Comparative Genomic Analysis of *Labrenzia aggregata* (*Alphaproteobacteria*) Strains Isolated From the Mariana Trench: Insights Into the Metabolic Potentials and Biogeochemical Functions

Haohui Zhong[1,2], Hao Sun[1], Ronghua Liu[1], Yuanchao Zhan[1], Xinyu Huang[3,4], Feng Ju[3,4] and Xiao-Hua Zhang[1,2,5]*

[1]College of Marine Life Sciences, and Frontiers Science Center for Deep Ocean Multispheres and Earth System, Ocean University of China, Qingdao, China, [2]Laboratory for Marine Ecology and Environmental Science, Qingdao National Laboratory of Marine Science and Technology, Qingdao, China, [3]School of Engineering, Westlake University, Hangzhou, China, [4]Institute of Advanced Technology, Westlake Institute for Advanced Study, Hangzhou, China, [5]Institute of Evolution & Marine Biodiversity, Ocean University of China, Qingdao, China

Hadal zones are marine environments deeper than 6,000 m, most of which comprise oceanic trenches. Microbes thriving at such depth experience high hydrostatic pressure and low temperature. The genomic potentials of these microbes to such extreme environments are largely unknown. Here, we compare five complete genomes of bacterial strains belonging to *Labrenzia aggregata* (*Alphaproteobacteria*), including four from the Mariana Trench at depths up to 9,600 m and one reference from surface seawater of the East China Sea, to uncover the genomic potentials of this species. Genomic investigation suggests all the five strains of *L. aggregata* as participants in nitrogen and sulfur cycles, including denitrification, dissimilatory nitrate reduction to ammonium (DNRA), thiosulfate oxidation, and dimethylsulfoniopropionate (DMSP) biosynthesis and degradation. Further comparisons show that, among the five strains, 85% gene functions are similar with 96.7% of them encoded on the chromosomes, whereas the numbers of functional specific genes related to osmoregulation, antibiotic resistance, viral infection, and secondary metabolite biosynthesis are majorly contributed by the differential plasmids. A following analysis suggests the plasmidic gene numbers increase along with isolation depth and most plasmids are dissimilar among the five strains. These findings provide a better understanding of genomic potentials in the same species throughout a deep-sea water column and address the importance of externally originated plasmidic genes putatively shaped by deep-sea environment.

Keywords: *Labrenzia aggregata*, the Mariana Trench, sulfur cycle, nitrogen cycle, metabolic capacity, hadal zone, comparative genomics

INTRODUCTION

One of the largest extreme environments on Earth is the water under high hydrostatic pressure (HHP), i.e., the deep sea, which contributes up to approximately 95% of the oceans' volume. Among the deep-sea ecosystems, hydrostatic pressure elevates greatly in oceanic trenches belonging to the hadal zone (deeper than 6,000 m). The Challenger Deep of the Mariana Trench is the deepest oceanic ecosystem on earth, in which the hydrostatic pressure elevates up to 110 MPa and the temperature remains low (1–2°C; Jamieson et al., 2010). The first obligate piezophilic microorganism isolated from the Challenger Deep was a strain MT-41 (Yayanos et al., 1981) which were classified into genus *Colwellia* later (DeLong et al., 1997). Later, many more microorganisms were recovered from the Mariana Trench thanks to developments in deep-sea sampling technologies, such as those from sediments (Kato et al., 1998; Pathom-Aree et al., 2006; Zhang et al., 2019a), animals (Kusube et al., 2017), and seawater (Zhou et al., 2018). Previous studies based on pure cultures focused on comparing piezophilic isolates from the same water depth, such as *Actinomycetes* isolates (Pathom-Aree et al., 2006) and *Shewanella benthica* (Zhang et al., 2019a), or on comparing deep-sea isolates with their counterparts from other surface waters (Peoples et al., 2020). Few studies considered the difference between the hadal and surficial isolates recovered throughout the same water column. Although recent development of non-cultivation methods (such as single cell sequencing and genome-resolved metagenomics) has greatly expanded the metabolic potentials of microbes from trenches (León-Zayas et al., 2015; Gao et al., 2019; Zhong et al., 2020), the mobile elements in genomes, such as plasmids, could not be retrieved by metagenomics (binning) because of their disproportionate sequencing depths (Maguire et al., 2020) and the structure of chromosomes and plasmids could not be observed by the highly fragmented draft genomes from these methods. Thus, an in-detail comparison based on pure cultures between vertically distributed ecotypes of the same species, especially between their mobile elements, is required to uncover the metabolic potentials and differentiation of these widely distributed bacteria in hadal environment.

Labrenzia aggregata (formerly *Stappia aggregata*) is heterotrophic, facultative anaerobic *Alphaproteobacteria* ubiquitous in marine environments, participating in nitrogen and sulfur cycles, such as denitrification (King, 2003; Weber and King, 2007; Wyman et al., 2013) and dimethylsulfoniopropionate (DMSP) production (Curson et al., 2017). DMSP is an important osmoprotectant (Cosquer et al., 1999; Kiene et al., 2000) or a carbon source (Yoch, 2002; Curson et al., 2011; Zhang et al., 2019b) for bacteria. *L. aggregata* has also been reported involving in carbon monoxide oxidation (Weber and King, 2007). Furthermore, an ammonia monooxygenase gene was predicted related to nitrification in the type strain *L. aggregata* IAM 12614 (Coates and Wyman, 2017). Besides, there was a speculation of carbon fixation *via* the Calvin cycle in *L. aggregata* considering the existence of the key enzyme ribulose-1,5-bisphosphate carboxylase-oxygenase (RuBisCO; Weber and King, 2007). Due to the incompleteness

of the previously reconstructed genomes, these predictions require further examination.

In this study, we isolated four strains of *L. aggregata* at the Challenger Deep from surface water down to 9,600 m depth. Another *L. aggregata* strain LZB033 with a draft genome assembly isolated previously from the surface water of East China Sea (Curson et al., 2017) was also included for comparison. Complete genomes of the five strains were sequenced and assembled, resulting in more detailed chromosome-plasmid level insights into functional prediction of this species and the locations of the key genes. Eventually, we draw an overview of the metabolic potentials on how *L. aggregata* differs between the surficial and hadal strains and survives in the deepest ocean.

MATERIALS AND METHODS

Sampling, Isolation, Cultivation, and DNA Extraction

All seawater samples were collected using Niskin bottles. SDL044 was isolated from the surface seawater samples (25° 0.527′N, 145° 59.514′E) of the Mariana Trench in October 2017. RF14 was isolated from the seawater samples (11° 23.036′N, 142° 30′E) of the Mariana Trench at a depth of 4,000 m in April 2016. ZYF703 and ZYF612 were isolated from seawater samples (11° 22.160′N, 142° 20.524′E) of the Mariana Trench in November 2016 at depths of 9,600 m. All these four strains were routinely grown on marine agar 2216E (MA; Becton Dickinson) at 28°C. A single colony was picked from original plates, and streaked and re-streaked at least three times to ensure purification (Zhao et al., 2020). Genomic DNA was extracted following the phenol-chloroform-isoamylic alcohol extraction protocol of Marmur (1961). For the pre-test of cultivation under HHP, liquid media inoculated with the cultures were diluted into the same optical densities as above. Next, these media were aspirated into 2.5 ml syringes before transfer into high pressure reactors for 2 weeks at 40, 60, and 80 MPa, 28°C as shown in Zheng et al. (2020). In total, there were three syringes per strain (triplicates) per pressure. After that, the media in the 1x, 10x, and 100x dilutions were cultivated on MA using the spread plate technique for 2 days at 28°C, and subsequently, the average colony numbers were counted and calculated. Growth curves of these five strains under different temperatures (4°C, 16°C, 28°C, and 37°C) were recorded by measuring the optical densities under atmospheric pressure.

Sequencing and Assembling

All the five strains were sequenced by Illumina Hiseq 4,000 with a 270 bp pair-end library and PacBio with 20 kb library at the Beijing Genomics Institute (BGI; Shenzhen, China). Quality control of reads was performed at the same institute. As for assembling, firstly, the PacBio reads were assembled with canu (v1.3; Koren et al., 2017), and Illumina reads were used later as correction by REAPR (v1.0.18; Hunt et al., 2013). Settings in assembling were all default. Later, the Illumina reads were mapped back to the assembled genomes with BBMap (Bushnell, 2014) to calculate the sequencing depths, which

were mapped read numbers multiplying by read length then divided by the total length of the chromosome or plasmid.

