal vents are remarkably heterogeneous. These sites, in which seawater percolates through the ocean floor into the subsurface, are characterized by chimneys propelling hot, anoxic hydrothermal fluids into the oxic, cold seawater. The combination of these 2 fluxes and the release of different chemicals, shape the biodiversity of the vents and the surrounding area. The different thermic, oxic, and chemical gradients favor the growth of variegate microbial communities enriched in chemolithoautotrophs (Reysenbach et al. 2000; Sievert and Vetriani 2012; Nakamura and Takai 2014), either free living or involved in symbiotic association with vent animals (Van Dover 2021).

Of the different bacteria that can be found at deep-sea vents, members of the *Vibrio* genus are among those who can adapt to a wider range of conditions, as they are ubiquitous in different environments, including seawater, sediments, and associated with marine organisms (Doni et al. 2023 a; Doni et al. (2023)). In particular, piezo-tolerant *Vibrio* strains have been previously isolated from deep-sea environments (Raguénès et al. 1997; Hasan et al. 2015), including hydrothermal vents (Hasan et al. 2015). Genomic analysis of such strains can provide insights into the molecular basis of their adaptation to such extreme conditions: the genome of *V. antiquarius* highlighted the presence of genes, such as zinc-dependent carboxypeptidases and alkyl hydroperoxide reductase, which can be useful when facing high concentrations of oxygen and heavy metals (Karl et al. 1988). Quite surprisingly, despite living in what can be considered a remote environment, *V. antiquarius* possess virulence genes commonly found in pathogenic *Vibrio* species, raising questions regarding the relationship between *Vibrio* species distribution, their lifestyle, and the evolutionary forces selecting for the genomic determinants of pathogenesis (Hasan et al. 2015).

There has been growing interest regarding the isolation of novel strains from remote environments, not only for the evolutionary implications that might be derived from the study of their genomes but also due to their possible biosynthetic potential. The identification of novel molecules from pristine environments, i.e. bioprospecting (Blicharska et al. 2019), represents a promising source of new therapeutics and antimicrobials (Michalska et al. 2015; Stokke et al. 2020), which are more than needed as the burden of antibiotic resistance is increasingly becoming a global health-care emergency (Antimicrobial Resistance Collaborators 2022). Previous analyses of bacteria isolated from hydrothermal vents highlighted the presence of Antimicrobial Resistance Genes (ARGs) and resistant phenotypes (Andrianasolo et al. 2009; Turner et al. 2018), suggesting that molecules with bacteriocide or bacteriostatic activity might be effectors of the ecological interactions within the microbial communities.

In this study, we report the isolation, sequencing, and characterization of 10 *Vibrio* strains recovered from two different locations, the East Pacific Rise and the Guaymas Basin. Both sites are deep-sea hydrothermal vents, located ~2,000 km apart. The microbes have been collected from different substrates, such as rocks, biofilms, mats, and hydrothermal fluids. To define their peculiarities and traits related to deep-sea adaptation, we compared them with the representative genomes of the other *Vibrio* species, defining a pan-genome for the whole genus which allowed us to taxonomically classify the isolates and to better highlight their unique functional properties, as well as features that could be associated to phenotypes of interest, such as antimicrobial resistance and virulence. The virulence of these strains was also tested using bivalve models, hinting at their potential as marine microbial pathogens.

Materials and Methods

Vibrio Strains Isolation and Sequencing

Samples of microbial biofilms were collected during cruises AT 15-15 (East Pacific Rise, 9N; January–February 2007) and AT 15-25A (Guaymas Basin, October 2007) on the Research Vessel *Atlantis* using the Deep-Submergence Vehicle *Alvin*. Bacteria used in this study were isolated from (i) biofilms collected on experimental microbial colonization devices made of stainless steel mesh or scraped from basalt rocks; (ii) sediment cores; (iii) *Beggiatoa*-dominated microbial mats; and (iv) hydrothermal fluids (Table 1). Enrichment cultures were carried out in liquid Medium 142 (https://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium142.pdf) from the German culture collection (DSMZ), modified by the elimination of acetate (142-A), or in liquid low-strength artificial seawater medium (LS ASW), containing (per liter): 24 g NaCl, 0.7 g KCl, 7.0 g, MgCl2, 0.15 g yeast extract, and 0.125 g peptone. Pure cultures were obtained by 3 consecutive isolations of single colonies on either medium 142-A or LS ASW solidified with 15 g of agar per liter.

Genomic DNA was extracted using the High Pure PCR Template Preparation Kit (Roche Diagnostics), following the manufacturer's protocol for genomic DNA isolation. DNA concentration and quality were determined fluorimetrically with QuantiFluorTM dsDNA System using a QuantiFluorTM fluorometer (Promega Italia srl, Milano, Italy). Sizing of genomic DNA was also conducted in an Agilent Bioanalyzer 2100 (Agilent, Palo Alto, CA, USA) using the High Sensitivity DNA kit (Agilent Technologies).

Table 1 Table of metadata and genomic features for EPRs

Genomes	Reads	Size	CDSs	Annotated	GC	Site temperature (°C)	Sample type	Coordinates	Depth	Vent site
EPR117	10691788	5136642	4596	4479	44,56	21	Biofilm on rock	9°49′N 104°17′W	2506m	East Pacific Rise 9°N
EPR119	9588392	5136458	4598	4481	44,56	27 to 45	Biofilm on rock	9°49′N 104°17′W	2500m	East Pacific Rise 9°N
EPR121	9286246	5135941	4594	4478	44,56	20 to 30	Biofilm on colonization device	9°49′N 104°17′W	2506m	East Pacific Rise 9°N
EPR123	10925950	5137111	4594	4480	44,56	2 to 13	Biofilm on colonization device	9°49′N 104°17′W	2503m	East Pacific Rise 9°N
EPR136	10539626	5135698	4594	4479	44,56	27 to 45	Biofilm on rock	9°49′N 104°17′W	2500m	East Pacific Rise 9°N
EPR218	11681554	5135759	4596	4478	44,56	8 to 20	Sediment core, surficial	27°0′N 111°24′W	2004m	Guaymas Basin, Gulf of California
EPR219	8173652	5135920	4597	4481	44,56	6	Beggiatoa-dominated microbial mat	27°0′N 111°24′W	2001m	Guaymas Basin, Gulf of California
EPR224	11815126	5136589	4598	4482	44,56	9	Sediment core with Beggiatoa-dominated microbial mat	27°0′N 111°24′W	2011m	Guaymas Basin, Gulf of California
EPR225	12841266	5232040	4727	4530	44,63	9	Sediment core with Beggiatoa-dominated microbial mat	27°0′N 111°24′W	2011m	Guaymas Basin, Gulf of California
EPR246	6920774	5135157	4597	4481	44,56	30	Hydrothermal fluids from vent	9°49′N 104°17′W	2500m	East Pacific Rise 9°N

The table reports genomic features and metadata related to isolation site for the 10 EPR strains reported in this work.

Libraries for next-generation sequencing on the Illumina platform (Illumina, Inc.) using the KAPA HyperPlus Kit for Illumina (Roche Diagnostics, Mannheim, Germany). The genomes were then sequenced using short-read Illumina technology on a MiSeq platform (2 × 150 bp).

Genome Analysis

The read quality of the sequenced genomes was assessed with FASTQC, and low-quality regions were removed with Trimmomatic v. 0.32 (Bolger et al. 2014). Assembly was performed using the CLC Genomics workbench. The representative genomes of *Vibrio* species have been obtained with ncbi-genome-download v. 0.3.3 (Blin 2023), using the following command line: "ncbi-genome-download –genera Vibrio –refseq-categories representative." Then, the assembled genomes and those downloaded from NCBI have been analyzed in the same way. First, genomes have been annotated with Prokka v. 1.14.6 (Seemann 2014), using this command line: "prokka –outdir SAMPLE/prokka/–force –prefix SAMPLE –genus Vibrio," where SAMPLE denotes the identifier associated with the genome (e.g. GCF_014050145.1 or EPR119). Then, the translations of the predicted coding sequences (i.e. the Prokka output files with the.faa suffix) have been annotated according to the EggNOG database (v. 5.0.2) using the emapper utility (v. 2.1.10) and diamond (v. 2.1.4) (Huerta-Cepas et al. 2019; Buchfink et al. 2021; Cantalapiedra et al. 2021). The annotation was a 2-step analysis, performing first a highly sensitive diamond run, followed by an annotation on the files generated with diamond. The commands used for the 2 steps are, respectively: "emapper.py -m diamond –sensmode more-sensitive –no_annot -i FAA -o OUT1 –cpu 0 –override" and "emapper.py -m no_search –annotate_hits_table OUT1.emapper.seed_orthologs -o OUT2 –dbmem –report_orthologs –override –target_orthologs one2one." Here, FAA, OUT1, and OUT2 denote the input file with the aminoacid sequences, the output prefix of the diamond search, and the output prefix of annotation files, respectively.

The presence of genes related to virulence has been assessed using Abricate (v. 1.0.1) (Seemann) to map annotated genes on the Virulence Factors Database (VFDB; Chen et al. 2016) by running the command "abricate –nopath –datadir/home/bosi/pipelord/resources/dbs/abricate –threads 30 –db vfdb GBK," using the GenBank file generated by Prokka (GBK) as input. Putative ARGs have been called with AMRFinderPlus (v. 3.11.2), as it follows: "amrfinder -n FNA -p FAA –gff GFF -o OUT –plus -d $(dirname $(which amrfinder))/data/latest -a prokka." The names FNA, FAA, and GFF indicate the Prokka output files reporting the nucleotide sequences, aminoacid sequences, and the annotation in GFF format, while OUT corresponds to the output directory. Biosynthetic clusters for secondary metabolites were called AntiSMASH (v. 7.0.0) with strictness relaxed (Blin et al. 2023).

