AEM
Journals.ASM.org

Single Cells within the Puerto Rico Trench Suggest Hadal Adaptation of Microbial Lineages

Rosa León-Zayas,[a] Mark Novotny,[b] Sheila Podell,[a] Charles M. Shepard,[c] Eric Berkenpas,[c] Sergey Nikolenko,[d,e] Pavel Pevzner,[d,f] Roger S. Lasken,[b] Douglas H. Bartlett[a]

Marine Biology Research Division, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA[a]; Microbial and Environmental Genomics, J. Craig Venter Institute, San Diego, California, USA[b]; National Geographic Society, Washington, DC, USA[c]; Algorithmic Biology Laboratory, St. Petersburg Academic University, Russian Academy of Sciences, St. Petersburg, Russia[d]; National Research Institute Higher School of Economics, St. Petersburg, Russia[e]; Computer Science and Engineering, University of California San Diego, La Jolla, California, USA[f]

Hadal ecosystems are found at a depth of 6,000 m below sea level and below, occupying less than 1% of the total area of the ocean. The microbial communities and metabolic potential in these ecosystems are largely uncharacterized. Here, we present four single amplified genomes (SAGs) obtained from 8,219 m below the sea surface within the hadal ecosystem of the Puerto Rico Trench (PRT). These SAGs are derived from members of deep-sea clades, including the *Thaumarchaeota* and SAR11 clade, and two are related to previously isolated piezophilic (high-pressure-adapted) microorganisms. In order to identify genes that might play a role in adaptation to deep-sea environments, comparative analyses were performed with genomes from closely related shallow-water microbes. The archaeal SAG possesses genes associated with mixotrophy, including lipoylation and the glycine cleavage pathway. The SAR11 SAG encodes glycolytic enzymes previously reported to be missing from this abundant and cosmopolitan group. The other SAGs, which are related to piezophilic isolates, possess genes that may supplement energy demands through the oxidation of hydrogen or the reduction of nitrous oxide. We found evidence for potential trench-specific gene distributions, as several SAG genes were observed only in a PRT metagenome and not in shallower deep-sea metagenomes. These results illustrate new ecotype features that might perform important roles in the adaptation of microorganisms to life in hadal environments.

Little is known about the microbial communities in the deepest ocean environment, the hadal zone, which is located at a depth of 6,000 m below sea level and below. Most hadal zones are true trenches, lying within convergent margins where an oceanic plate is subducted below an oceanic or continental plate (1). Extreme environmental conditions, such as a lack of sunlight, magmacrustal physical and chemical interactions, near-freezing temperatures, and high pressure, give rise to unique ecosystems with distinct biological diversity (2). Hadal environments include specialized megafauna that are distinct from their shallower deep-sea relatives, such as a ubiquitous scavenger, the lysianassoid amphipod (3–5). An investigation of deep-sea dwelling biota and their associated microbial communities might provide a better understanding of the hadal environment and the specific metabolic adaptations needed for success in this biome.

One required adaptation is the ability to grow under elevated pressure conditions. Piezophilic microorganisms possessing optimal growth rates at pressures above atmospheric pressure are present in trenches and other deep-sea habitats. Investigations of a few cultured piezophiles have indicated that they respond to pressure changes by altering their transcriptome and proteome, membrane fatty acids, respiratory system, and osmolytes (6–12). The increased production of unsaturated fatty acids has been demonstrated to be required for high-pressure growth (13–15). Additional putative adaptations have been inferred using metagenomic studies of 3- to 6-km-depth deep-ocean environments to examine the metabolic processes of as-yet-uncultivated microbes on a larger scale (16–20). For example, it has been observed that deep-sea metagenomes contain an overabundance of *cox* genes, suggesting that many deep-sea bacteria may be capable of using the aerobic oxidation of carbon monoxide as an additional energy source (16, 18, 20). The Puerto Rico Trench (PRT) metagenome is thus far the only hadal metagenome (16). Compared to surface metagenomes, the PRT metagenome has an overabundance of genes associated with signal transduction mechanisms, particularly those associated with internal sensing of redox potential and oxygen and genes encoding sulfatases for the degradation of complex polysaccharides. Inorganic ion transport and metabolism, along with transporters encoding outer membrane porins and genes involved in heavy metal efflux, are also abundant in the PRT metagenome (16).

In recent years, the study of uncultivated microbial communities has benefitted from the evolution of single-cell genomics techniques (21, 22). This approach was employed in the examination of four single-cell genomes derived from cells recovered in the PRT. They belonged to the *Gammaproteobacteria*, *Alphaproteobacteria*, *Bacteriodetes*, and *Planctomycetes* (16). The data obtained were in agreement with findings in the PRT metagenome highlighted by the presence of metal efflux systems, tripartite ATP-independent periplasmic (TRAP) and ABC-like transporters, and

Received 18 May 2015 **Accepted** 12 September 2015

Accepted manuscript posted online 18 September 2015

Citation León-Zayas R, Novotny M, Podell S, Shepard CM, Berkenpas E, Nikolenko S, Pevzner P, Lasken RS, Bartlett DH. 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.

Editor: K. E. Wommack

Address correspondence to Douglas H. Bartlett, dbartlett@ucsd.edu.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.01659-15.

sulfatases (16). The described metabolic processes harbored within these single cells provided an initial look into hadal-zone-associated metabolism.

Here, we describe single-cell genomes derived from newly collected seawater samples and from amphipod-associated microbes at a depth of >8,000 m within the PRT. The genomes included are those derived from cells belonging to marine group I (MGI) *Thaumarchaeota*, the SAR11 clade, and two other proteobacteria associated with previously cultured piezophilic organisms from the genus *Psychromonas* and the order *Rhodobacterales*. These genomes are representative of deep-sea clades known previously only through 16S rRNA gene surveys. A comparison of the deep-sea single amplified genome (SAG) sequences to related surface-associated microbial genome sequences provided clues to those metabolic characteristics harbored in the SAGs that may contribute to the adaptation of microbes to the hadal environment.

MATERIALS AND METHODS

Collection and sorting. Seawater and amphipods were collected 3.3 m off the seafloor using a free-falling vehicle (FFV) (see Fig. S1 in the supplemental material) deployed during November 2010 aboard the *Makai* (50-ft catamaran) over a water column depth of 8,219 m (±66 m) within the PRT (19°46.022′N, 66°55.432′W). Microbial samples were collected with a pair of baited 30-liter Niskin water sampling bottles. Recovered seawater and amphipods were placed inside polyethylene bags, pressurized to 62 MPa (9,000 lb/in.^{2}), and held at a temperature of 4°C. The samples were transferred to the J. Craig Venter Institute (JCVI) for single-cell sorting. The collected amphipod was homogenized using an autoclaved pestle in a microcentrifuge tube. The sample was filtered and stained with SYBR green I fluorescent dye (Invitrogen, Carlsbad, CA), sorted using a cooled FACSAria II flow cytometer (Becton, Dickinson and Company, Franklin Lakes, NJ), and stored at −80°C for later processing.

Genome amplification and sequencing. Genomic material was amplified using multiple displacement amplification (MDA) in a 384-well format using a GenomiPhi kit (GE Healthcare, Waukesha, WI) and a custom BioCel robotic system (Agilent Technologies, Santa Clara, CA), as described by McLean et al. (23). 16S rRNA genes were PCR amplified and cleaned, and amplicons were sent for Sanger sequencing at the Joint Technology Center (JTC) (J. Craig Venter Institute, Rockville, MD). The resulting 16S rRNA gene sequences were evaluated for evidence of contaminated sequences, and those were removed from consideration for whole-genome sequencing. 16S rRNA gene sequences were compared to sequences in the NCBI nr/nt database using BLASTN (24), and organisms of interest were selected based on their phylogenetic novelty. DNA recovered from 40 cells was prepared for whole-genome sequencing via the Illumina HiSeq 2000 platform. For comparison and validation, an *Escherichia coli* sample subjected to MDA was processed along with all other selected genomes (data not shown). Libraries were prepared using the multiple-barcode technology of the Nextera DNA sample prep kit (Illumina, San Diego, CA) and sent to the JTC for sequencing.