Genomic Annotation and Analyses

Annotations of these genomes were conducted by RASTtk online service with default settings (Brettin et al., 2015) and most of the analyses (including CDS numbers and gene count) are based on this result. The genomes were also annotated by eggnog-mapper on EggNOG database v5.0 (Huerta-Cepas et al., 2019) to eliminate suspicious results. Ambiguous annotations were manually checked against the KEGG database and non-redundant (NR) database in NCBI. Further clustering of 95% identical genes was performed by CD-HIT v 4.8.1 (Fu et al., 2012) with parameters "-c 0.95 -g 1."

For the phylogenomic analysis of the five strains, the DNA sequences encoding the 52 ribosomal proteins were selected and aligned individually by a codon-specific multiple-sequence aligner MACSE (Ranwez et al., 2018) with parameters "-gc_def 11." Phylogenetic inference was conducted by IQ-TREE v 2.1.3 (Minh et al., 2020) based on a manually concatenation of the ribosomal protein genes with parameters "-B 1000 -m MFP --seqtype CODON11," which could automatically select the best-fit codon model according to Bayesian information criterion. ANIs between these strains were calculated by JspeciesWS (Richter et al., 2016). For the phylogeny of *rbcL* genes, all of Swiss-Prot reviewed reference genes were downloaded from UniProt with key word rbcL/RuBisCO or rpl/RuBisCO-like proteins. After downloading, these amino-acid sequences were de-replicated with an 80% identity cutoff by CD-HIT as mentioned above with parameter "-c 0.8." Then, the amino-acid sequences of the *rbcL* genes from *L. aggregata* and reference were aligned by MAFFT v7.471 with default parameters (Katoh and Standley, 2013). After that, we deleted the sites that contains gaps in more than 50% sequences from the alignment. Finally, phylogenetic analysis was conducted by IQ-TREE 2 based on the curated alignment with parameters "-B 1000 -m MFP --msub nuclear".

Prophages in the genomes were detected by PHASTER with default settings (Arndt et al., 2016). To identify the genes encoding secondary metabolites and antibiotic resistance genes (ARGs), the genes were firstly searched against the structured antibiotic resistance gene database (SARGs) as described by Yang et al. (2013). The annotation of candidate ARGs was further confirmed by comparing with those derived from the aforementioned RAST, EggNOG, and KEGG database to eliminate ambiguous assignment. The ARGs were further manually assigned into resistance types and subtypes (Yang et al., 2016). Biosynthetic gene clusters of secondary metabolites, including NRPS, were mined by the antiSMASH database and tools with default settings (Blin et al., 2019).

Disk Diffusion Tests

Disks with diameters of 6 mm containing the same concentration of antibiotics were put on seeded agar plates. The concentration of the five strains was adjusted to the same (the optical density) before the test to make sure the results were comparable. The plates were cultured for 24 h at 28°C before the diameters of the inhibition zones were documented for downstream analysis.
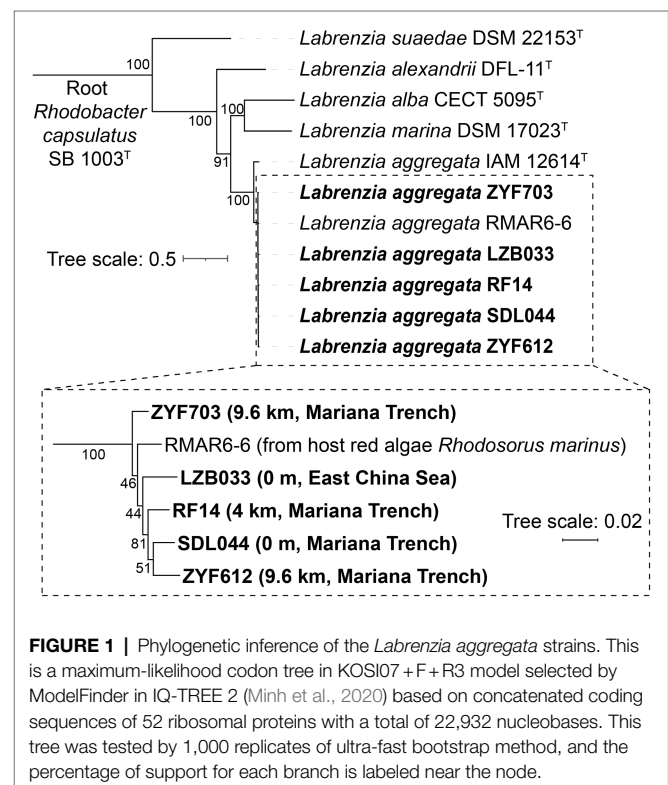
RESULTS AND DISCUSSION

Phylogenomic Analysis Confirms That the Five Strains Belong to the Same Species

Four *L. aggregata* strains SDL044, RF14, ZYF703, and ZYF612 were isolated from seawater samples from the Mariana Trench at depths of 0 m, 4,000 m, 9,600 m, and 9,600 m, respectively. Since each of these strains contains three copies of 16S rRNA genes and resolution of these genes is not high enough between strains, 52 ribosomal protein genes, which encode some of the most conserved proteins in all lives and are commonly used for phylogenomic analysis, were chosen for their classification, including those from *L. aggregata* strains and several reference genomes (**Figure 1**). In addition to LZB033, all the five strains are clustered on the same branch in the tree and located distantly from other species, indicating they all belong to *L. aggregata* with 100% branch support. Moreover, average nucleotide identities (ANIs) between each of the strains and the representative genome of *L. aggregata* are between 97 and 98%, which is higher than the empirical intra-species threshold – 95% (Goris et al., 2007; **Supplementary Table S1**), furtherly confirming the classification of the four isolates from the Mariana Trench.

Geographical Origin of the Hadal Strains

A pressurization test shows all the four strains from the Mariana Trench could partially tolerate HHP with higher survival rate in the hadal strains (**Supplementary Figure S1**). Another cultivation experiments measured by optical densities suggest

FIGURE 1 | Phylogenetic inference of the *Labrenzia aggregata* strains. This is a maximum-likelihood codon tree in KOSI07 + F + R3 model selected by ModelFinder in IQ-TREE 2 (Minh et al., 2020) based on concatenated coding sequences of 52 ribosomal proteins with a total of 22,932 nucleobases. This tree was tested by 1,000 replicates of ultra-fast bootstrap method, and the percentage of support for each branch is labeled near the node.

that under atmospheric pressure (0.1 MPa), all the five strains grow optimally under 37°C (**Supplementary Figure S2**). We could not observe any noticeable change in optical densities of these strains during a cultivation under 4°C, 80 MPa for two weeks. Thus, all these results indicate that the deep-sea *L. aggregata* strains are not piezophilic and they have not completely adapted to hadal HHP. It is more likely that the two hadal strains sank from surficial water recently and were inactive (or grew extremely slowly) under HHP at 9,600 m depth. Previously reported non-piezophilic trench species *Pseudomonas bathycetes*, which was the first isolate from sediments of the Mariana Trench, grew slowly with generation time up to 33 days under *in situ* condition: 2°C, 100 MPa (Schwarz and Colwell, 1975). Several recent studies have shown the connectivity of deep-sea and surficial microbial communities and most of the dominant prokaryotes from the deep ocean can also be detected in surface waters according to operational taxonomic unit analysis (Agustí et al., 2015; Mestre et al., 2018). Our isolations could be living examples of these sinking microbes.