Pan-genome Analysis

The files generated by emapper report for each protein its annotated OG, spanning different taxonomic levels. The files, each corresponding to an individual genome, have been merged into a single table with a column specifying the sample of each entry. The table has been used to compute for each different OG the relative frequency across the considered genomes, i.e. OGs with a frequency equal to 1 correspond to genes which are present in all genomes. Of note, multiple copies of an OG (paralogs) in a given genome were counted only once for frequency computation. The frequency has been used to classify OGs into 3 categories: Core genes, present in 95% or more genomes, shell genes, with frequency values between 5% and 95%, and cloud genes, present in less than 5% of genomes.

Core-genome phylogeny was performed considering single-copy genes which were present in all genomes. The aminoacid translations of these genes have been separately aligned using Muscle (Edgar 2004) as implemented in Biopython, using the parameters "clwstrict" and "diags." The multiple sequence alignments (MSAs) generated in this way have been concatenated into a single MSA fasta file, with one concatamer for each genome. The concatamer has been used to compute a phylogeny with iqtree2 (Minh et al. 2020) with the parameters –alrt 1,000 -B 1000, resulting in a maximum-likelihood phylogeny with 1,000 bootstraps. Model selection, as implemented in iqtree2, was performed with ModelFinder (Kalyaanamoorthy et al. 2017) by testing different models and selecting the best fit according to Bayesian Information Criteria minimization. Phylogenetic tree visualization and annotation was performed with ITOL (Letunic and Bork 2021). Pairwise ANI was computed with fastANI (v. 1.33 [Jain et al. 2018]).

Phylogenetic Analysis of the Clade *Vibrio alginolyticus*

To improve the phylogenetic resolution within the clades alginolyticus and diabolicus, the genomes of the 10 sequenced strains were used, together with those of the species *V. antiquarius*, *V. diabolicus*, and *V. chemagurensis* (genome identifiers: GCF_000024825.1, GCF_024748065.1, and GCF_012275705.1), as well as those of *V. alginolyticus* strains with a complete genome, to identify conserved polymorphic SNP sites. To accomplish this we applied Snippy (https://github.com/tseemann/snippy), using the representative genome of the species *V. alginolyticus* (GCF_023650915.1) as a reference to identify the variants, and the snippy-core utility to extract universally shared variant sites. The resulting alignment has been used to derive a phylogeny with iqtree2 as above.

Interaction With Bivalves

EPR strains were cultured in Zobell medium at 23 °C, respectively, under static conditions. After overnight growth, cells were harvested by centrifugation (4500 g, 10 min), washed with ASW (salinity 35 ppt) and resuspended to an A 600 = 1.2 (about 10^8 CFU/mL). TCBS agar (Scharlau Microbiology, Italy) was used for culturing both strains.

Mussels (*Mytilus galloprovincialis*, La Spezia, Italy) and Oysters (*Crassostrea gigas*, Bretagne, France) were purchased in autumn 2023 and acclimated for 24 h in static tanks containing aerated ASW, (35 ppt salinity for mussel and 31 ppt for oysters; 1 L/animal) at 18 °C. Bactericidal activity and LMS were determined in bivalve hemocyte monolayers as previously described (Lasa et al. 2021; Auguste et al. 2022).

Results

Sequencing and Assembly of *Vibrio* spp. From Deep-Sea Vents

Sampling expeditions with deep-sea submersibles enabled the collection of 10 different *Vibrio* isolates from various sites at 2 deep-sea hydrothermal vent locations. These ten isolates represent the entirety of *Vibrio* strains collected from these sites, and most of them, detailed in Table 1, are from the East Pacific Rise (EPR). Consequently, they will hereafter be referred to collectively as EPR strains. The isolates are mesophiles, being able to grow on media such as Thiosulfate Citrate Bile Sucrose (TCBS), similar to other *Vibrio* strains isolated from nonextreme environments. Genome sequencing and assembly delivered 10 draft genomes, whose general features are reported in Table 1. To better understand these genomes in the context of the *Vibrio* genus, we built a genomic dataset with a representative genome for each *Vibrio* species. This dataset, including a total of 141 species, represents—to the best of our knowledge—the most comprehensive and diverse dataset for the *Vibrio* genus, providing a view of the genomic variability of this taxon (supplementary table S1, Supplementary Material online). Considering the average genome size (Fig. 1A), EPRs display larger genomes (5.1 Mbp) than most other vibrios (4.7 Mbp), which is reflected also in the average number of genes (4,609 vs. 4,220). Both differences are statistically significant ($P < 0.05$). From a compositional perspective, the average GC content of EPRs is in line with that of the genus (46.4% vs. 46.3%).

The Pan-Genome of the Genus *Vibrio*

To gain insights into the taxonomic affiliation of EPRs and their evolutionary relationships with other vibrios, the *Vibrio* dataset has been used, along with the 10 EPR genomes, to define the pan-genome of the genus. Briefly, EggNOG was used to cluster coding sequences (CDS) in Orthologous Groups (OGs). Each OG was classified as core, shell, or cloud according to its frequency across *Vibrio* genomes, such that core OGs correspond to genes found in more than 95% of *Vibrio* genomes, shell genes are present in a number of

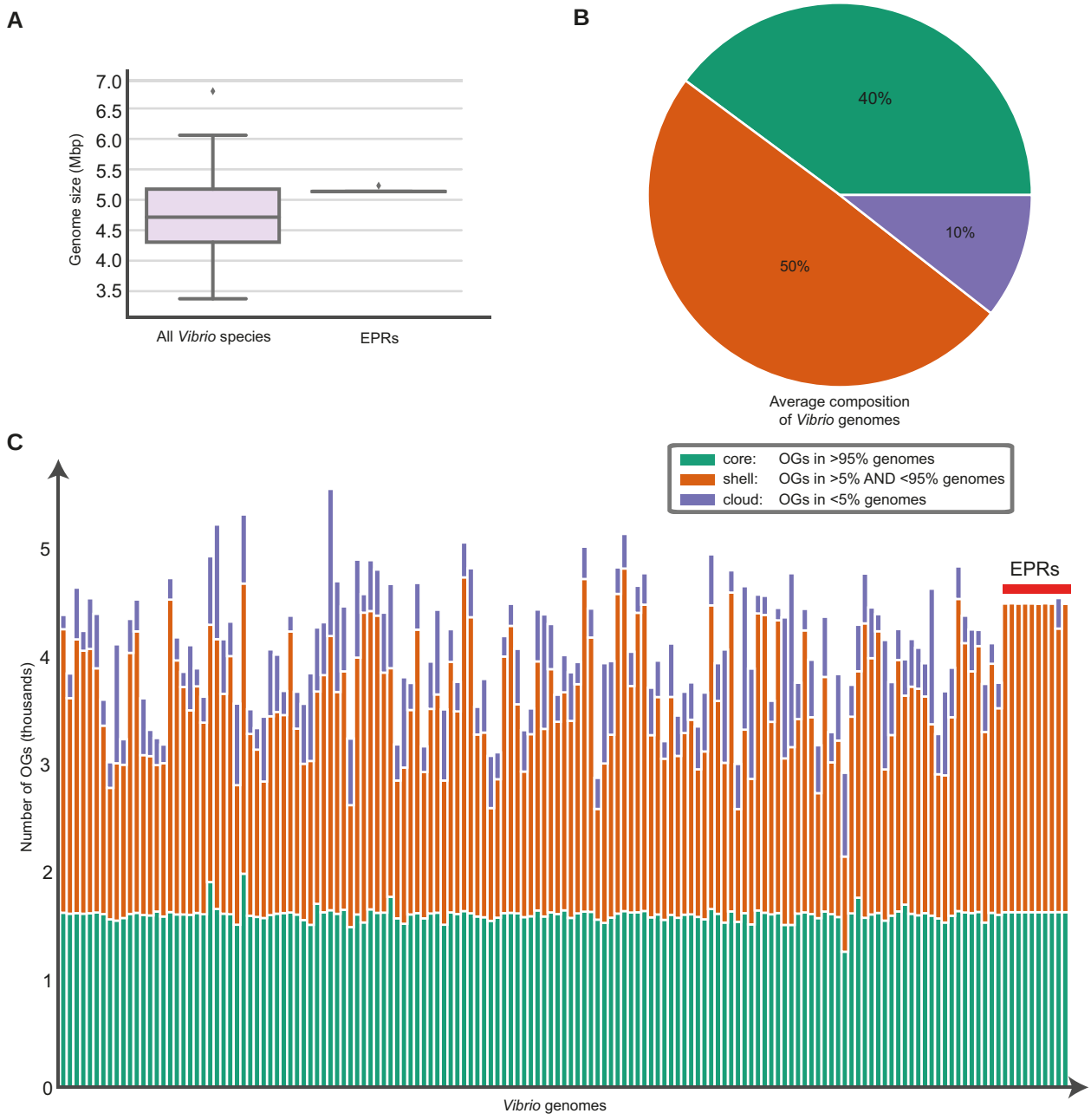

Fig. 1. The pan-genome of the *Vibrio* genus: A) The boxplot reports the distribution of genome size of *Vibrio* representatives compared with EPRs. B) The proportion of core, shell, and cloud genes in the average *Vibrio* genome. C) The number of OGs for each pan-genome category is reported for each genome.

genomes comprised of 5% and 95%, and cloud genes are identified in <5% of *Vibrio* genomes.