Assembly, annotation, and genome completion. Sequences were assembled using SPAdes 2.3.0 (78). Four genomes were selected for further processing and annotation based on the amount of genome sequence recovered and environmental relevance. The assembled genomes were uploaded to the IMG/ER platform (https://img.jgi.doe.gov/cgi-bin/mer/main.cgi) (25) for genome annotation.

The 16S rRNA gene sequences recovered from each SAG were analyzed by BLASTN against sequences in the NCBI nr/nt database (24). Sequence matches at a 95% similarity cutoff were extracted and used for phylogenetic reconstruction; in some cases, other phylogenetically relevant species were added as references. All reference sequences extracted from the NCBI were also annotated for their associated environmental source, including water column depth, if available. Sequences were aligned with the SINA aligner (http://www.arb-silva.de/aligner/) (26), and a maximum-likelihood tree was created using FastTree (27).

Genome-encoded protein predictions were obtained from IMG/ER and classified taxonomically using the DarkHorse software, version 1.4 (29; http://darkhorse.ucsd.edu/). The DarkHorse results were used to identify potential contaminating sequences among SAG contigs, based on whether or not taxonomic lineages were associated with the JCVI in-house database of potential contaminants (29). Contigs identified as contaminants were removed from the analysis. Estimated genome completeness was calculated using the Human Microbiome Project protocol for bacterial genomes (http://hmpdacc.org/doc/sops/reference_genomes/metrics/Bacterial_CoreGenes_SOP.pdf) and the protocol of Podell et al. (30) for the archaeal genome. Genome estimation was performed with the above-mentioned protocols after verifying that the marker genes were also found within the comparison genomes. In the case of "*Candidatus* Pelagibacter ubique," it had only 65 out of 66 of the markers, so the estimation of genome completeness for PRT SAR11 SAG was calculated based on those 65 marker genes only.

SAG comparisons to known genomes and metagenomes. Functional comparisons were performed using the IMG/ER 4 platform and manually with BLAST searches of the nr database (25). A most closely related genome (MCRG) was assigned to each SAG based on predicted protein similarity. Using the MCRG of each SAG, side-by-side comparisons of sequences in the categories of the Clusters of Orthologous Groups (COGs) and of sequences of genes for proteins of the Pfam and KEGG pathways were performed. In addition, SAGs and respective MCRGs were evaluated using the FR-HIT read recruitment software, with default parameters (31), against 96 metagenomic sets from the Global Ocean Survey (GOS) (32, 33) and metagenomic samples from the Hawaii Ocean Time series (HOT) (17, 19) and Mediterranean Sea (18, 20) (see Table S1 in the supplemental material). Recruitments were normalized based on the SAG size and total number of reads of the analyzed metagenome. To obtain a more detailed representation of the PRT SAG genes in the metagenomes, best reciprocal BLAST searches (BRBs) were performed with these data. The sequences from these FR-HIT results that contained >60 bp in alignment and 75% sequence similarity for each specific SAG were selected. A small database that included each SAG and closely related genomes was generated. The FR-HIT-recovered sequences were compared to the sequences in the databases generated using BLASTN. Reads that preferentially matched the SAGs were considered true read recruitments, while reads that preferentially matched any of the other comparison genomes were excluded from the analysis. BRB hits were quantified and normalized as described above.

Nucleotide sequence accession numbers. These single-cell genome sequences have been deposited at DDBJ/EMBL/GenBank under the accession numbers JPUE00000000, JPUP00000000, JPUQ00000000, and JPUR00000000. The annotated genomes are also available in the IMG/ER platform with the genome identification numbers (IDs) 2518645502, 2518645503, 2518645501, and 2518645504.

RESULTS AND DISCUSSION

Phylogenetic placement and biogeography of single cells subjected to MDA. PRT trench seawater and amphipods were collected in baited traps at a depth of approximately 8,200 m below sea level. The amphipods belonged to the genus *Hirondellea* of the superfamily *Lysianassoidea* (J. Carvajal and G. Rouse, unpublished data). A total of 15,720 single cells were sorted from trench seawater and amphipod-associated microbial communities. The trench seawater had a notably low biomass, and as such, the trench amphipods endemic to the area were collected as additional sources of microbial biomass (34, 35). The DNA from 2,880 sorted cells was subjected to MDA, with 22% of the amplified DNA yielding positive amplification. The 16S rRNA gene sequences obtained from the MDA-treated single cells were examined, and any

TABLE 1 Genomic characterization of 4 PRT SAGs[a]

Genome name	Sequencing genome size (bp)	Estimated genome size (bp) (% difference from MCRG)	No. of genes	% genome completeness	G+C content (%)	No. of 16S rRNAs	No. of COGs categories	No. of transposases
PRT *Nitrosopumilus* SAG	660,313	917,101 (43)	832	72	35	1	682	1 (+)
PRT SAR11 SAG	700,642	973,113 (25)	704	72	30	1	899	0
PRT *Marinosulfonomonas* SAG	17,36,454	2,255,135 (34)	1,980	77	52	1	1,904	162 (+)
PRT *Psychromonas* SAG	920,308	1,559,844 (65)	975	59	38	3	1,125	25 (−)

[a] The number of genes, number of contigs, G+C content, number of COGs categories, and coding region percentage were obtained from the IMG Web application. The estimated genome size was calculated based on the percent completion and the size of the MCRG. The percentages of difference in genome size reflect the small sizes of the PRT SAGs relative to the sizes of the MCRGs. The sign (+ or −) in the "No. of transposases" column indicates whether the value for the SAG is more (+) or less (−) than that for its respective MCRG. All estimated genome sizes were smaller than those of the MCRGs (the percentages in the "Estimated genome size" column reflect the percent decrease relative to the MCRG size).

sequences related to *Pseudomonas* spp. or other lab contaminants were discarded. Remaining sequences of >1,000 bp included a large number of microbial phyla and classes, among them *Alphaproteobacteria*, *Betaproteobacteria*, *Gammaproteobacteria*, *Bacteriodetes*, *Fusobacteria*, *Firmicutes*, *Actinobacteria*, and *Acidobacteria* (see Fig. S2 in the supplemental material). Forty single cells were selected based on phylogenetic novelty for complete genome sequencing. Of these, four contained the greatest genome completeness and as such were selected for further annotation and analysis: two derived from cells present in undiluted trench seawater, PRT *Nitrosopumilus* and PRT SAR11; and two derived from amphipod-associated cells, PRT *Marinosulfonomonas* and PRT *Psychromonas* (Fig. 1).

The archaeal PRT *Nitrosopumilus* falls within the I.1a group of the *Thaumarchaeota*, the same group as *Nitrosopumilus maritimus* (36, 37). Most of the members of this clade have been recovered from sediment and water column samples at depths ranging from 3,500 m to 5,000 m below the sea surface (38) (Fig. 1). The other SAG collected from undiluted trench seawater is related to the SAR11 group. Previous studies have suggested that deep-sea SAR11 (770 m) falls within the SAR11 group 1c (39). In contrast, the PRT SAR11 SAG falls within SAR11 group II, a sister clade to the SAR11 group I (1a, b, and c), which includes the genus *Pelagibacter*. SAR11 group II is divided into two distinct clusters (40), with the SAR11 SAG falling within a subclade of group IIa (41). Although the majority of group IIa organisms are associated with shallow-water environments, the sequences in this specific subclade are predominantly derived from deeper marine environments ranging from 2,400 to 3,300 m below the sea surface (Fig. 1).