Genomic Composition: Similar Core Chromosomic Genes but Differential Plasmids

Each of the five strains consists of one large chromosome (5.83 to 5.89 Mbp) and 2–4 plasmids (**Table 1**). These plasmids and chromosomes are distinguished by their sizes and differential sequencing depths (**Supplementary Table S2**). To avoid ambiguity, the plasmids are sorted and named by their sizes (**Table 1**, P1 ~ P4). All of these genomes share similar GC contents of $59.0 \pm 0.2\%$ and total sizes of 6.49–6.96 Mbp, whereas the strains from greater depths have larger total size of plasmids (the largest one is 76.6% larger than the smallest one). With these genomes getting functionally annotated, more than 6,000 coding DNA sequences (CDSs) are predicted in each of these strains. Consistent with the trends of genome size, the greater the depth from which the strain was recovered, the more CDSs are predicted; the 9.6 km strain ZYF612 contains 8.12% more CDSs (predominantly contributed by the plasmids) than the 0 m strain SDL044. When removing the redundancy of the annotation, the gene function numbers of the five strains differ slightly, ranging from 3,048 to 3,121, which are inconsistent with their sizes: the 4 km strain RF14 (third largest in size) contains only 3,048 annotated gene functions. Examinations based on Venn diagrams (**Figures 2A–C**) suggest 90% gene annotations on the chromosomes are similar, while only 27% are shared by the plasmids. The 9.6 km strain ZYF612 harbors most specific genes majorly contributed by its plasmids (120 or 20% of total plasmidic gene functions, **Figure 2B**). Since the ZYF612 contains an extra of more than 400 CDSs than other strains (6,825 vs. 6,248 ~ 6,456) but similar gene function numbers (3,121 vs. 3,048 ~ 3,088), these extra genes should mainly from multiple copies and hypothetical protein genes (more than 12.6%; 2,215 vs. 1,863 ~ 1,967). Further investigation suggests most of these multiple copies are related to phages and thus, these gene duplications might be a sign of phage

infections in ZYF612. Details of the phage-related genes will be discussed in a following section "Horizontal Gene Transfers in *L. aggregata*". Another analysis based on 95% nucleotide identity gene clustering (**Figure 2D**) suggests the pan-genome of *L. aggregata* could still be fast-expanding if adding more strains while the core-genome is relatively stable.

Metabolic Reconstruction: Core Carbon Metabolism

As shown in their metabolic reconstruction maps (**Figure 3**), the core metabolisms of the five strains are highly similar. Gluconeogenesis (GNG) and non-oxidative pentose phosphate pathway (PPP) are predicted in all of them. All the five genomes are devoid of the key gene encoding 6-phosphofructokinase in glycolysis (Embden–Meyerhof–Parnas pathway) or the gene encoding 6-phosphogluconolactonase participating in oxidative PPP and classical Entner-Doudoroff (ED) pathway (Ahmed et al., 2005), making these strains incapable of degrading glucose through these canonical pathways. Instead, a modified pathway incorporating glucose dehydrogenase (*gcd*), gluconolactonase (*gnl*), and gluconokinase (*gk*) is encoded in four strains except for RF14, generating 6-phospho-gluconate for subsequent non-oxidative PPP or the latter half of ED pathway. Similar modified glucose degrading pathways were reported previously in several *Cyanobacteria* (Chen et al., 2016). The genes *gcd* and *gnl* also involve in the archaeal semi-phosphorylative and non-phosphorylative ED pathway (Lamble et al., 2005; Reher and Schönheit, 2006), but a following key enzyme gluconate dehydratase is absent in all the five strains, making the modified ED pathway in *L. aggregata* different from these variants. Interestingly, only *gnl* genes are encoded on the largest plasmid of each strain (referred to as P1 hereafter, including LZB033P1, SDL044P1, RF14P1, ZYF703P1, and ZYF612P1), while all other genes are on the chromosomes. Overall, genetic predictions suggest that at least four strains (except for RF14) could conserve energy through complete glucose (and related carbohydrates) degradation into CO_2 *via* the modified ED pathway, the tricarboxylic acid cycle, and the respiratory chain.

CO oxidation is predicted in all the five strains according to the presence of aerobic carbon monoxide dehydrogenase complex (*coxL, coxM, and coxS*). The generated CO_2 was previously speculated being assimilated by the autotrophic Calvin cycle in the type strain of *L. aggregata* considering the presence of the gene *rbcL* (also known as *cbbL*, encoding the key enzyme RuBisCO; King and Weber, 2007). However, a phylogenetic analysis of the *rbcL* (**Supplementary Figure S3**) indicates that these *rbcL* genes encode form IV RuBisCO proteins, also known as RuBisCO-like proteins, which are incapable of catalyzing CO_2 fixation (Tabita et al., 2008); another gene encoding phosphoribulokinase (*prk*) essential in the Calvin cycle is also absent in all the five strains. In addition to the absence of other complete carbon fixation pathways, these five strains therefore are unlikely mixotrophic or capable of carbon fixation. The CO oxidation in *L. aggregata* might only function as a supplementary source of electrons for the respiratory chain, and the CO_2 molecules generated in this process are more likely being released instead of being assimilated.

TABLE 1 | Genomic information of the five completely sequenced *Labrenzia aggregata* strains.

Strain	Isolation Location	Isolation Depth (m)	Isolation Source	Genome (Mbp)	Chromosome (Mbp)	Plasmids Total (Mbp)	Each Plasmid (Mbp)	G + C (%)	CDS	Unique Gene Annotations	Hypothetical Proteins	tRNA
LZB033	East China Sea	0	Water	6.49	5.85	0.64	0.45, 0.19	59	6,248	3,052	1,863	52
SDL044	Mariana Trench	0	Water	6.55	5.89	0.66	0.46, 0.10, 0.07, 0.04	59.1	6,312	3,067	1,913	53
RF14	Mariana Trench	4,000	Water	6.62	5.87	0.75	0.27, 0.25, 0.13, 0.10	59	6,421	3,048	1,952	52
ZYF703	Mariana Trench	9,600	Water	6.69	5.88	0.81	0.45, 0.19, 0.17	59	6,456	3,088	1,967	51
ZYF612	Mariana Trench	9,600	Water	6.96	5.83	1.13	0.47, 0.25, 0.22, 0.19	58.8	6,825	3,121	2,215	54

The plasmids are sorted in descending order of their sizes and named after this order herein (LZB033P1, LZB033P2; SDL044P1 ~ SDL044P4; RF14P1 ~ RF14P4; ZYF703P1 ~ ZYF703P3; ZYF612P1 ~ ZYF612P4; and P1 = largest in size).

CO oxidation in deep-sea microbes (putatively in *Bacteroidetes* or *Chloroflexi*) has been reported previously in the Mediterranean (Martin-Cuadrado et al., 2009) and our isolates represent an *alphaproteobacterial* origin of these genes in deep sea.

Genes Related to Respiration and Biogeochemical Cycles

An electron transport chain is encoded in each genome of strains with oxygen or nitrate as final electron acceptors. Under anoxic conditions, genetically nitrate could be reduced into nitrogen (denitrification) or ammonium (DNRA) according to the presence of dissimilatory periplasmic nitrate reductase (*napA and napB*), nitrite reductase (NO-forming; *nirS*), nitric oxide reductase (*norB and norC*), nitrous oxide reductase (*nosZ*), and NADH-dependent nitrite reductase (*nirB and nirD*). All these genes have been detected in *L. aggregata* type strain IAM 12614 previously (Coates and Wyman, 2017), and denitrification in *L. aggregata* was also reported with experimental evidence (Weber and King, 2007). Interestingly, the denitrification genes are all located on the largest plasmid of each strain (P1), while the DNRA genes are encoded on chromosomes, indicating they might have different origins. Also an extra copy of *nosZ* and related genes (operon *nosRZDFYL*) could be predicted on the plasmid ZYF612P2. Another previously reported gene encoding ammonia monooxygenase related to nitrogen cycle (nitrification) in the strain IAM 12614 (Coates and Wyman, 2017) is absent in all our five strains.

When it comes to sulfur metabolism, all the four Mariana Trench strains contain the key DMSP biosynthetic gene *dsyB*, >99% amino-acid identical to the experimentally confirmed one in LZB033 (Curson et al., 2017). The gene for DMSP cleavage (*dddL*) is also universal among them, both suggesting a common role for *L. aggregata* in DMSP metabolism throughout the ocean. Other sulfur-related genes include a complete *sox* thiosulfate oxidation pathway that is encoded in each of the strains, possibly serving as a supplementary source of electrons for ATP biosynthesis. Moreover, genetically sulfate could be reduced for (homo-)cysteine biosynthesis (assimilatory sulfate reduction) in the five strains according to the presence of related genes.

Specific Genes in Hadal Strains

According to the genome annotations, only nine gene functions are shared by the two hadal strains and are absent in other strains (as shown in red fonts in **Figure 2C**). There are more strain-specific gene functions in the two hadal strains (90 in ZYF612 and 35 in ZYF703, shown in brown fonts in **Figure 2C**). The functions of these genes could not be summarized by one or two pathways and are divided into various functional categories, some of which related to hadal environment will be discussed separately in the following paragraphs.