Considering all 151 genomes, there are 41,215 OGs, divided into 32,403 clouds, 7,235 shells, and 1,577 core genes (supplementary table S2, Supplementary Material online). Considering the proportion of these categories in each genome, on average there are 40% core, 50% shell, and 10% cloud genes (Fig. 1B). The EPRs display a significantly ($P < 0.05$) different proportion with 36% of core, 63% of shell, and 1% of cloud genes. Such a low number of cloud genes are not observed in the other species (Fig. 1C). Of note, since the EPR group is larger ($N = 10$) than the frequency threshold to define cloud genes (lower than 5% or 7.55 genomes), shell genes can include genes present only in EPR genomes. Although the low number of cloud genes alone provides a strong indication that EPRs are very similar in gene content, we analyzed the distribution of OGs within this group to have a more comprehensive

picture. From a total of 4,823 OGs, 3,799 (79%) are present in 10/10 genomes, 490 (11%) are present in all genomes but EPR225 and 534 (10%) are present in a lower number of genomes, most of which (530) are specific to EPR225.

Phylogenomics of *Vibrio*

The 1,577 core genes included OGs which were present in multiple copies in a single genome (i.e. paralogs), and/or genes which were not found in a few strains, either due to recent gene loss events or fragmented genome assembly. Excluding genes falling in these categories, we obtained a total of 380 single-copy genes present in all genomes, which were aligned, concatenated, and used to build a maximum-likelihood phylogeny (Fig. 2A). The concatamer embeds 82,197 sites, of which 35,869 constant, 8,307 singletons, and 38,201 parsimony-informative. The phylogenetic tree was constructed using the "Q.yeast + I + G + R10" model, identified as the best-fitting through model selection, and supported by 1,000 pseudo-bootstrap replications, shows how 9/10 EPRs fall in a distinct clade, close to the species *V. alginolyticus*, whereas the other strain (EPR225) was placed in a clade with the species *V. chemaguriensis*, *V. diabolicus*, and *V. antiquarius*. These clades, which will be referred to as clade A and D, respectively, fall within the *V. harveyi* clade, well-known for including pathogenic *Vibrio* species such as *V. vulnificus*, *V. parahemolyticus*, and *V. alginolyticus*.

The phylogenetic analysis strongly implies that 9 EPRs belong to the species *V. alginolyticus*. The clade *diabolicus*, comprising 3 different species in addition to EPR225, lies close to the *alginolyticus* clade, consistent with a previous comparative analysis (Turner et al. 2018) on the species *V. alginolyticus*, *V. diabolicus*, and *V. antiquarius*, which highlighted a substantial synonymy between representatives of these taxa. To better resolve the phylogenetic relationships within the clade comprising *alginolyticus* and *diabolicus* (clade AD), we assembled a new dataset in which we included all representatives of *V. alginolyticus* with a complete genome available ($N = 46$), as well as those from the clade AD ($N = 14$), using the genomic sequences to identify single nucleotide polymorphisms (SNPs). There were a total of 379,566 conserved SNP variant sites which were aligned and used to perform a phylogenetic analysis. In parallel, the genomes were compared pairwise to compute their Average Nucleotide Identity (ANI) (supplementary table S3, Supplementary Material online).

The results of ANI and SNP phylogeny (Fig. 2B and C) are consistent, in that both highlight a clear separation between a clade which includes the members of the *diabolicus* clade mixed with five *V. alginolyticus* strains. In the phylogenetic tree, the strain EPR225 has been placed together with *V. alginolyticus* LHF01. The other EPRs cluster together in a distinct clade, clearly separated from the representative strain that was included in clade A. The information provided by this tree strongly implies the presence of two distinct species within this dataset. Considering the ANI results, EPR225 displays a higher similarity (ANI > 97%) with the other members of the mixed clade, with respect to the other strains (> 92%), whereas the other EPRs are extremely similar to each other (> 99%) and highly similar to the other *V. alginolyticus* strains (> 98%).

Functional Genomics

Annotation of functional elements in EPR genomes may provide important insights regarding phenotypic traits of interest, such as adaptation to deep-sea conditions or production of secondary metabolites. Considering the similarities of EPRs with pathogenic vibrios, their genomes were scanned also for genes encoding putative virulence factors (VFs) and ARGs. Moreover, the annotation has been extended to the whole pan-genome dataset to better highlight peculiar features of EPRs in relation to other representatives of the genus.

Distribution of ARGs

Considering the whole pan-genome, there were 381 ARGs, divided into 52 OGs, identified in 138 different genomes (supplementary table S4, Supplementary Material online). The distribution of such elements, averaging at 2.5 ARGs per genome, is uneven, with most vibrios having few ARGs, if any, and with few strains harboring a large number of genes (Fig. 3A). For instance, the genome with the most ARGs is *V. hepatarius*, with 21 genes, followed by *V. gangliei* with 11. Compared with other vibrios, EPRs possess a relatively high number of ARGs, with all strains having 5 ARGs, except EPR225 with 6. The genes are highly conserved among EPRs: all strains have *ampC* and *blaCARB*, encoding beta-lactamases, *tet(35)* and *tet(34)*, conferring resistance to tetracycline, and catC, encoding a chloramphenicol acetyltransferase; the gene *fos* is present only in EPR225. Analysis of the gene distribution within the pan-genome, as depicted in Fig. 3B, reveals distinct patterns of ARGs among the EPR strains: The gene *tet(34)*, encoding a tetracycline inactivation enzyme, is widely distributed across the genus, the functionally associated gene *tet(35)* is mostly present in the clade *harveyi*. As for the other genes, i.e. the ampC and blaCARB, encoding beta-lactamases, and catC, conferring resistance to chloramphenicol, they are highly specific to the clades A and D.

Annotation of VFs

In the pan-genome dataset we identified a total of 1,479 VFs in 134 genomes (supplementary table S5, Supplementary Material online). The distribution of these elements, as for ARGs, is uneven, with most of the species possessing few VFs or none, and with a few species having more than a hundred VFs, e.g. *V. cholerae* with 134.

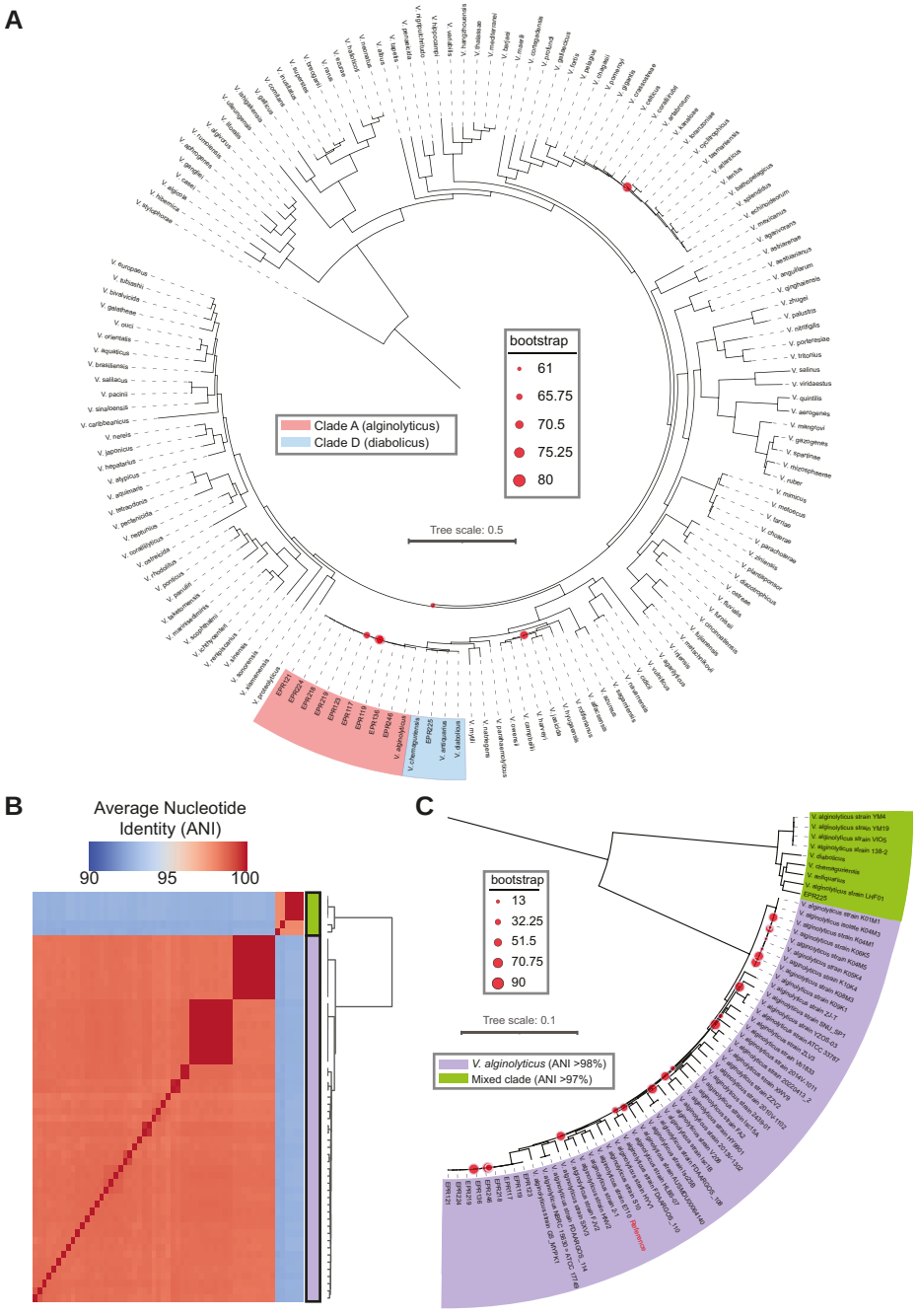

Fig. 2. Phylogenomics of *Vibrio*: A) The phylogenetic analysis was conducted using a concatenated dataset of 380 single-copy genes, supported by 1,000 bootstrap replications. Nodes on the tree with bootstrap support values ≤ 80% are specifically noted. Clades containing EPRs are colored. B) Matrix reporting pairwise ANI values obtained for a dataset including members of the clade A, D, and all complete *V. alginolyticus* strains available. Colored bars denote clusters of genomes with ANI ≥ 95%. C) The phylogenetic analysis was conducted with core SNPs on the same dataset of panel B, with 1,000 bootstrap replicates. Support values ≤ 90% are reported. Clades are colored according to the ANI results.