One of the amphipod-associated cells, PRT *Marinosulfonomonas*, clustered within a major clade of primarily uncultivated environmental samples isolated from deep trenches, hydrothermal vent plumes, and *Riftia* tube worms (42). PRT *Marinosulfonomonas* is closely related to the alphaproteobacterium *Rhodobacterales* strain PRT1, a piezophilic microorganism cultured from the PRT (Fig. 1) (43). PRT1 and PRT *Marinosulfonomonas* belong to a phylogenetically distinct clade dominated by deep-ocean bacteria within the *Roseobacter* lineage of the *Rhodobacterales*. The other amphipod-associated SAG, that of PRT *Psychromonas*, falls within the genus *Psychromonas* in the *Gammaproteobacteria*. It clusters together with the cultivated piezophile *Psychromonas* sp. strain CNPT3 and other microbes belonging to deep-ocean environments (e.g., trenches and whale falls [44]) (Fig. 1). CNPT3 was the first piezophile ever isolated, and it is known to possess various adaptations for growth at high pressure (45, 46). Members of all clades can clearly be assigned to deep-sea habitats.

Genomic characterization. The assembled genome sizes ranged in size from 0.6 Mbp to 1.7 Mbp. Genome completion varied from 59% to 77% (Table 1). Previous genome size calculations of cultured piezophiles, which are genomes retrieved from metagenomic and single-cell genomic analyses, have led to the conclusion that deep-sea microbes have larger genomes than those of their surface-dwelling relatives (8, 16, 17, 19, 39). This has been ascribed to the reduction in purifying selection encountered in deep-sea versus shallow-water habitats (19), although determining purifying selection for lineages that are not highly divergent can be misleading (47). However, all four of the SAGs described here have genome sizes smaller than their MCRGs (Table 1). Certainly, the four genomes examined in this study are too small in number to be a statistically significant data set; however, reduced genome sizes have been interpreted as reflecting genome streamlining and adaptation to oligotrophic environments (48). While the PRT is considered to be an oligotrophic environment (49), the incomplete SAGs and the methodological difficulty in accurately determining genome sizes from incomplete data do not allow further conclusions to be made.

General metabolic comparisons. To assess whether the PRT SAGs encode unique metabolic characteristics, proteins were assigned COGs, KEGG, and Pfam categories and used in comparisons between the SAGs and their associated MCRGs (Fig. 2; see also Table S2 in the supplemental material). For the sake of brevity, only the COGs categories are described below.

The PRT *Nitrosopumilus* SAG shares 358 of 406 of its annotated COGs with the thaumarchaeon *N. maritimus* SCM1. Among the shared COGs are several carbon fixation genes involved in the 3-hydroxypropionate-4-hydroxybutyrate pathway and general metabolic genes involved in cell functions. Forty-eight COGs unique to the SAG were distributed among 13 different categories. Approximately 30% of the 48 unique COGs were unknown (category S), followed by those for amino acid transport and metabolism (category E; 20%), energy production and conversion (category C; 7%), and posttranslational modification, protein turnover, and chaperones (category O; 7%) (Fig. 2). The PRT *Nitrosopumilus* SAG can be differentiated from the *N. maritimus* genome by the presence of unique sequences that may provide a selective advantage in the deep ocean, including enzymes for urea degradation, lipoic acid synthesis, glycine cleavage, and, remarkably, fatty acid synthesis (Table 2).

The PRT SAR11 SAG shares 496 of its 536 annotated COGs with "*Candidatus* Pelagibacter ubique" SAR11 HTCC1062. The 40 COGs unique to this SAG are distributed among 12 different categories (Fig. 2). These include novel metabolic properties not present in most other described SAR11 members, including genes

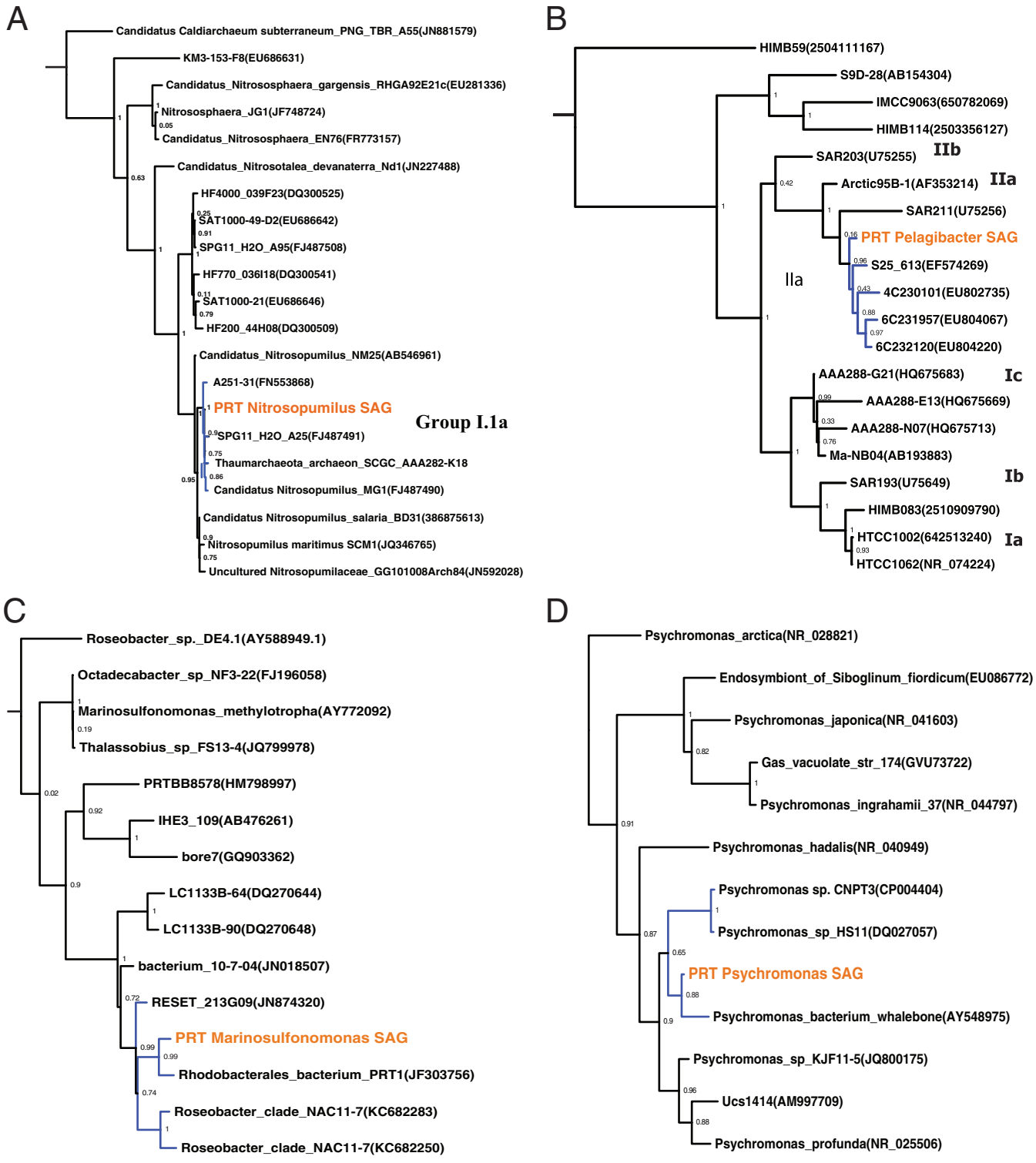

FIG 1 Phylogenetic trees (rooted to *E. coli* [not shown]) of 16S rRNA genes from the PRT *Nitrosopumilus*, PRT SAR11, PRT *Marinosulfonomonas*, and PRT *Psychromonas* SAGs. (A) PRT *Nitrosopumilus* (orange) falls within group I.1a in a cluster of other deep-sea-associated sequences (blue branches). (B) PRT SAR11 (orange) falls within group IIa in a cluster of other deep-sea-associated sequences (blue branches). (C) PRT *Marinosulfonomonas* (orange) falls within the *Roseobacter* group and clusters with other deep-sea-associated sequences (blue branches). (D) PRT *Psychromonas* (orange) groups closely with other isolated piezophiles and other deep-sea-associated sequences (blue branches).