Bacteria adapt to HHP through various changes, such as osmoregulation and respiratory chain (Simonato et al., 2006), the former of which relies on organic osmolytes or osmoprotectants (Galinski, 1995; Wood, 2011). Biosynthetic pathways of osmoprotectants in the five *L. aggregata* strains include DMSP, betaine (trimethylglycine), N-acetylglutaminylglutamine amide (NAGGN), and osmoregulated periplasmic glucans (OPGs). DMSP
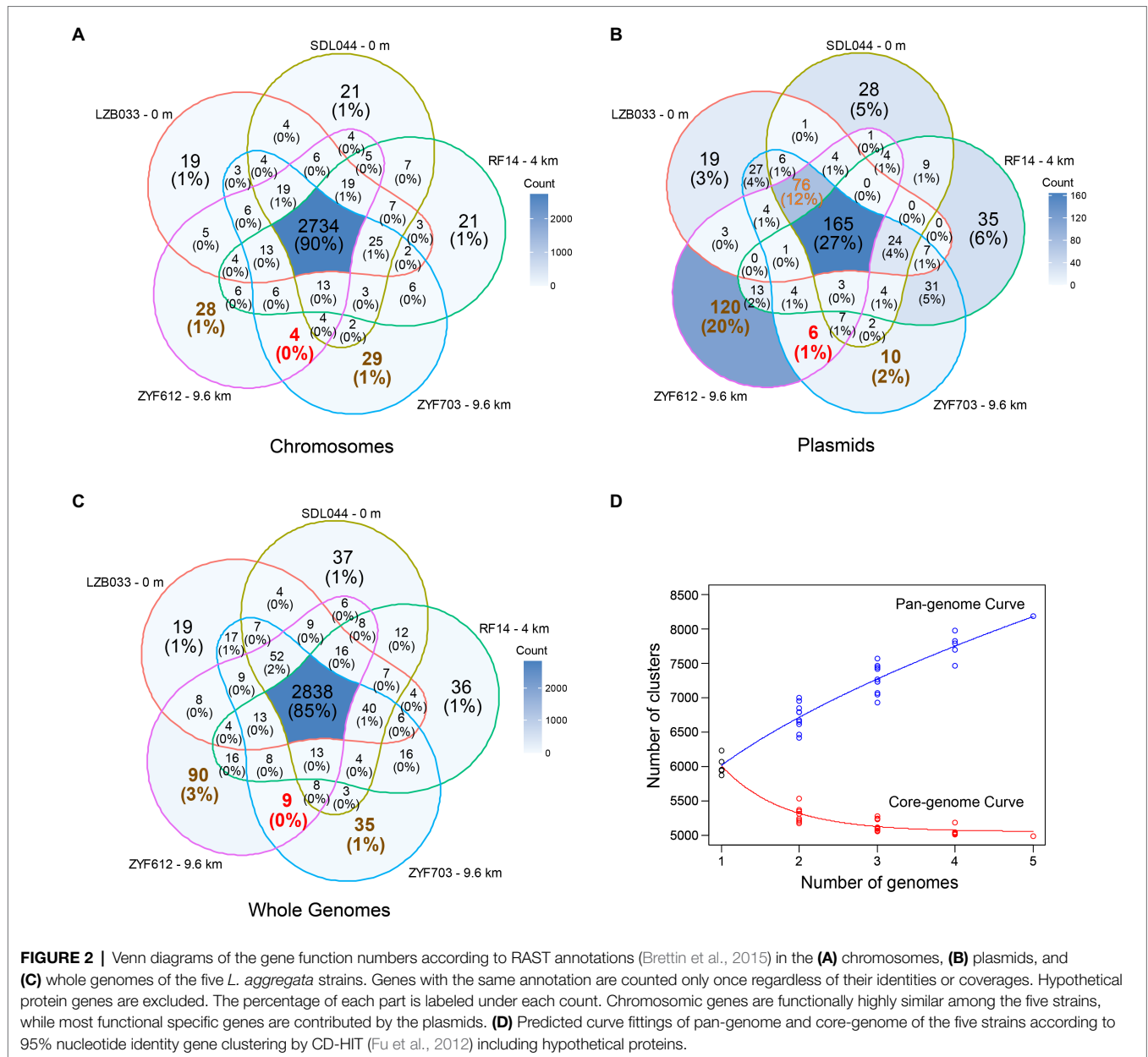
FIGURE 2 | Venn diagrams of the gene function numbers according to RAST annotations (Brettin et al., 2015) in the **(A)** chromosomes, **(B)** plasmids, and **(C)** whole genomes of the five *L. aggregata* strains. Genes with the same annotation are counted only once regardless of their identities or coverages. Hypothetical protein genes are excluded. The percentage of each part is labeled under each count. Chromosomic genes are functionally highly similar among the five strains, while most functional specific genes are contributed by the plasmids. **(D)** Predicted curve fittings of pan-genome and core-genome of the five strains according to 95% nucleotide identity gene clustering by CD-HIT (Fu et al., 2012) including hypothetical proteins.

has been discussed in the previous section. All the five strains encode two distinct betaine synthesis pathways: oxidation of choline (Cánovas et al., 1996) and methylation of glycine (Kerovuo et al., 2000). The previously identified NAGGN biosynthetic genes (Sagot et al., 2010) are also shared by all the strains. OPGs are located in the periplasm of Proteobacteria and are essential components in extreme environments (Bohin, 2000). Most of the OPG genes are shared by the five strains on chromosomes, whereas giant gene encodes the synthetase for a unique OPG, i.e., cyclic beta-1,2-glucan, which is predicted exclusively in the plasmid of the strain ZYF612 (**Supplementary Figure S4**). Cyclic beta-1,2-glucan is previously found in other *Alphaproteobacteria*, such as *Agrobacterium, Rhizobium,* and *Brucella* as an osmolyte or virulence factor (Dylan et al., 1990; Arellano-Reynoso et al.,

2005). In ZYF612, this gene is most similar to that from *Labrenzia suaedae* (93% amino-acid identity) isolated from a plant, *Suaeda maritima* that grows on tidal flats (Bibi et al., 2014). As for respiratory chain, in the 9.6 km strain ZYF612, several genes encoding *ccb3* cytochrome c oxidase, which are essential to the respiratory chain, are multiple copies located both on the chromosome and the plasmids with identities smaller than 95%, indicating different origins of these genes. Previous studies also found multiple clusters of such oxidase genes in piezophilic bacteria *Photobacterium profundum* SS9 (Vezzi et al., 2005) and *Shewanella benthica* (Zhang et al., 2019a). Although the hadal *L. aggregata* strains have not completely adapt to hadal environment, these genomic features may help them survive under the extremely HHP at 9,600 m depth.
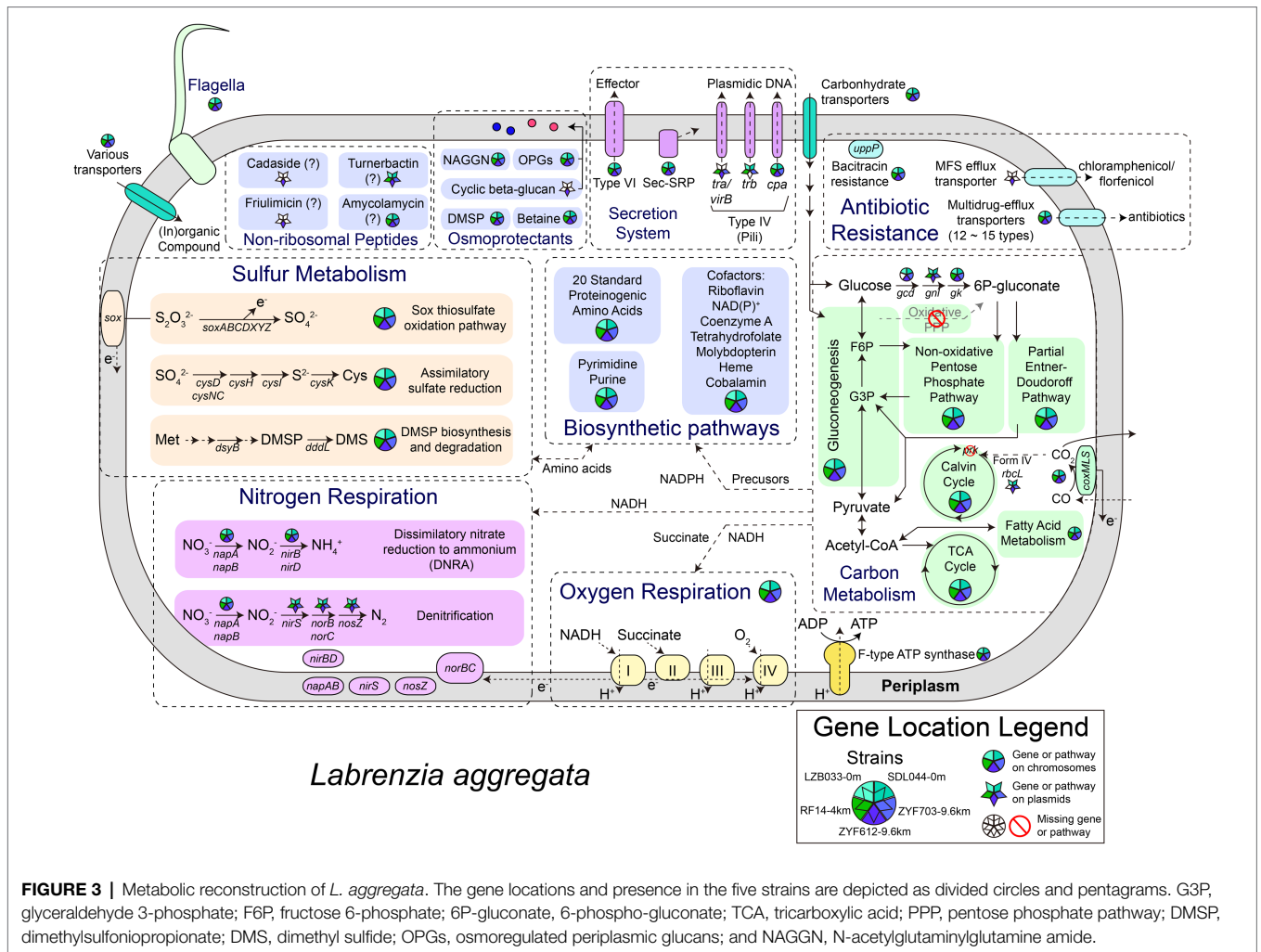
FIGURE 3 | Metabolic reconstruction of *L. aggregata*. The gene locations and presence in the five strains are depicted as divided circles and pentagrams. G3P, glyceraldehyde 3-phosphate; F6P, fructose 6-phosphate; 6P-gluconate, 6-phospho-gluconate; TCA, tricarboxylic acid; PPP, pentose phosphate pathway; DMSP, dimethylsulfoniopropionate; DMS, dimethyl sulfide; OPGs, osmoregulated periplasmic glucans; and NAGGN, N-acetylglutaminylglutamine amide.