Compared to the average number of VFs (9.8), EPRs have a higher number of VFs, i.e. 18 for EPR225 and 12 for the other strains (Fig. 3C). From a functional point of view, we identified 3 main categories present in EPRs: Type 3 Secretion System 1 (T3SS1), flagella and chemotaxis. The T3SS1 VFs form an operon which encodes a molecular system previously characterized in pathogens such as *Salmonella enterica* and *V. parahaemolyticus*. These genes

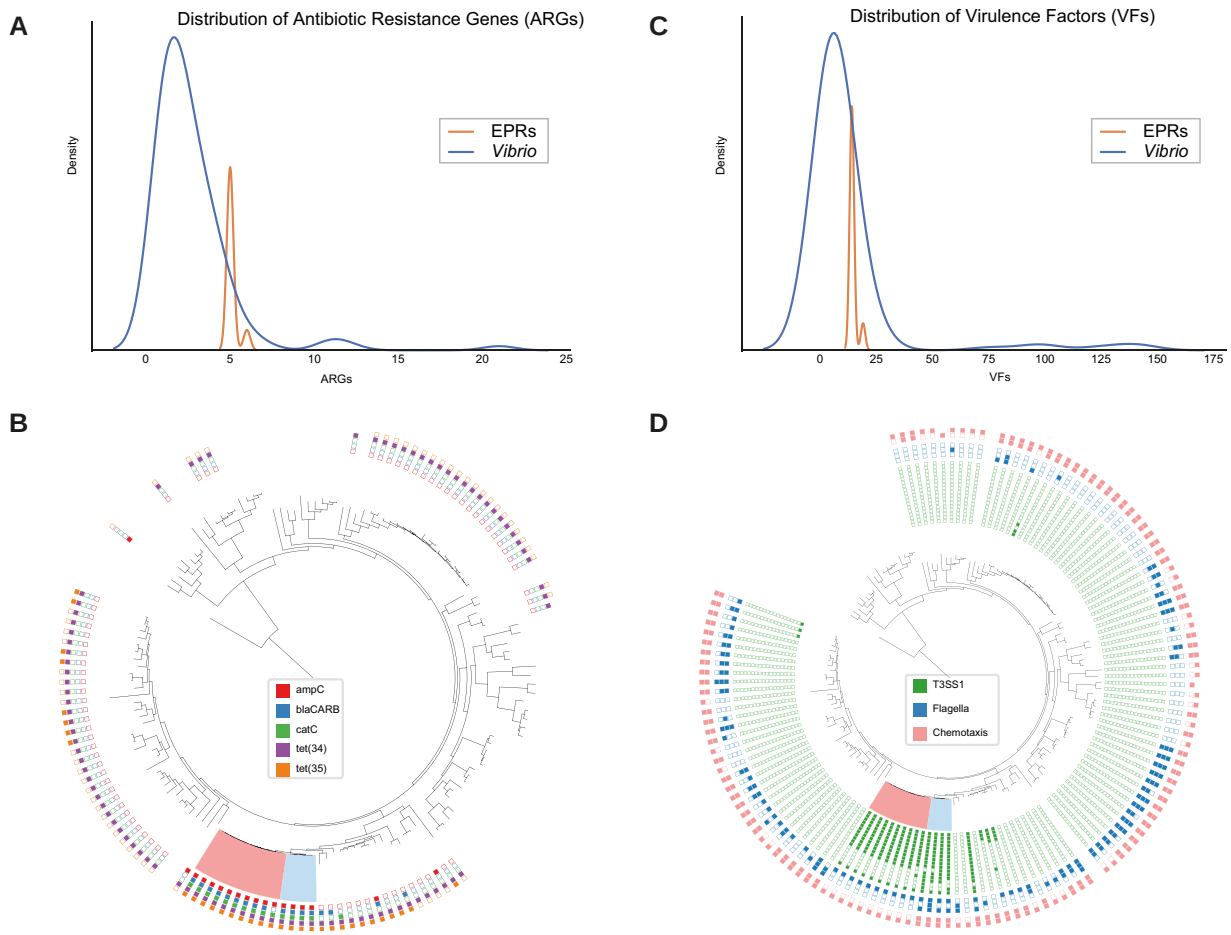

Fig. 3. Distribution of ARGs and VFs in *Vibrio*. A) Distribution of ARGs in EPRs compared with the rest of *Vibrio*. B) Phyletic pattern of ARGs found in EPRs within the *Vibrio* phylogeny. C) Distribution of VFs in EPRs compared with the rest of *Vibrio*. D) Phyletic pattern of VFs found in EPRs within the *Vibrio* phylogeny.

are almost absent in *Vibrio*, as they are highly specific to the clades A and D (Fig. 3D), in addition to the species *V. parahaemolyticus* which possesses an additional copy of a similar system (T3SS2). Although many *Vibrio* species harbor multiple VFs for the synthesis of flagellar proteins, EPR225 possesses 3 of such VFs, which are flgB, flgC, and fliG, while other EPRs have only flgB. Chemotaxis is a trait widely associated with *Vibrio*, and 2 VFs involved in this process, i.e. cheY and cheW, are widely distributed in a large clade (131 nodes) within the genus (Fig. 3D), with the former being present in almost all genomes (127/131), and the latter being less prevalent (89/131). Notably, the chemotaxis gene cheA is almost absent in *Vibrio*, being present in 4 species. Within the EPRs, clade D presents both cheY and cheW, while genomes from clade A only have cheY.

Production of Secondary Metabolites

The *Vibrio* genomes were probed to annotate biosynthetic clusters for secondary metabolites (supplementary table S6, Supplementary Material online). On average, vibrios contain 4.29 secondary metabolite clusters. All EPR genomes contain 4 biosynthetic clusters for producing the following molecule classes: ectoine, beta-lactone, arylpolyene, and NI-siderophore. These gene clusters are highly conserved, with EPR225 showing a slight variation in the cluster composition with respect to other EPRs (Fig. 4). The synthesis of such molecules is a feature shared with other *Vibrio* representatives: clusters associated with the production of beta-lactones are the most common, being found in a large number of species (130/151 genomes), followed by ectoine (103/151), while clusters for arylpolyene and NI-siderophore are less prevalent (81/151 and 55/151).

Genes Related to Extreme Environmental Conditions

Previous analyses on *Vibrio* strains isolated in deep-sea hydrothermal vents (*V. antiquarius*) or warm bathypelagic deep waters (*V. bathopelagicus*) highlighted genes possibly advantageous in these particular environments. By