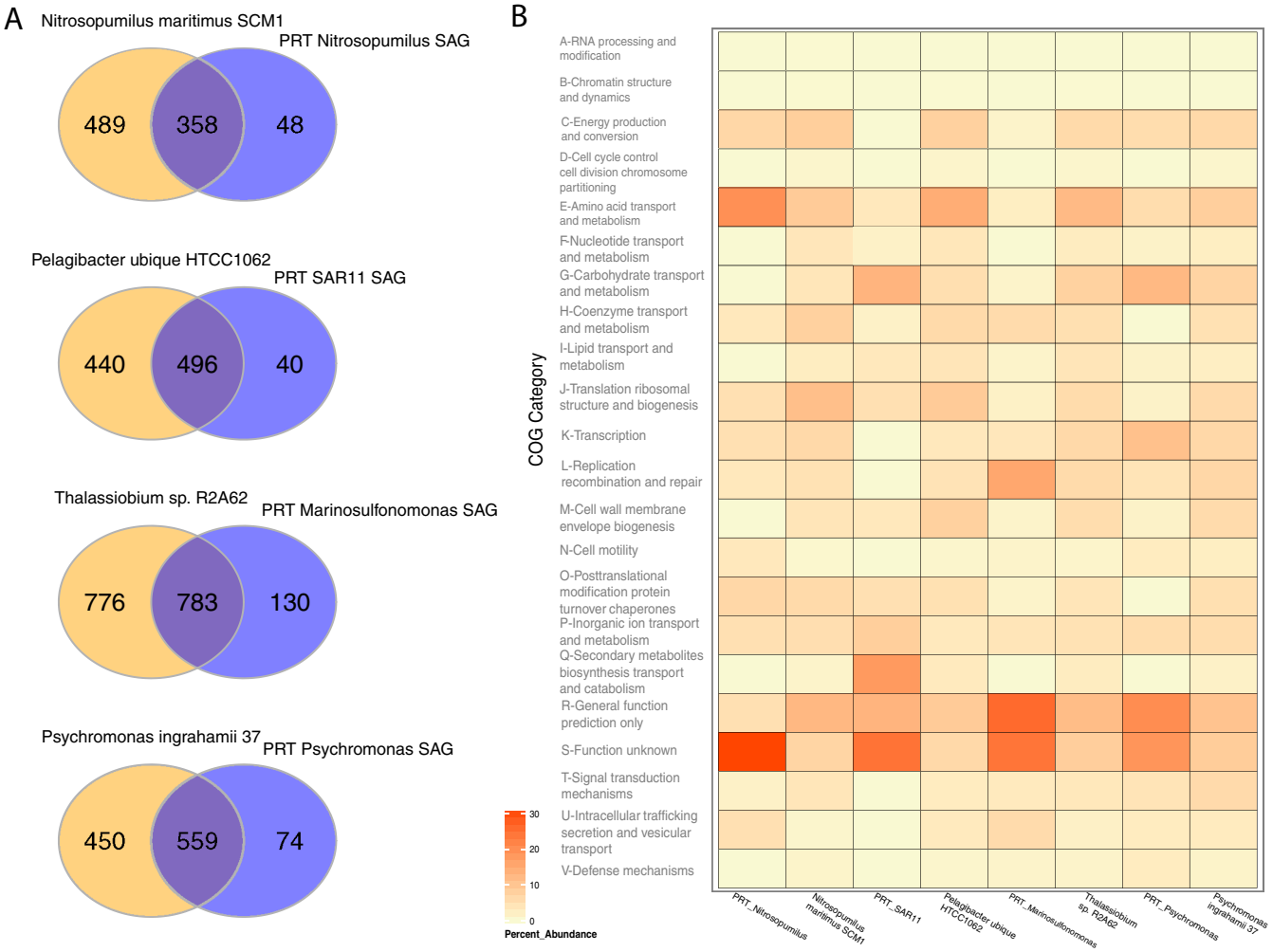

FIG 2 Comparison of SAGs to closely related genomes based on annotated COGs. (A) Side-by-side comparisons of PRT SAGs to MCRG. (B) Distribution of percentages of abundance of COGs categories found uniquely in PRT genomes versus the total COGs distribution of MCRG.

that encode the glycolytic enzymes phosphofructokinase and pyruvate kinase.

The PRT *Marinosulfonomonas* SAG exhibits extensive metabolic diversity, as is the case for the genomes of other members of the family *Rhodobacterales*. The PRT *Marinosulfonomonas* SAG shares 783 of 913 COGs with the alphaproteobacterium *Thalassiobium* R2A62. Among their shared metabolic properties are housekeeping functions associated with tRNA synthases, pilus synthesis and assembly, cellular shape, transport systems, and metabolic processes. One hundred thirty COGs are unique to the PRT *Marinosulfonomonas* SAG. These are distributed among the 19 different categories, with the most abundant having unknown function (category S) (39%) and general function prediction (category R) (15%) (Fig. 2).

The PRT *Psychromonas* SAG was compared with the surface water bacterium *Psychromonas ingrahamii* and the cultured piezophilic bacterium *Psychromonas* sp. strain CNPT3. The PRT *Psychromonas* SAG shares 559 of its 633 COGs with *P. ingrahamii* and 632 with CNPT3. Seventy-four COGs are unique to the SAG compared to *P. ingrahamii* and 31 compared to CNPT3 (Fig. 2). The PRT *Psychromonas* SAG shares a number of genes with CNPT3 that are not found in the shallow-water *P. ingrahamii*. This includes genes involved in motility and a number of transporters and permeases for iron, multidrug, sugar, and amino acid transport. These two microbes also share the ability to produce the periplasmic nitrate reductase system protein NapA and encode the carbon starvation protein CstA, which is suggested to be involved in peptide transport under stressed conditions (50).

Novel metabolic potential. (i) Lipoylation/glycine cleavage/ammonia acquisition. Characterized by their ability to autotrophically oxidize ammonia, members of the *Thaumarchaeota* have been suggested to play a major role in the nitrogen cycle, particularly in the deep ocean (19, 37, 51). PRT *Nitrosopumilus* is the deepest-dwelling member of the *Thaumarchaeota* to be studied at the genomic level in detail. The identification of genes associated with lipoylation, the glycine cleavage system (GCS), fatty acid metabolism, and lipid A biosynthesis implies that this archaeon contains properties not yet demonstrated in any thaumarchaeon (Table 2).

As part of the novel metabolic potential in the PRT *Nitrosopumilus* SAG, two genes are involved in lipoylation, lipoate-protein ligase A and lipoyl synthase genes. Lipoate or lipoic acid (LA)