Secondary metabolites, such as antibiotics, may give hadal microbes advantages when competing with other microbes, especially in the trenches where nutrients are accumulated by the physical topography (Jamieson et al., 2010) leading to more heterotrophic competitors. Several giant gene clusters encoding non-ribosomal peptide synthetases (NRPSs) were detected (**Supplementary Figure S5**) in the *L. aggregata* strains. These giant multi-domain enzymes synthesize peptide secondary metabolites without ribosomes or mRNA, and some of these metabolites are important antibiotics, antifungal, and antiparasitic agents (Cane and Walsh, 1999; Koglin and Walsh, 2009). Two NRPS gene clusters are present in each of the five strains and located on the chromosome and the largest plasmid of each strain (P1), while three other NRPS gene clusters are only predicted on the smallest plasmid of ZYF612 (ZYF612P4). According to functional prediction, the shared plasmidic NRPS might synthesize a siderophore involving in iron transport (turnerbactin-like, 30% of genes show similarity according to the default identity threshold in antiSMASH v5.0, Blin et al., 2019), while the shared chromosomic NRPS is suspiciously an antibiotic synthase (amycolamycin-like, only 2% of genes show similarity). The specific NRPS gene clusters in ZYF612 are all predicted to be antibiotic-related but their identities with other known clusters are relatively low (approximately 9% of genes show similarity, **Supplementary Figure S5**). These giant genes (such as the NRPS and cyclic beta-1,2-glucan synthase genes) demand a great amount of energy and amino-acid precursors (Reva and Tümmler, 2008). Although functions of these giant genes may be predicted tentatively, their existence may well be crucial for the survival of ZYF612 in the deepest water even with such high costs.

Apart from antibiotic biosynthesis, antibiotic resistance genes also differ in some hadal strains. In the four Mariana Trench strains, the antibiotic resistance gene (ARG) numbers (>50% identities against SARG database) range from 6 to 10, while the ZYF612 strain contains the most. A chloramphenicol/florfenicol resistance gene is specifically predicted in the plasmid ZYF612P1. Furthermore, a simple antibiotics disk diffusion test confirms the higher resistance against chloramphenicol in the three deep-sea strains than the surficial strain SDL044, while the two hadal strains are more resistant to quinolones (**Supplementary Figure S6**). Although currently there is no strong evidence of high antibiotic concentration in the Mariana Trench, the antibiotic resistance genes might be an advantage competing against other bacteria.

Horizontal Gene Transfers in *L. aggregata*

As discussed in the previous sections, several genes are located in the plasmids, such as the denitrification genes. Plasmids could be transferred horizontally between strains by bacterial conjugation *via* pili, which belong to the type IV secretion system (Llosa et al., 2002; Kline et al., 2010; Maier and Wong, 2015). Thus, these plasmidic genes are more tend to be external origin than the chromosomic genes. At least three types of putative pili genes are predicted in the five strains, with one *cpa* type ubiquitously in the chromosomes and two other types uniquely on the plasmids of some strains (**Figure 3**). Since the annotation-based Venn diagram (**Figure 2B**) is not enough for uncharacterized (hypothetical) proteins, an identity-based clustering of the plasmidic genes is conducted, and the result (**Figure 4**) suggests the largest plasmid of each strain (P1) is relatively conserved while the other plasmids differ greatly. It is tempting to speculate that the P1 plasmids in all the five strains are a stable component of the genome while all other plasmids are more frequently transferred horizontally. Since the size of the plasmid RF14P1 is smaller than the others (0.27 Mbp vs. 0.45–0.47 Mbp, **Table 1**) and the shared genes are likewise fewer (shown as the narrower edges in **Figure 4**), it is likely that the plasmid RF14P1 experienced several gene deletion, resulting in the missing 76 gene functions that are shared in the other four strains (shown in orange fonts in **Figure 2B**), such as the gene *rbcL* mentioned above. Gene acquirements of the 76 genes independently in the four strains are unlikely because these shared genes are highly similar (> 95% identity as shown in the edges of **Figure 4**).

A gene origin analysis based on gene annotations (**Supplementary Table S3**) suggests that the extra genes in the hadal strains predominantly come from species in *Bradyrhizobiaceae*, *Oceanicola*, *Sphingomonadales* (all belong to *Alphaproteobacteria*), and *Actinobacteria*. But the numbers of these annotated extra genes in ZYF612 (totally 94 genes) are much less than the extra 400 CDSs. It is reasonable to consider functionally similar proteins with different origins or hypothetical proteins as other contributors of the difference between the annotation-based and identity-based results (**Figures 2B**, **4**). For example, in ZYF612, the copy numbers of genes of external origin, such as phage-associated and transposase genes, are much higher than those in other strains. In the plasmids of ZYF612, there are 66 copies of mobile element proteins (with domains similar to transposases), higher than the average copy number of these in the other four strains (16 copies), which could partially
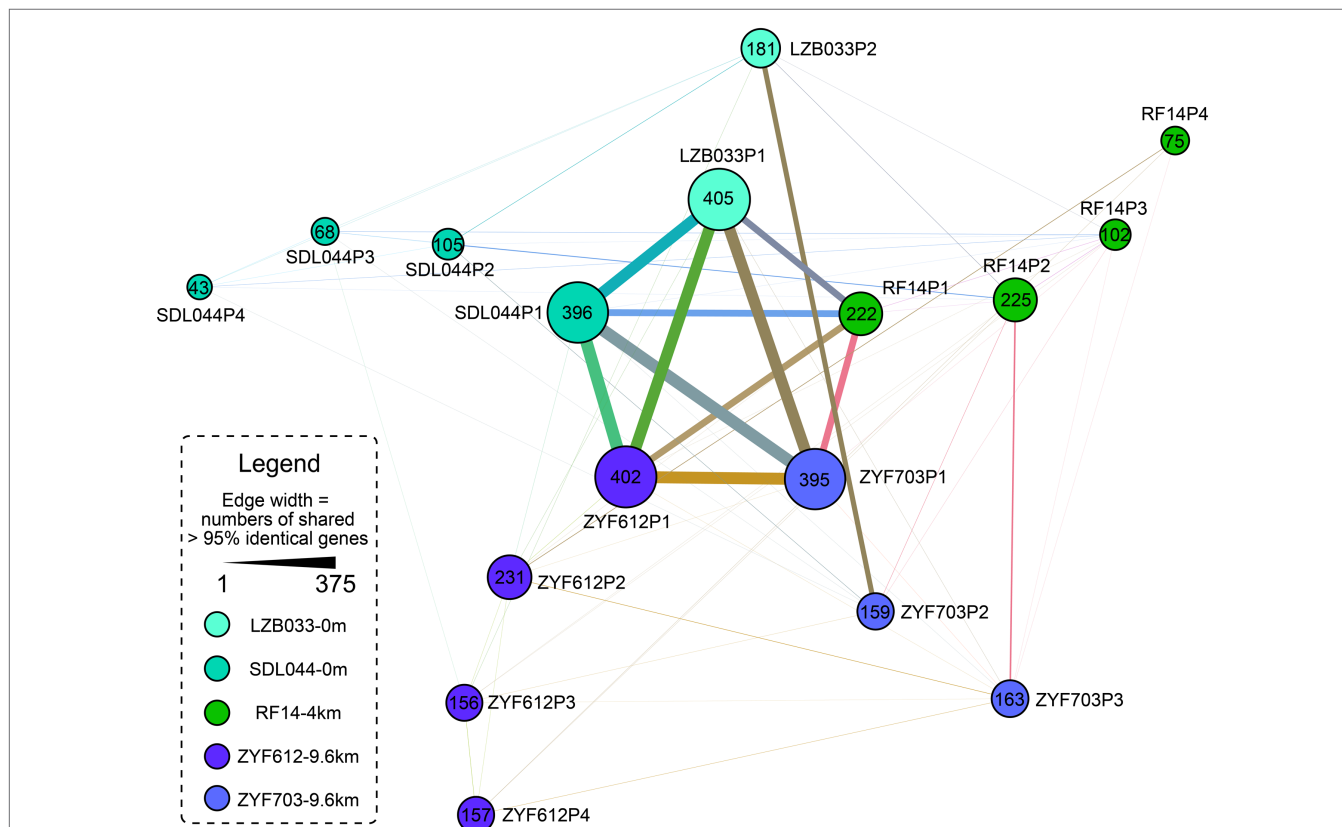