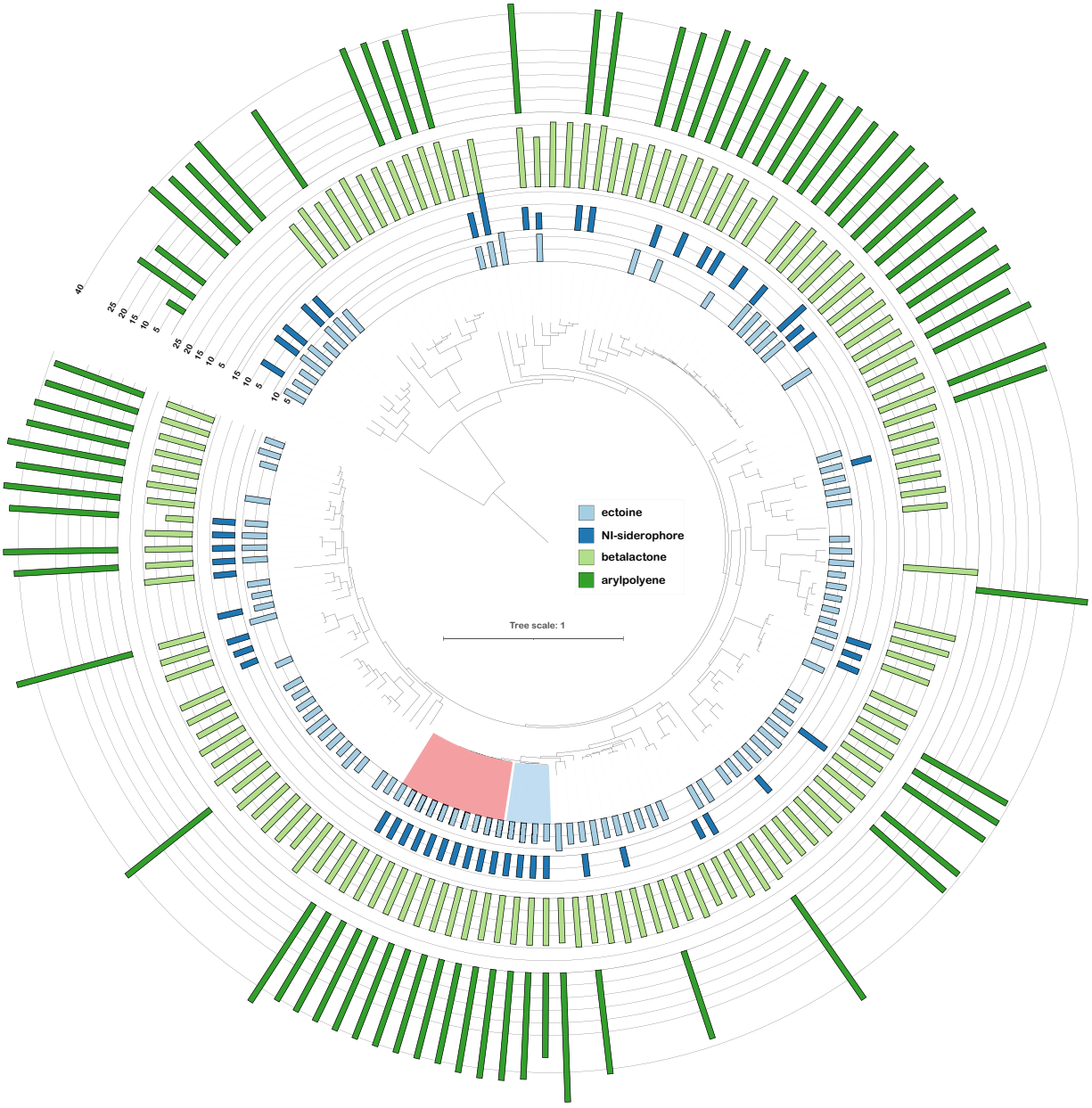

Fig. 4. Distribution of biosynthetic clusters found in EPRs in the *Vibrio* phylogeny. The number of genes associated with each of the 4 classes of secondary metabolites (ectoine, beta-lactone, arylpolyene, and NI-siderophore) are reported across the *Vibrio* tree.

searching these sequences in the pan-genome, we assessed that they can be found not only in EPRs but also in several other *Vibrio* species (supplementary fig. S1, Supplementary Material online). To find genes possibly associated with deep-sea vent lifestyle we performed an enrichment analysis, leading to the identification of 147 genes significantly (FDR $\leq$ 0.05) over-represented in deep-sea genomes (supplementary fig. S2, Supplementary Material online), which have been sorted according to their enrichment score (supplementary table S7, Supplementary Material online). The 3 genes with the highest enrichment in deep-sea vibrios are: (i) luxO, encoding a response regulator of the two-component family of signal transduction molecules involved with quorum sensing, luminescence and biofilm formation (Freeman and Bassler 1999), (ii) an unnamed hypothetical protein, and (iii) rpoE, encoding a sigma factor of the 70 families (Lonetto et al. 1992). Focusing on the functional roles of the top 60 genes, mostly found within the clades A and D (Fig. 5A), revealed that their main KEGG categories are "metabolic pathways," followed by "Valine, leucine and isoleucine degradation" and "Two-component system" (Fig. 5B). Considering the COG

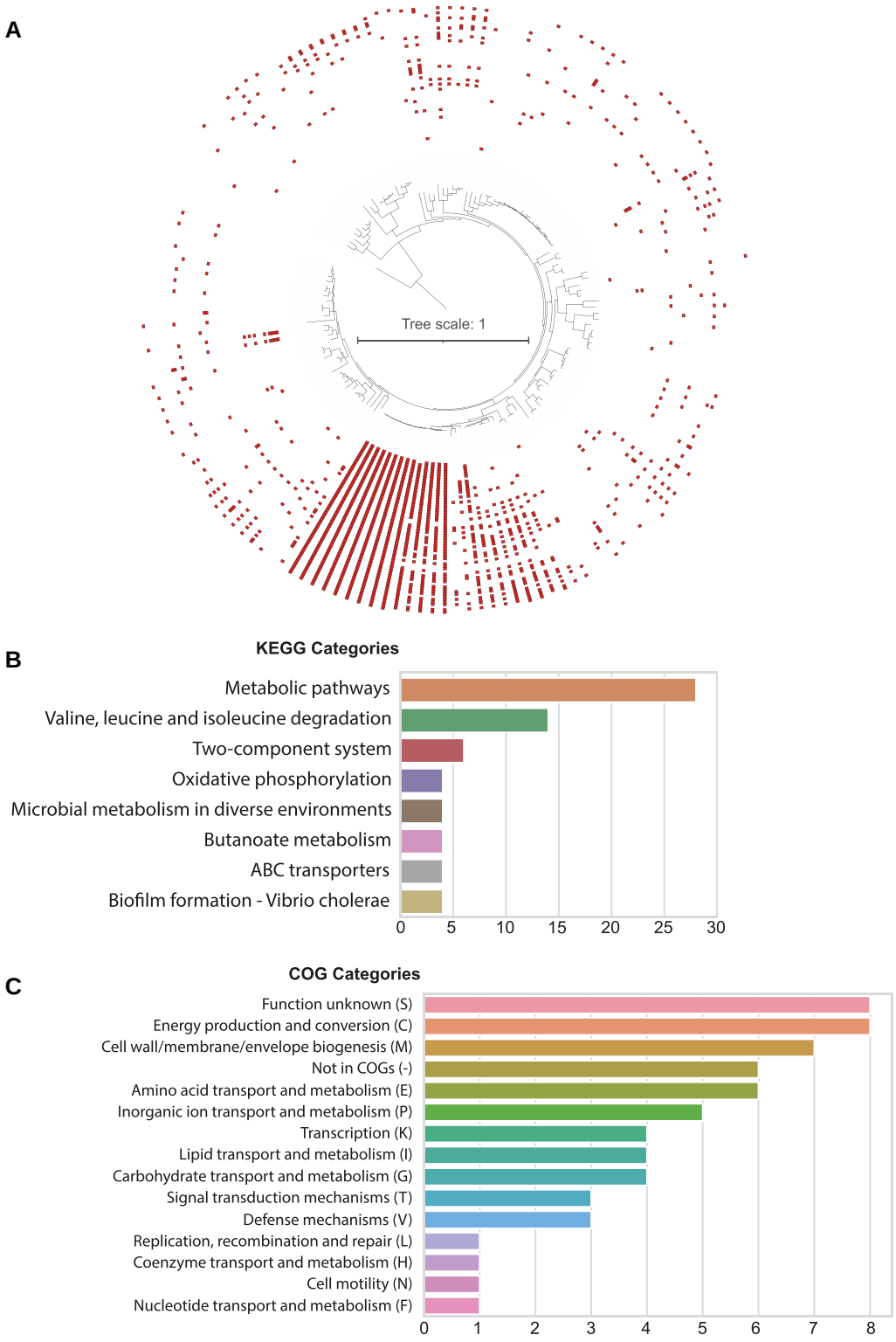

Fig. 5. Top 60 genes enriched in deep-sea hydrothermal vents Vibrios: A) Phyletic pattern of the 60 genes most significantly enriched in the genomes of 12 *Vibrio* isolates from deep-sea hydrothermal vents. B) Distribution of KEGG terms from the 60 most enriched genes. C) Distribution of GO terms from the 60 most enriched genes.

ontology, the most represented categories are "Function unknown (S)," "Energy production and conversion (C)," and "Cell wall/membrane/envelope biogenesis (M)" (Fig. 5C).

Since phylogenetic signals might confound the enrichment analysis, we also highlighted genes potentially associated with deep-sea lifestyles using a more stringent evolutionary method based on Pagel's model (Pagel 1994). This approach identified a total of 17 genes significantly (FDR ≤ 0.05) enriched in deep-sea genomes (supplementary table S8, Supplementary Material online) that have been ranked according to the enrichment score. The most enriched genes include an orthologue of ParB, a transcription factor involved in chromosome partitioning (Ireton et al. 1994), a gene encoding a trypsin-like serine protease, and a gene with a WYL domain, possibly involved in the regulation of CRISPR–Cas adaptive immunity (Makarova et al. 2014).

In Vitro Models of EPRs Interactions With Bivalves

The bactericidal activity of mussel and oyster hemocyte monolayers was evaluated after in vitro incubation of EPRs in artificial seawater medium (ASW [10^6 CFU/mL]) for 90 min. Given the extreme relatedness of most EPRs, the tests were performed on EPR225 and EPR117, as reference strains for the clades D and A, respectively. Both strains showed increased survival when incubated with oyster hemocytes compared to those with mussel hemocytes (Fig. 6). Strain EPR117 showed the lowest survival rate, reaching 57% and 67% after 90 min of incubation with mussel and oyster hemocytes, respectively, whereas the values for EPR225 were 75% and 83%.

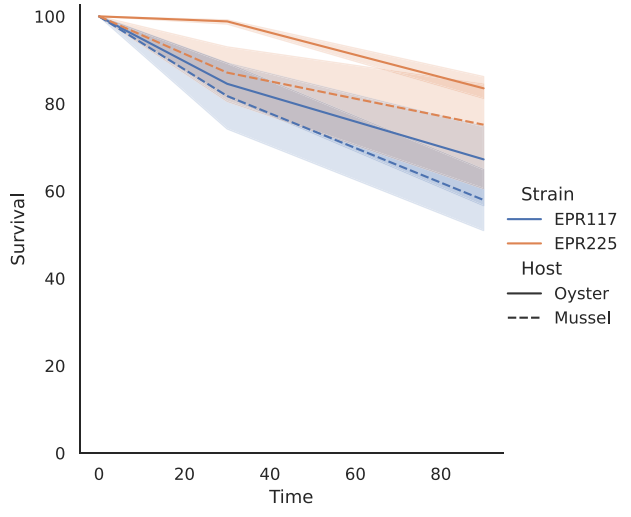

Fig. 6. In vitro bactericidal activity in hemocyte monolayers of the *M. galloprovincialis* and *C. gigas*. Percentages of survival of bacterial cells (106 CFU/mL) incubated on hemocyte monolayers for different periods of time (30 to 90 min), compared to values obtained at time 0.