TABLE 2 Unique metabolic properties of the SAG genomes

Organism and function[a]	Function ID	Top BLAST match[b]
PRT *Nitrosopumilus*		
FOG: PAS/PAC domain[c]	COG2202	Signal transduction histidine kinase, with PAS, phosphoacceptor, and ATP binding domain ("*Candidatus* Nitrososphaera gargensis")
NhaD-type sodium/proton antiporters	COG1055	Putative arsenical pump membrane protein (*Nitrososphaera viennensis* EN76)
Lipoate-protein ligase	COG0095	Lipoate-protein ligase A (*Aciduliprofundum* sp. strain MAR08-339)
Lipoate synthase	COG0320	Lipoyl synthase (*Corallococcus coralloides*)
Glycine cleavage system H protein	COG0509	Glycine cleavage system protein H (*Dictyoglomus thermophilum*)
Glycine dehydrogenase subunit 1	COG0403	Glycine dehydrogenase subunit 1 (*Chloroherpeton thalassium*)
Glycine dehydrogenase subunit 2	COG1003	Glycine dehydrogenase subunit 2 (*Carboxydothermus hydrogenoformans*)
Aminomethyltransferase	COG0404	Glycine cleavage system protein T (*Kosmotoga olearia*)
Dihydrolipoamide dehydrogenase	COG1249	Dihydrolipoamide dehydrogenase (*Thermotoga lettingae*)
Urease subunit gamma	COG0831	Urease subunit gamma (*Nitrosopumilus* sp. strain AR)
Urease accessory protein	COG2371	Urease accessory protein (*Cenarchaeum symbiosum*)
Urease accessory protein	COG0830	Urease accessory protein (*C. symbiosum*)
Urease accessory protein	COG0378	Urease accessory protein UreG (*C. symbiosum*)
Urease accessory protein	COG0829	Urease accessory protein UreD ("*Candidatus* Nitrososphaera gargensis")
Urease subunit alpha	COG0804	Urease subunit alpha (*C. symbiosum*)
Urease subunit beta	COG0832	Urease subunit beta (*N. viennensis* EN76)
Nitrogen regulatory protein P-II 2	KO:K04752	Hypothetical protein ("*Candidatus* Nitrosoarchaeum limnia")
UDP-3-*O*-[3-hydroxymyristoyl] glucosamine *N*-acyltransferase	COG1044	Acetyltransferase ("*Candidatus* Nitrosopumilus salaria")
3-Oxoacyl-[acyl-carrier-protein] synthase III	COG0332	3-Oxoacyl-ACP synthase (*Microcystis aeruginosa*)
Aquaporin-4	COG0580	Glycerol transporter ("*Candidatus* Nitrosoarchaeum limnia")
PRT SAR11		
6-Phosphofructokinase	COG0205	6-Phosphofructokinase (alphaproteobacterium HIMB59)
Pyruvate kinase	COG0469	Pyruvate kinase (*Hydrogenivirga* sp. strain 128-5-R1-1)
Taurine dioxygenase	COG2175	Taurine catabolism dioxygenase TauD ("*Candidatus* Pelagibacter sp." strain HTCC7211)
TRAP-type C_4-dicarboxylate transport system, large permease component	COG1593	TRAP transporter, DctM subunit (*Clostridium bolteae*)
TRAP-type C_4-dicarboxylate transport system, periplasmic component	COG1638	DctP family TRAP transporter solute receptor (*C. bolteae*)
TRAP-type C_4-dicarboxylate transport system, small permease component	COG3090	C_4-dicarboxylate ABC transporter permease (*Thermosinus carboxydivorans*)
TRAP-type uncharacterized transport system, periplasmic component	COG2358	C_4-dicarboxylate ABC transporter substrate-binding protein (alphaproteobacterium HIMB114)
TRAP-type uncharacterized transport system, fused permease components	COG4666	C_4-dicarboxylate ABC transporter (alphaproteobacterium HIMB114)
ABC-type Mn/Zn transport systems, ATPase component	COG1121	Zinc transporter (alphaproteobacterium HIMB5)
ABC-type sugar transport system, ATPase component	COG1129	Sugar ABC transporter ATP-binding protein (*Rhizobium* sp. strain CF142)
Ribose/xylose/arabinose/galactoside ABC-type transport systems, permease components	COG1172	Ribose/xylose/arabinose/galactoside ABC-type transport systems, permease components ("*Candidatus* Pelagibacter ubique" HIMB058)
ABC-type sugar transport system, periplasmic component	COG1879	LacI family transcriptional regulator (*Kiloniella laminariae*)
ABC-type Zn^{2+} transport system, periplasmic component/surface adhesin	COG4531	Zinc transporter ("*Candidatus* Pelagibacter sp." HTCC7211)
Maltose/maltodextrin transport system ATP-binding protein	COG3839	Sugar ABC transporter ATP-binding protein (*Rhizobium* sp. CF142)
Na^+/melibiose symporter and related transporters	COG2211	Symporter (alphaproteobacterium HIMB59)
3-Oxoacyl-[acyl-carrier-protein] synthase III	COG0332	Hypothetical protein (*Mariprofundus ferrooxydans*)
PRT *Marinosulfonomonas*		
Taurine dioxygenase	COG2175	Gamma-butyrobetaine dioxygenase (*Labrenzia* sp. strain DG1229)
Aquaporin Z	COG0580	Aquaporin (*Rhodobacter sphaeroides*)
Thiamine monophosphate synthase	COG0352	Thiamine-phosphate pyrophosphorylase (*Ahrensia* sp. strain 13_GOM-1096 m)
Thiamine biosynthesis protein, ThiC	COG0422	Phosphomethylpyrimidine synthase (*Pseudovibrio* sp. strain FO-BEG1)
Membrane-associated lipoprotein involved in thiamine biosynthesis	COG1477	Thiamine biosynthesis protein ApbE (*Pannonibacter phragmitetus*)
Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase	COG0351	Hydroxymethylpyrimidine kinase (*Sulfitobacter* sp. strain MM-124)
Thiamine biosynthesis, ThiG	KO:K03149	Thiazole synthase (*Sulfitobacter* sp. MM-124)
Biotin synthase and related enzymes	COG0502	Biotin synthase (*Ruegeria* sp. strain R11)
Dethiobiotin synthetase	COG0132	Dethiobiotin synthetase (*Phaeobacter gallaeciensis*)
Cobalamin biosynthesis protein, CbiG	COG2073	Precorrin-3B methylase (*Roseobacter* sp. strain AzwK-3b)
Nitrous oxide reductase	COG4263	Nitrous oxide reductase (*Ruegeria lacuscaerulensis*)
Regulator of nitric oxide reductase transcription	COG3901	FMN-binding protein (*Roseobacter* sp. strain SK209-2-6)
Collagenase U32	COG0826	Peptidase U32 (*Phaeobacter arcticus*)
Co/Zn/Cd cations	COG0053	ABC transporter permease (*Rhodobacter* sp. strain SW2)
Na^+/H^+ antiport	COG2111	Cation:proton antiporter (*Rhodobacteraceae* bacterium HTCC2150)
ABC-type transporters for many drugs	COG0842	ABC-type multidrug transport system, permease component (*Thalassobacter arenae*)
ABC-type transporters for dipeptide/oligopeptide/nickel	COG1124	Oligopeptide transport ATP-binding protein OppF (*T. arenae*)
ABC-type transporters for long-chain fatty acids	COG1133	Transporter ("*Candidatus* Pelagibacter sp." HTCC7211)
ABC-type transporters for 2-aminoethylphosphonate	COG1178	Iron ABC transporter permease (*Labrenzia* sp. DG1229)
PRT *Psychromonas*		
[NiFe]-hydrogenase I small subunit	COG1740	Quinone-reactive [NiFe]-hydrogenase small chain (*Shewanella halifaxensis*)
[NiFe]-hydrogenase I large subunit	COG0374	Hydrogenase 2 large subunit (*Shewanella frigidimarina*)
[NiFe]-hydrogenase I cytochrome *b* subunit	COG1969	Hydrogenase (*Shewanella loihica*)
Nitrate reductase cytochrome *c*-type subunit, NapB	COG3043	Nitrate reductase (*Psychromonas* sp. CNPT3)
Periplasmic nitrate reductase chaperone, NapD	COG3062	Sorbose reductase (*Psychromonas* sp. CNPT3)
Periplasmic nitrate reductase maturation protein NapF	COG1149	Ferredoxin (*Psychromonas* sp. CNPT3)
Trimethylamine *N*-oxide reductase system, NapE	COG4459	TorE protein (*Psychromonas* sp. CNPT3)
ABC-type multidrug transport system, permease component	COG0842	ABC transporter (*Psychromonas* sp. CNPT3)
Chemotaxis signal transduction protein	COG0835	Nitrate/nitrite sensor protein NarQ (*Psychromonas* sp. CNPT3)

[a] Metabolic potential is listed as novel when it is not found in the most closely related genome (MCRG).
[b] ACP, acyl carrier protein; FMN, flavin mononucleotide; ABC, ATP binding cassette; TRAP, tripartite ATP-independent periplasmic; UDP, uridine-diphosphate.
[c] Regulatory domains involved in signaling pathways.

is a highly conserved cofactor in the aerobic metabolism of 2-oxoacids and C_1 compounds (52). These enzymes catalyze the attachment of the lipoyl moiety to dihydrolipoyl acyltransferase (E2), which is required for the function of several key enzyme complexes in oxidative and one-carbon metabolism. In this case, E2 is missing from the PRT *Nitrosopumilus* SAG, although this might stem from the fact that the genome is only 72% complete. However, E2 is also absent from the genomes of most *Thermococcus* species (one of the two archaeal genera in which lipoylation has been studied). In the absence of E2, the glycine cleavage system protein H is used as a lipoylation target (53). The PRT *Nitrosopumilus* SAG has all the genes for GCS, suggesting that protein H is the likely target for lipoylation. The catabolism of glycine involves a reversible reaction whereby glycine is cleaved to carbon dioxide, ammonia, and a methylene group (—CH_2—). The methylene group is accepted by tetrahydrofolate (THF) to form 5,10-methylene-THF, which is involved in purine and methionine biosynthesis. The regeneration of THF produces NADH, which can be used directly to yield energy, and ammonia, which can be utilized for a variety of processes, including energy generation via ammonia oxidation. GCS has been studied only in hyperthermophilic and halophilic archaea (54, 55).