FIGURE 4 | Network of the plasmidic identical genes (considering nucleotide identity >95% as identical) in the five *L. aggregata* strains. Each node represents a plasmid and the size of it reflects the numbers of genes after removing multiple copies. The widths of the edges indicate the numbers of shared genes. Different from **Figure 2B**, this is a network based on identity regardless of annotation; thus, the results also consider previously ignored hypothetical proteins. Most genes are shared by the longest plasmids of the five strains and the other plasmids encode more specific genes. The plasmids are named after their lengths as mentioned in **Table 1**. All these plasmids encode less than 20 identical genes with the chromosomes (mainly in P1; data not shown).

explain why ZYF612 harbors 400 more CDSs than other strains while the gene function number varies slightly. A previous study has also reported transposase being the most overrepresented COG categories in deep-sea microbial communities by fosmid clones (DeLong et al., 2006). As expected, two intact prophages could be predicted on the chromosomes of the two hadal strains ZYF703 and ZYF612 and other two intact ones on the plasmid ZYF612P2 (**Supplementary Table S4**). More incomplete or questionable prophages are also predicted in all the three deep-sea strains including RF14 (**Supplementary Table S4**). These prophages could be another origin of the larger size of the three deep-sea strains; for example, all the prophage proteins contribute to at least 0.11 Mbp genome size and 133 CDSs in ZYF612. Previous studies also reported phage insertion in *Geobacillus* genomes from the Mariana Trench (Takami et al., 2004) and in *Colwellia* genomes from various locations including the Mariana Trench (Peoples et al., 2020). Raised virus prokaryote ratio in hadal water compared to that in surface water has also been reported in the Mariana Trench (raised from 10.8-fold to 41.1-fold, Nunoura et al., 2015) and the Japan Trench (raised from 20~50-folds to 80~260-fold, Nunoura et al., 2016). Therefore, it could be postulated that HGT by virus may be another important source of external genes and partially shape the pan-genome of hadal species.

Overall, external genes were acquired by the hadal strains to confront the hostile environment in trenches, such as the higher hydrostatic pressure. Compared to our recently report of microorganisms that thrive in the hadal zone by consuming hydrocarbons (Liu et al., 2019), the hadal strains of *L. aggregata* have not changed their core metabolism greatly and still retain similar catabolic potentials like their counterparts from shallower depths. Unlike obligate piezophiles, these deep-sea ecotypes sinking from surface water still reserve many genetic potentials to survive in various other environments.

CONCLUSION

The completely sequenced genomes reveal how *L. aggregata* strains may change to resist higher hydrostatic pressure. Genomically, almost identical chromosomes but dissimilar plasmidic composition suggest extra genetic potentials are horizontally acquired partially by phages and mainly through plasmid exchanges (conjugation), such as those large genes encoding cyclic beta-1,2-glucan synthase and NRPS on ZYF612's plasmids. Genes involving in biogeochemical cycles, such as the denitrification genes, may

also be overlooked by binning in certain lineages due to their plasmidic locations. Comparisons of these five strains offer a vertical genomic view of a microbial species along the seawater column, which show that instead of one single common solution, many different strategies have been evolved by the microorganisms concurrently to confront hostile deep-sea environments. In conclusion, our findings provide a metabolic framework for future studies of *L. aggregata*, and moreover, a template of ubiquitously distributed marine bacteria throughout oceans.

DATA AVAILABILITY STATEMENT

The five Labrenzia genome assemblies were uploaded onto NCBI RefSeq with the following accession numbers: GCF_001932055.2 for LZB033, GCF_020882605.1 for SDL044, GCF_020882585.1 for RF14, GCF_020882625.1 for ZYF703 and GCF_020882645.1 for ZYF612.

AUTHOR CONTRIBUTIONS

X-HZ designed the experiments, analyzed the data, and wrote the manuscript. HZ performed the experiments, analyzed the data, prepared all figures, and wrote the manuscript. HS performed the DNA extraction and the genome analysis. RL performed the phenotypic experiments and data analysis. YZ performed the bacterial phage analysis. XH and FJ performed antibiotic resistance analysis. All authors edited and approved the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.770370/full#supplementary-material

REFERENCES

Agustí, S., González-Gordillo, J. I., Vaqué, D., Estrada, M., Cerezo, M. I., Salazar, G., et al. (2015). Ubiquitous healthy diatoms in the deep sea confirm deep carbon injection by the biological pump. *Nat. Commun.* 6:7608. doi: 10.1038/ncomms8608

Ahmed, H., Ettema, T. J., Tjaden, B., Geerling, A. C., van der Oost, J., and Siebers, B. (2005). The semi-phosphorylative Entner-Doudoroff pathway in hyperthermophilic archaea: a re-evaluation. *Biochem. J.* 390, 529–540. doi: 10.1042/BJ20041711

Arellano-Reynoso, B., Lapaque, N., Salcedo, S., Briones, G., Ciocchini, A. E., Ugalde, R., et al. (2005). Cyclic β-1, 2-glucan is a Brucella virulence factor required for intracellular survival. *Nat. Immunol.* 6, 618–625. doi: 10.1038/ni1202

Arndt, D., Grant, J., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21. doi: 10.1093/nar/gkw387

Bibi, F., Jeong, J. H., Chung, E. J., Jeon, C. O., and Chung, Y. R. (2014). Labrenzia suaedae sp. nov., a marine bacterium isolated from a halophyte, and emended description of the genus Labrenzia. *Int. J. Syst. Evol. Microbiol.* 64, 1116–1122. doi: 10.1099/ijs.0.052860-0

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., et al. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47, W81–W87. doi: 10.1093/nar/gkz310

Bohin, J. P. (2000). Osmoregulated periplasmic glucans in Proteobacteria. *FEMS Microbiol. Lett.* 186, 11–19. doi: 10.1111/j.1574-6968.2000.tb09075.x

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., et al. (2015). RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* 5:8365. doi: 10.1038/srep08365

Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner (No. LBNL-7065E). Berkeley, CA (United States): Lawrence Berkeley National Lab.(LBNL).