In parallel, lysosomal membrane stability (LMS) was measured as a proxy for cellular stress in hemocytes challenged with EPRs at different concentrations (10^6, 10^7, and 10^8 CFU/mL), and compared with untreated cells. The results obtained (supplementary table S9, Supplementary Material online) showed a strong decrease in LMS with increasing EPR concentrations (until almost complete destabilization at the highest concentration tested) for both bivalve species. Of the two strains, EPR117 induced the strongest LMS reduction, whereas for both strains the larger effect was observed in oysters rather than mussels (supplementary fig. S3, Supplementary Material online).

Discussion

Deep-sea hydrothermal vents are among the most pristine and remote ecosystems on Earth (Grassle 1987). Fueled by microbial chemosynthesis these areas harbor unique microbial communities (Dick 2019), displaying active chemical cycling and peculiar conditions which could potentially mirror those that occurred at the origin of life on our planet, as they represent one of the earliest types of environments that existed on Earth (Martin et al. 2008). Representatives of the *Vibrio* genus are heterogeneous and can colonize different niches including deep-sea environments. *Vibrio* spp. are responsible for the majority of human diseases attributed to the indigenous microbiota in coastal marine environments (Baker-Austin et al. 2018). According to these features, studies investigating the biology of *Vibrio* spp. in hydrothermal vent settings may help to elucidate fundamental aspects concerning their origin, evolution and pathogenic potential (Hasan et al. 2015; Lasa et al. 2021).

To the best of our knowledge, the pan-genome we defined represents the first attempt at leveraging *Vibrio* diversity to investigate phylogeny and genomic peculiarities of *Vibrio* isolates. When constructing the dataset for our study, we included a representative sequence for each species to avoid redundancies, which could artificially distort the landscape of genetic variability within the genus. While for most of the hundreds of available species, there are few available genomes, there are a few species (e.g. *V. cholerae* and *V. parahemolyticus*) for which thousands of genomes are available, often with negligible differences. Taxonomic classification of Vibrios is challenging due to their elevated genomic plasticity and recombination rates (Thompson et al. 2006; Hunt et al. 2008); however, the phylogenetic analysis we performed using more than 300 shared protein-coding genes delivered a robust reconstruction which placed EPRs in 2 closely related clades. The SNP-based phylogeny provided an increased resolution to inspect the relationships within these clades, showing that one of these, including most (9/10) EPRs, encompassed only the species *V. alginolyticus*, while the other embedded four different species, i.e. *V. alginolyticus*, *V. chemaguriensis*,

V. antiquarius, and *V. diabolicus*, with the last 2 previously isolated from deep-sea vents. This mixed clade was first described by Turner et al. (2018), who considered it an extended *diabolicus* species. We expanded this clade with the species *V. chemaguriensis*, as well as another *V. alginolyticus* strain and EPR225. The results obtained for ANI are consistent with such groups, in that the 2 clades represent distinct, although closely related, species.

Considering the distribution of strains isolated from deep-sea vents, these are paraphyletic, i.e. EPR225, *V. antiquarius* and *V. diabolicus* belong to the clade D, whereas the other EPRs belong to the clade A, implying that colonization of such niches occurred recently after the speciation of these 2 clades. With this classification, we are also able to highlight a biogeographic pattern, in that all representatives of the clade D have been isolated from East Pacific Rise hydrothermal vents located at 9N, whereas members of clade A were also found in Guaymas Basin, where the East Pacific Rise extends into the Gulf of California. While all strains isolated from the Pacific Ocean are similar and belong to the *harveyi* clade, the other deep-sea species present in our dataset, *V. bathopelagicus* (Lasa et al. 2021), differs in several aspects: Taxonomically, as it belongs to the *splendidus* clade, unrelated to *harveyi*, geographically (isolated from the Mediterranean Sea), and environmentally distinct (not isolated from a hydrothermal vent). However, it is important to consider that, with only 10 genomes from two sites, our ability to make definitive conclusions at the population level is limited; a larger sample size could reveal greater diversity and improve the evolutionary insights of deep-sea adaptation. As more isolates from different geographical locations will be available, it will be possible to provide a clearer picture of the biogeography of *Vibrio* from deep-sea hydrothermal vents.

The taxonomic relationship of EPRs with the clade *harveyi* is intriguing, as it embeds several species able to infect humans and animals (Doni et al. 2023), with potentially severe consequences [e.g. *V. vulnificus* can cause necrotizing fasciitis with fatal outcomes (Tsai et al. 2021)]. In particular, the species *V. alginolyticus*, the closest known relative of EPRs, is an established pathogen for both humans (Hornstrup and Gahrn-Hansen 1993; Citil et al. 2015) and aquatic animals (Ahmed et al. 2016), with relevance also for the mariculture industry due to its strong pathogenicity toward bivalves and shrimps (Dong et al. 2020). In particular, the *V. alginolyticus* strain included in our dataset, E110, was isolated from the hepatopancreas of diseased white-leg shrimp from aquaculture in South China (Li et al. 2022). The genomic analysis of EPRs provided a strong indication of their pathogenic potential, as EPRs have the same VFs of *V. alginolyticus* E110, including genes encoding a T3SS which are found also in *V. parahaemolyticus* (Wu et al. 2020). Regarding the other close relatives of EPRs, both *V. diabolicus* and *V. antiquarius* have been isolated from deep-sea hydrothermal vent environments (Raguénès et al. 1997; Hasan et al. 2015), while *V. chemaguriensis* was identified in the estuarine waters of the Chemaguri creek of the Sundarbans, Bengal bay, a brackish water environment (Ghosh and Bhadury 2019).

The virulence potential of EPR strains was further investigated by in vitro assay challenging bivalve hemocytes showing how inoculation of EPRs elicits a strong cellular stress compared with control conditions. As a comparison, the LMS reduction induced by the EPR strains was comparable, if not greater (in mussels), to *V. tasmaniensis*, an established oyster pathogen (Gay et al. 2004; Lasa et al. 2021). At the same time, results of bacterial killing indicated that hemocytes were not able to recognize and kill EPRs efficiently, although this could be influenced also by the cellular stress induced by bacteria.

Surprisingly, the genomes of EPRs present a large number of ARGs (5) compared with other vibrios (2.5). The distribution of such genes in the pan-genome highlighted how 4 genes (*blaCARB*, *ampC*, *catC*, and *tet35*) are prevalently found within the *harveyi* clade, unlike tet34, which is present in various *Vibrio* species. Although these genes were already associated with resistant phenotypes in *V. parahaemolyticus* and *V. harveyi* (Teo et al. 2002; Zhang et al. 2019; Deng et al. 2020), it is unclear the extent to which anthropogenic factors (e.g. exposure to antibiotics in mariculture and their possible over usage) contribute to their emergence in *Vibrio* strains. The presence of ARGs in EPR genomes, found in the relatively undisturbed and remote deep-sea hydrothermal vents, suggests they are not a result of exposure to antibiotics of human origin, but naturally occurring elements which might play a role in microbial warfare (D'Costa et al. 2011; Scott et al. 2020). The high concentration of heavy metals in hydrothermal vents could be another factor contributing to ARG selection, as previously reported (Xu et al. 2017; Anedda et al. 2023).

In our comprehensive analysis of *Vibrio* genomes, we annotated secondary metabolites biosynthetic clusters, revealing an average presence of 4.29 clusters per genome. Notably, all EPR genomes contained four specific biosynthetic clusters for synthesizing ectoine, beta-lactone, arylpolyene, and NI-siderophore. Such clusters were highly conserved across the EPR samples, with only minor compositional variations observed in EPR225. The ubiquity of these clusters highlights their potential role in environmental adaptation and stress response, particularly in osmoregulation and iron acquisition which are critical in marine environments: The beta-lactone cluster, the most prevalent across the surveyed genomes (130/151 genomes), may be involved in bacterial interactions since these molecules exhibit antibiotic properties (Scott 2017), which can be crucial in the competitive environments that *Vibrio* species often inhabit. Arylpolyene has been typically associated with protection against UV radiation or oxidative stress (Schöner

et al. 2016), while NI-siderophores play a role in iron acquisition, particularly important in aquatic environments since iron rapidly precipitates or is not readily bioavailable (Hopkinson and Morel 2009). Finally, ectoine is involved in osmoregulation, since this molecule is a compatible solute maintaining cell volume and protects cellular proteins under hyperosmotic stress conditions, which are common in marine environments (Graf et al. 2008).

Pan-genome analysis reveals the presence of genetic elements that are notably enriched in deep-sea *Vibrio* strains. In particular, the analysis yielded a list of 147 genes, including some previously recognized for their roles in adapting to extreme environments (supplementary table S7, Supplementary Material online). For example, rpoE, known to encode a stress-response sigma factor, has been linked to piezo- and psychro-tolerance in other deep-sea bacteria (Chi and Bartlett 1995), while luxO, associated with bioluminescence regulation (Miyamoto et al. 2000), is thought to contribute to baro-tolerance and the management of reactive oxygen species (Bao et al. 2023). Intriguingly, many of these genes have yet to be functionally characterized and should be subjected to future studies aimed at understanding the molecular mechanisms of deep-sea adaptation.