It has been suggested that ammonia-oxidizing *Thaumarchaeota* utilize ureases to catalyze the degradation of urea to carbon dioxide and ammonia when environmental ammonia concentrations are low (56). The PRT *Nitrosopumilus* SAG encodes all components of the urease enzyme and urease accessory proteins. In addition, the *Nitrosopumilus* SAG also possesses a *glnK* gene, which is a well-known regulator of ammonium transport and incorporation in *Eukarya*, *Bacteria*, and selected groups of *Archaea* (within the *Euryarchaeota*) (57).

(ii) Fatty acid metabolism and lipid synthesis. The *Nitrosopumilus* SAG has a number of genes associated with fatty acid and lipid synthesis, including 3-oxoacyl-[acyl-carrier-protein] synthase III (KAS III). No other archaeal KASIII enzyme has been reported (58). In a comparison with all public genomes available in IMG, the top hit was to a hypothetical protein from a single-cell-genome sequence of an environmental *Thaumarchaeota* archaeon (SCGC AAA282-K18; R. Stepanuskas, unpublished data). The *Nitrosopumilus* SAG also encodes an acetyl coenzyme A (acetyl-CoA) carboxylase, which catalyzes the conversion of acetyl-CoA to malonyl-CoA, providing the substrates needed (acetyl-CoA and malonyl-CoA) for KASIII to perform the first condensation step in fatty acid synthesis.

The *Nitrosopumilus* SAG contained an acetyltransferase involved in lipid A synthesis, a structural component of lipopolysaccharide (Table 2). Given that archaeal cells are not known to produce lipopolysaccharides, it is difficult to speculate on the significance of finding enzymes belonging to lipid A biosynthesis.

(iii) Carbon and energy acquisition. Many bacteria and archaea have the ability to reduce nitrous oxide (N_2O) to nitrogen gas (N_2) without performing the complete denitrification pathway (59). Thus, since the PRT *Marinosulfonomonas* SAG contains genes associated with N_2O reduction (e.g., *nosZ*), it is possible that it employs N_2O in the oxidation of organic matter (Table 2). It has been proposed that nitrification leads to N_2O production in sinking particles (60). The PRT *Marinosulfonomonas*, which was recovered from an amphipod, may make use of the N_2O produced from incomplete denitrification processes associated with the amphipod microbiome, sediments, or particles ingested by amphipods.

Although other members of the SAR11 group have been shown to contain glycolytic operons (61), not all members encode phosphofructokinase and pyruvate kinase, key enzymes encoded in the PRT SAR11 SAG (Table 2). The only other described SAR11 genome that encodes these enzymes is the alphaproteobacterium HIMB59 (62, 63). More-detailed analysis of these two enzymes must be done to clearly understand their evolution and distribution within the SAR11. Nevertheless, the presence of these enzymes in the PRT SAR11 SAG together with the numerous ABC-type sugar transporters suggests that it may be capable of glycolysis via the Embden-Meyerhof-Parnas pathway, utilizing sugar substrates for carbon acquisition and energy production.

Analysis of the metabolic potential of the PRT *Psychromonas* SAG provided a unique opportunity to address functions shared among related piezophiles isolated from different ocean basins. Periplasmic nitrate reductase (NapBDEF) genes associated with carbon acquisition and energy generation are present in the PRT *Psychromonas* SAG and its deep-sea MCRG but not in its surface MCRG. Nitrate reductase initiates aerobic ammonification and is thought to be involved in the disposal of excess reductant power and as an electron sink to regenerate NAD^+ (64). Among the unique metabolic properties of the PRT *Psychromonas* SAG is a gene that codes for a [NiFe]-hydrogenase I enzyme. This is a membrane-bound protein that links H_2 oxidation to anaerobic or aerobic respiration, with the recovery of energy occurring via proton motive force (65). [NiFe]-Hydrogenases have been studied extensively in organisms associated with hydrothermal vents and anaerobic systems, but less is known about microorganisms inhabiting oxygenated marine environments (66).

(iv) Sulfur metabolism. Some microorganisms, including some members of the SAR11 clade, lack the ability to utilize sulfate and must acquire their sulfur using different sulfur-containing substrates (67). Two of the PRT SAGs, PRT SAR11 and PRT *Marinosulfonomonas*, have a taurine dioxygenase (TauD) gene. In *E. coli*, TauD is used to provide an alternate sulfur source under sulfur-deficient conditions (68). The taurine degradation pathway produces sulfite for cysteine biosynthesis. It is possible that the PRT *Marinosulfonomonas* SAG and PRT SAR11 SAG utilize taurine as a source of reduced sulfur.

(v) Osmoregulation. Genes coding for proteins associated with osmotic regulation are also present in several of the SAGs. For example, the NhaD-type sodium/proton antiporters, which have been proposed to serve osmoregulatory purposes (69), are found in the *Nitrosopumilus* SAG. Furthermore, the *Nitrosopumilus* and *Marinosulfonomonas* SAGs encode aquaporins. Aquaporins are known to be important in osmotic pressure adaptation by effluxing water from cells exposed to hypotonic environments. Aquaporins have also been suggested to be especially useful for the retention of small-molecule compatible solutes, like urea, glycerol, and glucose (70). Aquaporin-4, encoded by the PRT *Nitrosopumilus* SAG, has been described only as a mammalian protein, while the *Marinosulfonomonas* SAG encodes an aquaporin Z (AqpZ). The role of AqpZ in free-living marine microorganisms has not been fully characterized. This channel is selectively permeable to water, has a role in both the short-term and long-term osmoregulatory responses, and is required by rapidly growing cells. Given that osmotic pressure and hydrostatic pressure can have opposing effects on macromolecules (71), and deep-ocean