Cane, D. E., and Walsh, C. T. (1999). The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. *Chem. Biol.* 6, R319–R325. doi: 10.1016/s1074-5521(00)80001-0

Cánovas, D., Vargas, C., Csonka, L. N., Ventosa, A., and Nieto, J. J. (1996). Osmoprotectants in Halomonas elongata: high-affinity betaine transport system and choline-betaine pathway. *J. Bacteriol.* 178, 7221–7226. doi: 10.1128/jb.178.24.7221-7226.1996

Chen, X., Schreiber, K., Appel, J., Makowka, A., Fähnrich, B., Roettger, M., et al. (2016). The Entner-Doudoroff pathway is an overlooked glycolytic route in *Cyanobacteria* and plants. *Proc. Natl. Acad. Sci. U. S. A.* 113, 5441–5446. doi: 10.1073/pnas.1521916113

Coates, C. J., and Wyman, M. (2017). A denitrifying community associated with a major, marine nitrogen fixer. *Environ. Microbiol.* 19, 4978–4992. doi: 10.1111/1462-2920.14007

Cosquer, A., Pichereau, V., Pocard, J. A., Minet, J., Cormier, M., and Bernard, T. (1999). Nanomolar levels of dimethylsulfoniopropionate, dimethylsulfonioacetate, and glycine betaine are sufficient to confer osmoprotection to Escherichia coli. *Appl. Environ. Microbiol.* 65, 3304–3311. doi: 10.1128/AEM.65.8.3304-3311.1999

Curson, A. R., Liu, J., Martínez, A. B., Green, R. T., Chan, Y., Carrión, O., et al. (2017). Dimethylsulfoniopropionate biosynthesis in marine bacteria and identification of the key gene in this process. *Nat. Microbiol.* 2:e17009. doi: 10.1038/nmicrobiol.2017.9

Curson, A. R., Todd, J. D., Sullivan, M. J., and Johnston, A. W. (2011). Catabolism of dimethylsulphoniopropionate: microorganisms, enzymes and genes. *Nat. Rev. Microbiol.* 9, 849–859. doi: 10.1038/nrmicro2653

DeLong, E. F., Franks, D. G., and Yayanos, A. A. (1997). Evolutionary relationships of cultivated psychrophilic and barophilic deep-sea bacteria. *Appl. Environ. Microbiol.* 63, 2105–2108. doi: 10.1128/aem.63.5.2105-2108.1997

DeLong, E. F., Preston, C. M., Mincer, T., Rich, V., Hallam, S. J., Frigaard, N. U., et al. (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science* 311, 496–503. doi: 10.1126/science.1120250

Dylan, T. Y. L. E. R., Helinski, D. R., and Ditta, G. S. (1990). Hypoosmotic adaptation in rhizobium meliloti requires beta-(1----2)-glucan. *J. Bacteriol.* 172, 1400–1408. doi: 10.1128/jb.172.3.1400-1408.1990

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565

Galinski, E. A. (1995). Osmoadaptation in bacteria. *Adv. Microb. Physiol.* 37, 273–328. doi: 10.1016/S0065-2911(08)60148-4

Gao, Z. M., Huang, J. M., Cui, G. J., Li, W. L., Li, J., Wei, Z. F., et al. (2019). *In situ* meta-omic insights into the community compositions and ecological roles of hadal microbes in the Mariana trench. *Environ. Microbiol.* 21, 4092–4108. doi: 10.1111/1462-2920.14759

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., and Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91. doi: 10.1099/ijs.0.64483-0

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47, D309–D314. doi: 10.1093/nar/gky1085

Hunt, M., Kikuchi, T., Sanders, M., Newbold, C., Berriman, M., and Otto, T. D. (2013). REAPR: a universal tool for genome assembly evaluation. *Genome Biol.* 14:R47. doi: 10.1186/gb-2013-14-5-r47

Jamieson, A. J., Fujii, T., Mayor, D. J., Solan, M., and Priede, I. G. (2010). Hadal trenches: the ecology of the deepest places on earth. *Trends Ecol. Evol.* 25, 190–197. doi: 10.1016/j.tree.2009.09.009

Kato, C., Li, L., Nogi, Y., Nakamura, Y., Tamaoka, J., and Horikoshi, K. (1998). Extremely barophilic bacteria isolated from the Mariana trench, challenger deep, at a depth of 11,000 meters. *Appl. Environ. Microbiol.* 64, 1510–1513. doi: 10.1128/AEM.64.4.1510-1513.1998

Katoh, K., and Standley, D. M. (2013). DMAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kerovuo, J., Reinikainen, T., Nyyssölä, A., Kaukinen, P., and von Weymarn, N. (2000). Extreme halophiles synthesize betaine from glycine by methylation. *J. Biol. Chem.* 275, 22196–22201. doi: 10.1074/jbc.M910111199

Kiene, R. P., Linn, L. J., and Bruton, J. A. (2000). New and important roles for DMSP in marine microbial communities. *J. Sea Res.* 43, 209–224. doi: 10.1016/S1385-1101(00)00023-X

King, G. M. (2003). Molecular and culture-based analyses of aerobic carbon monoxide oxidizer diversity. *Appl. Environ. Microbiol.* 69, 7257–7265. doi: 10.1128/AEM.69.12.7257-7265.2003

King, G. M., and Weber, C. F. (2007). Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nat. Rev. Microbiol.* 5, 107–118. doi: 10.1038/nrmicro1595

Kline, K. A., Dodson, K. W., Caparon, M. G., and Hultgren, S. J. (2010). A tale of two pili: assembly and function of pili in bacteria. *Trends Microbiol.* 18, 224–232. doi: 10.1016/j.tim.2010.03.002

Koglin, A., and Walsh, C. T. (2009). Structural insights into nonribosomal peptide enzymatic assembly lines. *Nat. Prod. Rep.* 26, 987–1000. doi: 10.1039/b904543k

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly *via* adaptive k-mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi: 10.1101/gr.215087.116

Kusube, M., Kyaw, T. S., Tanikawa, K., Chastain, R. A., Hardy, K. M., Cameron, J., et al. (2017). *Colwellia* marinimaniae sp. nov., a hyperpiezophilic species isolated from an amphipod within the challenger deep, Mariana trench. *Int. J. Syst. Evol. Microbiol.* 67, 824–831. doi: 10.1099/ijsem.0.001671

Lamble, H. J., Theodossis, A., Milburn, C. C., Taylor, G. L., Bull, S. D., Hough, D. W., et al. (2005). Promiscuity in the part-phosphorylative Entner–Doudoroff pathway of the archaeon Sulfolobus solfataricus. *FEBS Lett.* 579, 6865–6869. doi: 10.1016/j.febslet.2005.11.028

León-Zayas, R., Novotny, M., Podell, S., Shepard, C. M., Berkenpas, E., Nikolenko, S., et al. (2015). Single cells within the Puerto Rico trench suggest hadal adaptation of microbial lineages. *Appl. Environ. Microbiol.* 81, 8265–8276. doi: 10.1128/AEM.01659-15

Liu, J., Zheng, Y., Lin, H., Wang, X., Li, M., Liu, Y., et al. (2019). Proliferation of hydrocarbon-degrading microbes at the bottom of the Mariana trench. *Microbiome* 7:47. doi: 10.1186/s40168-019-0652-3

Llosa, M., Gomis-Rüth, F. X., Coll, M., and Cruz, F. D. L. (2002). Bacterial conjugation: a two-step mechanism for DNA transport. *Mol. Microbiol.* 45, 1–8. doi: 10.1046/j.1365-2958.2002.03014.x

Maguire, F., Jia, B., Gray, K. L., Lau, W. Y. V., Beiko, R. G., and Brinkman, F. S. (2020). Metagenome-assembled genome binning methods with short reads disproportionately fail for plasmids and genomic islands. *Microb. Genom.* 6. [Epub ahead of Preprint. doi: 10.1099/mgen.0.000436

Maier, B., and Wong, G. C. (2015). How bacteria use type IV pili machinery on surfaces. *Trends Microbiol.* 23, 775–788. doi: 10.1016/j.tim.2015.09.002

Maier, B., and Wong, G. C. (2015). How bacteria use type IV pili machinery on surfaces. *Trends Microbiol.* 23, 775–788. doi: 10.1016/j.tim.2015.09.002

Marmur, J. (1961). A procedure for the isolation of deoxyribonucleic acid from micro-organisms. *J. Mol. Biol.* 3:208-IN1. doi: 10.1016/S0022-2836(61)80047-8

Martin-Cuadrado, A. B., Ghai, R., Gonzaga, A., and Rodriguez-Valera, F. (2009). CO dehydrogenase genes found in metagenomic fosmid clones from the deep Mediterranean Sea. *Appl. Environ. Microbiol.* 75, 7436–7444. doi: 10.1128/AEM.01283-09