Overall our study provides new insights into molecular phylogeny and ecology of bacterial strains related to pathogenic *Vibrio* species in deep-sea environments.

Statistics

For statistical analysis, data normality was tested using the Shapiro–Wilk test. Descriptive statistics were calculated to summarize the data. Differences between groups were assessed using ANOVA or Kruskal–Wallis tests, depending on the data distribution. Categorical data (i.e. deep-sea gene enrichment) was analyzed with the Fisher exact test. Phylogenetically aware detection of deep-sea gene enrichment was performed by fitting Pagel '94 model of correlated evolution of two binary characters (Pagel 1994), as implemented in phytools (Revell 2024). An enrichment score was defined by summing the value of the transition matrix corresponding to the states "0|Deep-sea" to "1|Deep-sea," and by subtracting the value corresponding to the states "0|Not Deep-sea" to "1|Not Deep-sea." For multiple test corrections, the Benjamini–Hochberg procedure was applied to control the false discovery rate. A *P*-value of <0.05 was considered statistically significant. Statistical analyses were conducted using either the module stats of the library scipy (Virtanen et al. 2020) or pandas (Mc Kinney).

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Funding

E.B., E.T., L.D., M.A., and L.V. were supported by "National Biodiversity Future Center-NBFC (Italy)" (PNRR CN00000033—Centro Nazionale Biodiversità), from the Italian Ministry of University and Research (MUR). C.O. and L.V. were supported by PRIN 2017 "Emergence of virulence and antibiotic-resistance vectors in coastal and deep-sea marine environments and analysis of the mechanisms and conditions underlying their spread and evolution" (201728ZA49_002), from the Italian Ministry of University and Research (MUR). J.M.-U. was funded by Grants PID2021-127107NB-I00 from Ministerio de Ciencia e Innovación (Spain), 2021 SGR 00526 from Generalitat de Catalunya. C.V. was supported by US National Science Foundation grants OCE 19-48623 and IOS 19-51690 and NASA grant 80NSSC21K0485.

Data Availability

The data underlying this article are available in NCBI and can be accessed with the following accessions: SAMN38691998, SAMN38691999, SAMN38692000, SAMN38692001, SAMN38692002, SAMN38692003, SAMN38692004, SAMN38692005, SAMN38692006, and SAMN38692007.

Literature Cited

Ahmed R, Rafiquzaman SM, Hossain MT, Lee J-M, Kong I-S. Species-specific detection of *Vibrio alginolyticus* in shellfish and shrimp by real-time PCR using the groEL gene. Aquac Int 2016:24(1):157–170. https://doi.org/10.1007/s10499-015-9916-5.

Andrianasolo EH, Haramaty L, Rosario-Passapera R, Bidle K, White E, Vetriani C, Falkowski P, Lutz R. Ammonificins A and B, hydroxyethylamine chroman derivatives from a cultured marine hydrothermal vent bacterium, Thermovibrio ammonificans. J Nat Prod. 2009:72(6):1216–1219. https://doi.org/10.1021/np800726d.

Anedda E, Farrell ML, Morris D, Burgess CM. Evaluating the impact of heavy metals on antimicrobial resistance in the primary food production environment: a scoping review. Environ Pollut. 2023:320:121035. https://doi.org/10.1016/j.envpol.2023.121035.

Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet. 2022:399(10325):629–655. https://doi.org/10.1016/S0140-6736(21)02724-0.

Auguste M, Rahman FU, Balbi T, Leonessi M, Oliveri C, Bellese G, Vezzulli L, Furones D, Canesi L. Responses of Mytilus galloprovincialis to challenge with environmental isolates of the potential emerging pathogen Malaciobacter marinus. Fish Shellfish Immunol. 2022:131:1–9. https://doi.org/10.1016/j.fsi.2022.09.048.

Baker-Austin C, Oliver JD, Alam M, Ali A, Waldor MK, Qadri F, Martinez-Urtaza J. Vibrio spp. infections. Nat Rev Dis Primers. 2018:4(1):8. https://doi.org/10.1038/s41572-018-0005-8.

Bao X-C, Tang H-Z, Li X-G, Li A-Q, Qi X-Q, Li D-H, Liu S-S, Wu L-F, Zhang W-J. Bioluminescence contributes to the adaptation of deep-sea bacterium photobacterium phosphoreum ANT-2200 to high hydrostatic pressure. Microorganisms. 2023:11(6):1362. https://doi.org/10.3390/microorganisms11061362.

Bar-On YM, Phillips R, Milo R. The biomass distribution on Earth. Proc Natl Acad Sci U S A. 2018:115(25):6506–6511. https://doi.org/10.1073/pnas.1711842115.
Blicharska M, Smithers RJ, Mikusiński G, Rönnbäck P, Harrison PA, Nilsson M, Sutherland WJ. Biodiversity's contributions to sustainable development. Nat Sustain. 2019:2(12):1083–1093. https://doi.org/10.1038/s41893-019-0417-9.
Blin K, Shaw S, Medema MH, Weber T. The antiSMASH database version 4: additional genomes and BGCs, new sequence-based searches and more. Nucleic Acids Res. 2023:52(D1):D586–D589. https://doi.org/10.1093/nar/gkad984.
Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics. 2014:30(15):2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
Buchfink B, Reuter K, Drost H-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods. 2021:18(4):366–368. https://doi.org/10.1038/s41592-021-01101-x.
Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol. 2021:38(12):5825–5829. https://doi.org/10.1093/molbev/msab293.
Chen L, Zheng D, Liu B, Yang J, Jin Q. VFDB 2016: hierarchical and refined dataset for big data analysis–10 years on. Nucleic Acids Res. 2016:44(D1):D694–D697. https://doi.org/10.1093/nar/gkv1239.
Chi E, Bartlett DH. An rpoE-like locus controls outer membrane protein synthesis and growth at cold temperatures and high pressures in the deep-sea bacterium Photobacterium sp. strain SS9. Mol Microbiol. 1995:17(4):713–726. https://doi.org/10.1111/j.1365-2958.1995.mmi_17040713.x.
Citil BE, Derin S, Sankur F, Sahan M, Citil MU. Vibrio alginolyticus associated chronic myringitis acquired in Mediterranean waters of Turkey. Case Rep Infect Dis. 2015:2015:187212. https://doi.org/10.1155/2015/187212.
D'Costa VM, King CE, Kalan L, Morar M, Sung WWL, Schwarz C, Froese D, Zazula G, Calmels F, Debruyne R, et al. Antibiotic resistance is ancient. Nature. 2011:477(7365):457–461. https://doi.org/10.1038/nature10388.
Deng Y, Xu L, Chen H, Liu S, Guo Z, Cheng C, Ma H, Feng J. Prevalence, virulence genes, and antimicrobial resistance of Vibrio species isolated from diseased marine fish in South China. Sci Rep. 2020:10(1):14329. https://doi.org/10.1038/s41598-020-71288-0.
Dick GJ. The microbiomes of deep-sea hydrothermal vents: distributed globally, shaped locally. Nat Rev Microbiol. 2019:17(5):271–283. https://doi.org/10.1038/s41579-019-0160-2.
Dong Y, Zhao P, Chen L, Wu H, Si X, Shen X, Shen H, Qiao Y, Zhu S, Chen Q, et al. Fast, simple and highly specific molecular detection of Vibrio alginolyticus pathogenic strains using a visualized isothermal amplification method. BMC Vet Res. 2020:16(1):76. https://doi.org/10.1186/s12917-020-02297-4.
Doni L, Martinez-Urtaza J, Vezzulli L. Searching pathogenic bacteria in the rare biosphere of the ocean. Curr Opin Biotechnol. 2023:80:102894–NaN. https://doi.org/10.1016/j.copbio.2023.102894.
Doni L, Tassistro G, Oliveri C, Balbi T, Auguste M, Pallavicini A, Canesi L, Pruzzo C, Vezzulli L. Plankton and marine aggregates as transmission vectors for V. aestuarianus 02/041 infecting the pacific oyster Crassostrea gigas. Environ Microbiol Rep. 2023:15:631–641. https://doi.org/10.1111/1758-2229.13206.
Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004:32(5):1792–1797. https://doi.org/10.1093/nar/gkh340.
Freeman JA, Bassler BL. A genetic analysis of the function of LuxO, a two-component response regulator involved in quorum sensing in Vibrio harveyi. Mol Microbiol. 1999:31(2):665–677. https://doi.org/10.1046/j.1365-2958.1999.01208.x.
Gay M, Berthe FCJ, Le Roux F. Screening of Vibrio isolates to develop an experimental infection model in the Pacific oyster Crassostrea gigas. Dis Aquat Organ. 2004:59(1):49–56. https://doi.org/10.3354/dao059049.
Ghosh A, Bhadury P. Vibrio chemaguriensis sp. nov., from Sundarbans, Bay of Bengal. Curr Microbiol. 2019:76:1118–1127. https://doi.org/10.1007/s00284-021-02363-6.
Graf R, Anzali S, Buenger J, Pfluecker F, Driller H. The multifunctional role of ectoine as a natural cell protectant. Clin Dermatol. 2008:26(4):326–333. https://doi.org/10.1016/j.clindermatol.2008.01.002.
Grassle JF. In: Blaxter JHS, Southward AJ, editors. The ecology of deep-sea hydrothermal vent communities. London/New York: Academic Press; 1987. p. 301–362.
Hasan NA, Grim CJ, Lipp EK, Rivera ING, Chun J, Haley BJ, Taviani E, Choi SY, Hoq M, Munk AC, et al. Deep-sea hydrothermal vent bacteria related to human pathogenic Vibrio species. Proc Natl Acad Sci U S A. 2015:112(21):E2813–E2819. https://doi.org/10.1073/pnas.1503928112.
Hopkinson BM, Morel FMM. The role of siderophores in iron acquisition by photosynthetic marine microorganisms. BioMetals. 2009:22(4):659–669. https://doi.org/10.1007/s10534-009-9235-2.
Hornstrup MK, Gahrn-Hansen B. Extraintestinal infections caused by Vibrio parahaemolyticus and Vibrio alginolyticus in a Danish county, 1987-1992. Scand J Infect Dis. 1993:25(6):735–740. https://doi.org/10.3109/00365549309008571.
Blin K. 2023. ncbi-genome-download. Zenodo. https://doi.org/10.5281/ZENODO.8192432.
Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 2019:47(D1):D309–D314. https://doi.org/10.1093/nar/gky1085.
Hunt DE, David LA, Gevers D, Preheim SP, Alm EJ, Polz MF. Resource partitioning and sympatric differentiation among closely related bacterioplankton. Science. 2008:320(5879):1081–1085. https://doi.org/10.1126/science.1157890.
Ireton K, Gunther NW, Grossman AD. spo0J is required for normal chromosome segregation as well as the initiation of sporulation in Bacillus subtilis. J Bacteriol. 1994:176:5320–5329. https://doi.org/10.1128/jb.176.17.5320-5329.1994.
Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90 K prokaryotic genomes reveals clear species boundaries. Nat Commun. 2018:9(1):5114. https://doi.org/10.1038/s41467-018-07641-9.
Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017:14(6):587–589. https://doi.org/10.1038/nmeth.4285.
Karl DM, Taylor GT, Novitsky JA, Jannasch HW, Wirsen CO, Pace NR, Lane DJ, Olsen GJ, Giovannoni SJ. A microbiological study of Guaymas Basin high temperature hydrothermal vents. Deep Sea Res A. 1988:35(5):777–791. https://doi.org/10.1016/0198-0149(88)90030-1.
Lasa A, Auguste M, Lema A, Oliveri C, Borello A, Taviani E, Bonello G, Doni L, Millard AD, Bruto M, et al. A deep-sea bacterium related to coastal marine pathogens. Environ Microbiol. 2021:23(9):5349–5363. https://doi.org/10.1111/1462-2920.15629.
Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021:49(W1):W293–W296. https://doi.org/10.1093/nar/gkab301.