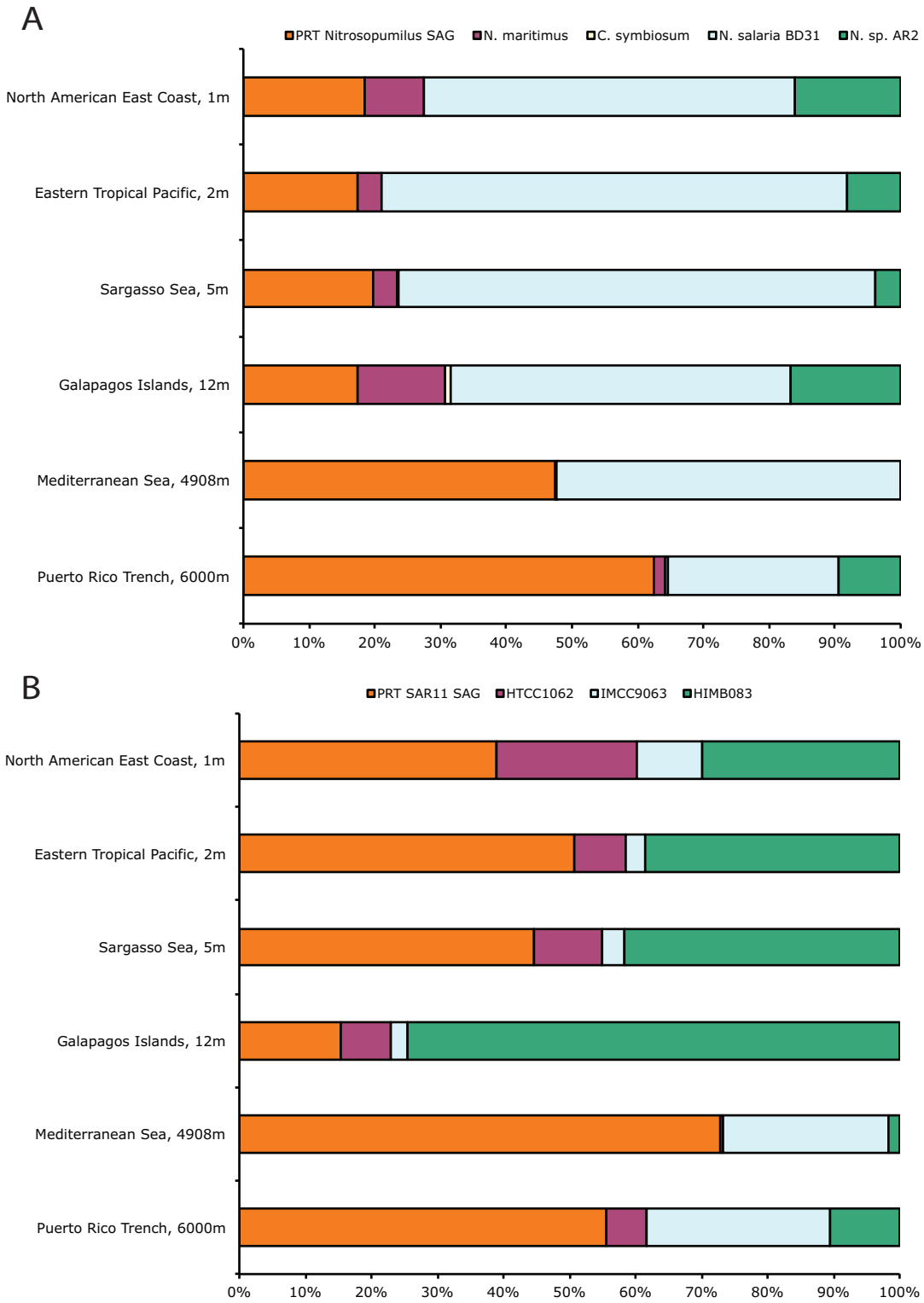

FIG 3 Best reciprocal BLAST analysis for PRT *Nitrosopumilus* SAG and PRT SAR11 SAG. The relative abundance of the PRT *Nitrosopumilus* SAG and the PRT SAR11 SAG as assessed by best reciprocal BLAST (BRB) analysis shows a trend preferentially recruiting reads from metagenome data sets associated with deep-sea environments compared to metagenomes from surface waters. Metagenomes are displayed on the *y* axis, and the *x* axis displays the percentages for top hits during recruitment. (A) PRT *Nitrosopumilus* SAG and related genomes: *N. maritimus*, *N. maritimus* SCM1; *C. symbiosum*, *Cenarchaeum symbiosum*; *N. salaria* BD31, "*Candidatus* Nitrosopumilus salaria" BD31; *N.* sp. AR2, "*Candidatus* Nitrosopumilus sp." strain AR2. (B) PRT SAR11 SAG and related genomes: HTCC1062, "*Candidatus* Pelagibacter ubique" SAR11 HTCC1062; IMCC9063, "*Candidatus* Pelagibacter sp." IMCC9063; HIMB083, "*Candidatus* Pelagibacter-like" (SAR11) HIMB083. RBB were normalized based on the genome size of the analyzed genome and the total number of reads of the analyzed metagenome.

TABLE 3 Genes unique to the Puerto Rico Trench metagenome[a]

Product name	COGs category, description	COGs ID	BLAST result protein[b]
PRT *Marinosulfonomonas* SAG			
Arsenite oxidase, small subunit	C, energy production and conversion	COG0723	Arsenite oxidase, small-subunit-domain-containing protein (*Pseudovibrio* sp. FO-BEG1)
Nitrous oxide reductase apoprotein	C, energy production and conversion	COG4263	Nitrous oxide reductase (*Ruegeria lacuscaerulensis*)
Intracellular septation protein A	D, cell cycle control, cell division, chromosome partitioning	COG2917	Multidrug transporter (*Roseobacter* sp. SK209-2-6)
Tryptophan 2,3-dioxygenase (vermilion)	E, amino acid transport and metabolism	COG3483	Tryptophan 2,3-dioxygenase (*Leisingera methylohalidivorans* DSM 14336)
4-Aminobutyrate aminotransferase and related aminotransferases	E, amino acid transport and metabolism	COG0160	4-Aminobutyrate aminotransferase (*Rhizobium leguminosarum*)
ABC-type sugar transport systems, ATPase components	G, carbohydrate transport and metabolism	COG3839	ABC transporter ATP-binding protein (*Roseovarius* sp. strain TM1035)
Hypothetical protein	K, transcription	COG2002	Transcriptional regulator ("*Candidatus* Nitrosopumilus salaria")
Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	K, transcription	COG2197	Transcriptional regulator (*Octadecabacter arcticus* 238)
3-Methyladenine DNA glycosylase	L, replication, recombination, and repair	COG2818	DNA-3-methyladenine glycosylase 1 (*Roseobacter litoralis* Och 149)
Xanthine and CO dehydrogenases maturation factor, XdhC/CoxF family	O, posttranslational modification, protein turnover, chaperones	COG1975	Hypothetical protein thalar_02047 (*Thalassobacter arenae* DSM 19593)
Peroxiredoxin	O, posttranslational modification, protein turnover, chaperones	COG0678	Peroxiredoxin (*R. litoralis* Och 149)
Hypothetical protein	R, general function prediction only	COG1058	Molybdenum cofactor biosynthesis protein (*Roseobacter* sp. strain CCS2)
Hypothetical protein			Membrane protein (*Nitrosopumilus* sp. strain SJ)
Polysaccharide lyase family 8, *N*-terminal alpha-helical domain			Silent information regulator protein Sir2 (*Elizabethkingia meningoseptica*)
Hypothetical protein			Sulfatase (*Jannaschia* sp. strain CCS1)
Hemerythrin HHE cation binding domain			LysR family HTH-type transcriptional regulator (*Phaeobacter gallaeciensis* 2.10)
Hypothetical protein			MORN repeat protein (endosymbiont of *Ridgeia piscesae*)
PRT *Psychromonas* SAG			
Ferredoxin	C, energy production and conversion	COG0633	Ferredoxin (*Psychromonas* sp. CNPT3)
Phosphotransferase system fructose-specific component IIB	G, carbohydrate transport and metabolism	COG1445	PTS, fructose-specific IIBC component (*Psychromonas* sp. CNPT3)
Hypothetical protein	N, cell motility	COG1344	Flagellin (*Psychromonas* sp. CNPT3)
PRT *Nitrosopumilus* SAG			
Hypothetical protein	C, energy production and conversion	COG3794	Blue (type 1) copper domain-containing protein ("*Candidatus* Nitrosopumilus sp." AR2)
Histones H3 and H4	B, chromatin structure and dynamics	COG2036	Transcription factor CBF/NF-Y/histone domain-containing protein (*N. maritimus*)
Hypothetical protein	K, transcription	COG2002	AbrB family transcriptional regulator (*N. maritimus* SCM1)
Hypothetical protein	K, transcription		Transcriptional regulator ("*Candidatus* Nitrosopumilus salaria")

[a] SAG sequence reads that recruited to the Puerto Rico Trench metagenome but did not recruit to the other metagenome samples (GOS-NHRI/PCPA/FIGI/SARA [32, 33], HOT 4000 m [17], and DeepMed 3000 m [18] [see Table S1 in the supplemental material]) were identified using FR-HIT (29). The PRT SAR11 SAG did not have any genes that uniquely recruited to the PRT metagenome. HTH, helix-turn-helix; HHE, histidine-histidine-glutamate.