Mestre, M., Ruiz-González, C., Logares, R., Duarte, C. M., Gasol, J. M., and Sala, M. M. (2018). Sinking particles promote vertical connectivity in the ocean microbiome. *Proc. Natl. Acad. Sci. U. S. A.* 115, E6799–E6807. doi: 10.1073/pnas.1802470115

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015

Nunoura, T., Hirai, M., Yoshida-Takashima, Y., Nishizawa, M., Kawagucci, S., Yokokawa, T., et al. (2016). Distribution and niche separation of planktonic microbial communities in the water columns from the surface to the hadal waters of the Japan trench under the Eutrophic Ocean. *Front. Microbiol.* 7:1261. doi: 10.3389/fmicb.2016.01261

Nunoura, T., Takaki, Y., Hirai, M., Shimamura, S., Makabe, A., Koide, O., et al. (2015). Hadal biosphere: insight into the microbial ecosystem in the deepest ocean on earth. *Proc. Natl. Acad. Sci. U. S. A.* 112, E1230–E1236. doi: 10.1073/pnas.1421816112

Pathom-Aree, W., Stach, J. E., Ward, A. C., Horikoshi, K., Bull, A. T., and Goodfellow, M. (2006). Diversity of *actinomycetes* isolated from challenger deep sediment (10,898 m) from the Mariana trench. *Extremophiles* 10, 181–189. doi: 10.1007/s00792-005-0482-z

Peoples, L. M., Kyaw, T. S., Ugalde, J. A., Mullane, K. K., Chastain, R. A., Yayanos, A. A., et al. (2020). Distinctive gene and protein characteristics of extremely piezophilic *Colwellia*. *BMC Genomics* 21, 692–618. doi: 10.1186/s12864-020-07102-y

Ranwez, V., Douzery, E. J., Cambon, C., Chantret, N., and Delsuc, F. (2018). MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. *Mol. Biol. Evol.* 35, 2582–2584. doi: 10.1093/molbev/msy159

Reher, M., and Schönheit, P. (2006). Glyceraldehyde dehydrogenases from the thermoacidophilic euryarchaeota *Picrophilus torridus* and *Thermoplasma acidophilum*, key enzymes of the non-phosphorylative Entner–Doudoroff pathway, constitute a novel enzyme family within the aldehyde dehydrogenase superfamily. *FEBS Lett.* 580, 1198–1204. doi: 10.1016/j.febslet.2006.01.029

Reva, O., and Tümmler, B. (2008). Think big–giant genes in bacteria. *Environ. Microbiol.* 10, 768–777. doi: 10.1111/j.1462-2920.2007.01500.x

Richter, M., Rosselló-Móra, R., Oliver Glöckner, F., and Peplies, J. (2016). JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 32, 929–931. doi: 10.1093/bioinformatics/btv681

Sagot, B., Gaysinski, M., Mehiri, M., Guigonis, J. M., Le Rudulier, D., and Alloing, G. (2010). Osmotically induced synthesis of the dipeptide N-acetylglutaminylglutamine amide is mediated by a new pathway conserved among bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 107, 12652–12657. doi: 10.1073/pnas.1003063107

Schwarz, J. R., and Colwell, R. R. (1975). Macromolecular Biosynthesis in *Pseudomonas bathycetes* at Deep-Sea Pressure and Temperature, Session 104, Abstr. K. 94. In *75th annual meeting of the American Society for Microbiology*.

Simonato, F., Campanaro, S., Lauro, F. M., Vezzi, A., D'Angelo, M., Vitulo, N., et al. (2006). Piezophilic adaptation: a genomic point of view. *J. Biotechnol.* 126, 11–25. doi: 10.1016/j.jbiotec.2006.03.038

Tabita, F. R., Satagopan, S., Hanson, T. E., Kreel, N. E., and Scott, S. S. (2008). Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *J. Exp. Bot.* 59, 1515–1524. doi: 10.1093/jxb/erm361

Takami, H., Nishi, S., Lu, J., Shimamura, S., and Takaki, Y. (2004). Genomic characterization of thermophilic *Geobacillus* species isolated from the deepest sea mud of the Mariana trench. *Extremophiles* 8, 351–356. doi: 10.1007/s00792-004-0394-3

Vezzi, A., Campanaro, S., D'angelo, M., Simonato, F., Vitulo, N., Lauro, F. M., et al. (2005). Life at depth: *Photobacterium profundum* genome sequence and expression analysis. *Science* 307, 1459–1461. doi: 10.1126/science.1103341

Weber, C. F., and King, G. M. (2007). Physiological, ecological, and phylogenetic characterization of Stappia, a marine CO-oxidizing bacterial genus. *Appl. Environ. Microbiol.* 73, 1266–1276. doi: 10.1128/AEM.01724-06

Wood, J. M. (2011). Bacterial osmoregulation: a paradigm for the study of cellular homeostasis. *Annu. Rev. Microbiol.* 65, 215–238. doi: 10.1146/annurev-micro-090110-102815

Wyman, M., Hodgson, S., and Bird, C. (2013). Denitrifying *alphaproteobacteria* from the Arabian Sea that express nosZ, the gene encoding nitrous oxide reductase, in oxic and suboxic waters. *Appl. Environ. Microbiol.* 79, 2670–2681. doi: 10.1128/AEM.03705-12

Yang, Y., Jiang, X., Chai, B., Ma, L., Li, B., Zhang, A., et al. (2016). ARGs-OAP: online analysis pipeline for antibiotic resistance genes detection from metagenomic data using an integrated structured ARG-database. *Bioinformatics* 32, 2346–2351. doi: 10.1093/bioinformatics/btw136

Yang, Y., Li, B., Ju, F., and Zhang, T. (2013). Exploring variation of antibiotic resistance genes in activated sludge over a four-year period through a metagenomic approach. *Environ. Sci. Technol.* 47, 10197–10205. doi: 10.1021/es4017365

Yayanos, A. A., Dietz, A. S., and Van Boxtel, R. (1981). Obligately barophilic bacterium from the Mariana trench. *Proc. Natl. Acad. Sci. U. S. A.* 78, 5212–5215. doi: 10.1073/pnas.78.8.5212

Yoch, D. C. (2002). Dimethylsulfoniopropionate: its sources, role in the marine food web, and biological degradation to dimethylsulfide. *Appl. Environ. Microbiol.* 68, 5804–5815. doi: 10.1128/AEM.68.12.5804-5815.2002

Zhang, W. J., Cui, X. H., Chen, L. H., Yang, J., Li, X. G., Zhang, C., et al. (2019a). Complete genome sequence of *Shewanella benthica* DB21MT-2, an obligate piezophilic bacterium isolated from the deepest Mariana trench sediment. *Mar. Genomics* 44, 52–56. doi: 10.1016/j.margen.2018.09.001

Zhang, X.-H., Liu, J., Liu, J., Yang, G., Xue, C. X., Curson, A. R., et al. (2019b). Biogenic production of DMSP and its degradation to DMS—their roles in the global sulfur cycle. *Sci. China Life Sci.* 62, 1296–1319. doi: 10.1007/s11427-018-9524-y

Zhao, X., Liu, J., Zhou, S., Zheng, Y., Wu, Y., Kogure, K., et al. (2020). Diversity of culturable heterotrophic bacteria from the Mariana trench and their ability to degrade macromolecules. *Mar. Life sci. Technol.* 2, 181–193. doi: 10.1007/s42995-020-00027-1

Zheng, Y., Wang, J., Zhou, S., Zhang, Y., Liu, J., Xue, C. X., et al. (2020). Bacteria are important dimethylsulfoniopropionate producers in marine aphotic and high-pressure environments. *Nat. Commun.* 11:4658. doi: 10.1038/s41467-020-18434-4

Zhong, H., Lehtovirta-Morley, L., Liu, J., Zheng, Y., Lin, H., Song, D., et al. (2020). Novel insights into the *Thaumarchaeota* in the deepest oceans: their metabolism and potential adaptation mechanisms. *Microbiome* 8:78. doi: 10.1186/s40168-020-00849-2

Zhou, S., Ren, Q., Li, Y., Liu, J., Wang, X., Wu, Y., et al. (2018). Abyssibacter profundi gen. Nov., sp. nov., a marine bacterium isolated from seawater of the Mariana trench. *Int. J. Syst. Evol. Microbiol.* 68, 3424–3429. doi: 10.1099/ijsem.0.002999LEGEND