Li X, Zhang C, Jin X, Wei F, Yu F, Call DR, Zhao Z. Temporal transcriptional responses of a Vibrio alginolyticus strain to podoviridae phage HH109 revealed by RNA-Seq. mSystems. 2022:7(2): e0010622. https://doi.org/10.1128/msystems.00106-22.
Lonetto M, Gribskov M, Gross CA. The sigma 70 family: sequence conservation and evolutionary relationships. J Bacteriol. 1992:174(12): 3843–3849. https://doi.org/10.1128/jb.174.12.3843-3849.1992.
Makarova KS, Anantharaman V, Grishin NV, Koonin EV, Aravind L. CARF and WYL domains: ligand-binding regulators of prokaryotic defense systems. Front Genet. 2014:5. https://doi.org/10.3389/fgene.2014.00102.
Martin W, Baross J, Kelley D, Russell MJ. Hydrothermal vents and the origin of life. Nat Rev Microbiol. 2008:6(11):805–814. https://doi.org/10.1038/nrmicro1991.
Mc Kinney W. Pandas: A foundational python library for data analysis and statistics. [accessed 2023 December 28]. https://www.dlr.de/sc/portaldata/15/resources/dokumente/pyhpc2011/submissions/pyhpc2011_submission_9.pdf.
Michalska K, Steen AD, Chhor G, Endres M, Webber AT, Bird J, Lloyd KG, Joachimiak A. New aminopeptidase from 'microbial dark matter' archaeon. FASEB J. 2015:29(9):4071–4079. https://doi.org/10.1096/fj.15-272906.
Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020:37(5):1530–1534. https://doi.org/10.1093/molbev/msaa015.
Miyamoto CM, Lin YH, Meighen EA. Control of bioluminescence in Vibrio fischeri by the LuxO signal response regulator. Mol Microbiol. 2000:36(3):594–607. https://doi.org/10.1046/j.1365-2958.2000.01875.x.
Nakamura K, Takai K. Theoretical constraints of physical and chemical properties of hydrothermal fluids on variations in chemolithotrophic microbial communities in seafloor hydrothermal systems. Prog Earth Planet Sci. 2014:1(1):5. https://doi.org/10.1186/2197-4284-1-5.
Pagel M. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. Proc Biol Sci. 1994:255:37–45. https://doi.org/10.1098/rspb.1994.0006.
Raguénès G, Christen R, Guezennec J, Pignet P, Barbier G. Vibrio diabolicus sp. nov., a new polysaccharide-secreting organism isolated from a deep-sea hydrothermal vent polychaete annelid, Alvinella pompejana. Int J Syst Bacteriol. 1997:47(4):989–995. https://doi.org/10.1099/00207713-47-4-989.
Revell LJ. phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). PeerJ. 2024:12:e16505–NaN. https://doi.org/10.7717/peerj.16505.
Reysenbach AL, Banta AB, Boone DR, Cary SC, Luther GW. Microbial essentials at hydrothermal vents. Nature. 2000:404(6780):835. https://doi.org/10.1038/35009029.
Rousk J, Bengtson P. The microbial regulation of global biogeochemical cycles. Lausanne, Switzerland: Frontiers E-books; 2014.
Schöner TA, Gassel S, Osawa A, Tobias NJ, Okuno Y, Sakakibara Y, Shindo K, Sandmann G, Bode HB. Aryl polyenes, a highly abundant class of bacterial natural products, are functionally related to antioxidative carotenoids. ChemBioChem. 2016:17(3):247–253. https://doi.org/10.1002/cbic.201500474.
Scott T. Doctoral thesis 2017-05-07.
Scott LC, Lee N, Aw TG. Antibiotic resistance in minimally human-impacted environments. Int J Environ Res Public Health. 2020:17(11):3939. https://doi.org/10.3390/ijerph17113939.
Seemann T. abricate: :mag_right: Mass screening of contigs for antimicrobial and virulence genes. Github [Accessed 2023 December 6]. https://github.com/tseemann/abricate.
Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014:30(14):2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
Sievert S, Vetriani C. Chemoautotrophy at deep-sea vents: past, present, and future. Oceanography. 2012:25(1):218–233. https://doi.org/10.5670/oceanog.2012.21.
Stokke R, Reeves EP, Dahle H, Fedøy AE, Viflot T, Lie Onstad S, Vulcano F, Pedersen RB, Eijsink VGH, Steen IH. Tailoring hydrothermal vent biodiversity toward improved biodiscovery using a novel in situ enrichment strategy. Front Microbiol. 2020:11:249. https://doi.org/10.3389/fmicb.2020.00249.
Teo JWP, Tan TMC, Poh CL. Genetic determinants of tetracycline resistance in Vibrio harveyi. Antimicrob Agents Chemother. 2002:46(4): 1038–1045. https://doi.org/10.1128/AAC.46.4.1038-1045.2002.
Thompson FL, Klose KE, AVIB Group. Vibrio2005: the first international conference on the biology of Vibrios. J Bacteriol. 2006:188(13): 4592–4596. https://doi.org/10.1128/JB.00141-06.
Tsai Y-H, Huang T-Y, Chen J-L, Hsiao C-T, Kuo L-T, Huang K-C. Bacteriology and mortality of necrotizing fasciitis in a tertiary coastal hospital with comparing risk indicators of methicillin-resistant Staphylococcus aureus and Vibrio vulnificus infections: a prospective study. BMC Infect Dis. 2021:21(1):771. https://doi.org/10.1186/s12879-021-06518-5.
Turner JW, Tallman JJ, Macias A, Pinnell LJ, Elledge NC, Nasr Azadani D, Nilsson WB, Paranjpye RN, Armbrust EV, Strom MS. Comparative genomic analysis of Vibrio diabolicus and six taxonomic synonyms: a first look at the distribution and diversity of the expanded species. Front Microbiol. 2018:9:1893. https://doi.org/10.3389/fmicb.2018.01893.
Van Dover CL. The ecology of deep-sea hydrothermal vents. Princeton, NJ: Princeton University Press; 2021.
Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, et al. Scipy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020:17(3):261–272. https://doi.org/10.1038/s41592-019-0686-2.
Wu C, Zhao Z, Liu Y, Zhu X, Liu M, Luo P, Shi Y. Type III secretion 1 effector gene diversity among vibrio isolates from coastal areas in China. Front Cell Infect Microbiol. 2020:10:301. https://doi.org/10.3389/fcimb.2020.00301.
Xu Y-B, Hou M-Y, Li Y-F, Huang L, Ruan J-J, Zheng L, Qiao Q-X, Du Q-P. Distribution of tetracycline resistance genes and AmpC β-lactamase genes in representative non-urban sewage plants and correlations with treatment processes and heavy metals. Chemosphere. 2017:170:274–281. https://doi.org/10.1016/j.chemosphere.2016.12.027.
Zhang G, Sun K, Ai G, Li J, Tang N, Song Y, Wang C, Feng J. A novel family of intrinsic chloramphenicol acetyltransferase CATC in Vibrio parahaemolyticus: naturally occurring variants reveal diverse resistance levels against chloramphenicol. Int J Antimicrob Agents. 2019:54(1): 75–79. https://doi.org/10.1016/j.ijantimicag.2019.03.012.

Associate editor: Allie Graham