[b] MORN, membrane occupation and recognition nexus; PTS, phosphotransferase system; CBF, CCAAT-binding factor; NF-Y, nuclear transcription factor Y; IIBC, enzyme II B and C combined domains.

organisms accumulate large amount of osmolytes (72) (sometimes referred to as piezolytes [73]), aquaporins might play a role in high-pressure adaptation.

BRBs. Phylogenetic analyses demonstrated that the PRT SAGs were most closely related to the genomes of other deep-ocean microorganisms (Fig. 1). BRB analysis also indicated that the PRT *Nitrosopumilus* and PRT SAR11 SAGs recruited preferentially to deep-ocean metagenomes over surface metagenomes (Fig. 3). This was not apparent for the other two SAGs, which may be due to the dramatically lower abundance of all BRB hits for all metagenomes examined when these SAGs and their comparison genomes were examined, potentially a reflection of their amphipod-associated habitat (see Fig. S3 in the supplemental material).

When the SAG genes were compared to those present in 96 publicly available metagenomes (see Table S2 in the supplemental material), as expected, many were found to be present in all the metagenomes, but no SAG genes were unique among deep-ocean metagenomes (PRT, HOT 4000 m, and DeepMed 3000 m [deep Mediterranean]). However, read recruitments did identify genes uniquely recruiting to the PRT metagenome (Table 3). Among the genes uniquely represented, 30% were related to genes involved in transcriptional regulation, and 15% were related to genes associated with transporters. Of these, no one gene or set of genes was found across all PRT SAGs. The lack of a conserved deep-sea gene might stem from the undersampling of ultradeep-ocean environments.

Conclusions. The use of single-cell genomics enabled the characterization of four partial genomes, all belonging to deep-sea clades. The two SAGs recovered from undiluted seawater are of special interest (PRT SAR11 SAG and PRT *Nitrosopumilus* SAG), as they are from the deepest-dwelling studied members of highly abundant groups with large-scale biogeochemical significance (74–77). The other two SAGs provide additional information for groups of previously cultured piezophiles.

At the functional level, genes present in the SAGs but absent in their comparison MCRGs revealed novel metabolic capabilities, including those associated with nitrogen, sulfur, carbon, and energy acquisition mechanisms. Some of the most significant findings are the potential for PRT *Thaumarchaeota* to synthesize fatty acids and utilize alternate ways to recover ammonia, CO_2, and methyl groups. Also significant is the discovery that within the ubiquitous SAR11 group, PRT SAR11 encodes a complete glycolytic pathway, and it contains the genes for phosphofructokinase and pyruvate kinase. From the samples recovered from the amphipod-associated cells, the PRT *Psychromonas* and PRT *Marinosulfonomonas* SAGs, it seems that genes involved in generating energy from H_2 or N_2O oxidation are potentially significant when considering survival in the PRT. The importance of osmoregulation, possibly related to high-pressure adaptation in the ultradeep ocean, is suggested by the finding of aquaporins in PRT *Marinosulfonomonas* and PRT *Nitrosopumilus* SAGs.

When we looked at the presence and abundance of SAG genes in different metagenomes, specific genes were uniquely found to be associated with the PRT. Many of these encoded functions associated with transport or transcription. Also, PRT SAR11 and PRT *Nitrosopumilus* SAG genes were more abundant in the metagenomes examined from deeper-dwelling organisms. Overall, the single cells collected from the PRT suggest that the development of added metabolic capabilities might be advantageous for survival in this and perhaps other hadal locations.

ACKNOWLEDGMENTS

We thank Jessica Blanton, Juan Ugalde, Greg Rouse, Jose Carvajal, Carlos Rios-Velazquez, Logan Peoples, Graham Wilhelm, Shelbi Randenberg, Jessica Wdowiarz, and Justin DeShields for their contribution to the discussion and collection of samples and data.

We are also grateful for the financial support provided by the National Science Foundation (grants 0801973 and 0827051 and graduate research fellowship 068775), the National Aeronautics and Space Administration (grant NNX11AG10G), the National Institute of General Medical Sciences of the National Institutes of Health (award T32GM067550), the UCSD Academic Senate, and Scripps Institution of Oceanography, as well as for a generous contribution from Joanie Nasher.

We declare no conflicts of interest.

REFERENCES

1. **Jamieson AJ.** 2001. Ecology of deep oceans: hadal trenches. *In* Encyclopedia of life sciences. John Wiley & Sons, Chichester, England. http://dx.doi.org/10.1002/9780470015902.a0023606.
2. **Nunoura T, Takaki Y, Hirai M, Shimamura S, Makabe A, Koide O, Kikushi T, Miyazaki J, Koba K, Yoshida N, Sunamura M, Takai K.** 2015. Hadal biosphere: insight into the microbial ecosystem in the deepest ocean on Earth. Proc Natl Acad Sci U S A **112:**E1230–E1236.
3. **Blankenship L, Yayanos A, Cadien D, Levin L.** 2006. Vertical zonation patterns of scavenging amphipods from the hadal zone of the Tonga and Kermadec Trenches. Deep-Sea Res **53:**48–61. http://dx.doi.org/10.1016/j.dsr.2005.09.006.
4. **France S.** 1993. Geographic-variation among 3 isolated populations of the hadal amphipod *Hirondellea gigas* (*Crustacea*, *Amphipoda*, *Lysianassoidea*). Mar Ecol Prog Ser **92:**277–287. http://dx.doi.org/10.3354/meps092277.
5. **Jamieson AJ, Kilgallen NM, Rowden AA, Fujii T, Horton T, Lörz AN, Kitazawa K, Priede IG.** 2011. Bait-attending fauna of the Kermadec Trench, SW Pacific Ocean: evidence for an ecotone across the abyssal-hadal transition zone. Deep-Sea Res **58:**49–62. http://dx.doi.org/10.1016/j.dsr.2010.11.003.
6. **Bartlett DH.** 2002. Pressure effects on in vivo microbial processes. Biochim Biophys Acta **1595:**367–381. http://dx.doi.org/10.1016/S0167-4838(01)00357-0.
7. **Vezzi A, Campanaro S, D'Angelo M, Simonato F, Vitulo N, Lauro F, Cestaro A, Malacrida G, Simionati B, Cannata N, Romualdi C, Bartlett DH, Valle G.** 2005. Life at depth: *Photobacterium profundum* genome sequence and expression analysis. Science **307:**1459–1461. http://dx.doi.org/10.1126/science.1103341.
8. **Lauro FM, Bartlett DH.** 2008. Prokaryotic lifestyles in deep sea habitats. Extremophiles **12:**15–25. http://dx.doi.org/10.1007/s00792-006-0059-5.
9. **Campanaro S, De Pascale F, Telatin A, Schiavon R, Bartlett D, Valle G.** 2012. The transcriptional landscape of the deep-sea bacterium *Photobacterium profundum* in both a *toxR* mutant and its parental strain. BMC Genomics **13:**567. http://dx.doi.org/10.1186/1471-2164-13-567.
10. **Tamegai H, Nishikawa S, Haga M, Bartlett DH.** 2012. The respiratory system of the piezophile *Photobacterium profundum* SS9 grown under various pressures. Biosci Biotechnol Biochem **76:**1506–1510.
11. **Le Bihan T, Rayner J, Roy MM, Spagnolo L.** 2013. *Photobacterium profundum* under pressure: a MS-based label-free quantitative proteomics study. PLoS One **8:**e60897. http://dx.doi.org/10.1371/journal.pone.0060897.
12. **Ohke Y, Sakoda A, Kato C, Sambongi Y, Kawamoto J, Kurihara T, Tamegai H.** 2013. Regulation of cytochrome *c*-and quinol oxidases, and piezotolerance of their activities in the deep-sea piezophile *Shewanella violacea* DSS12 in response to growth conditions. Biosci Biotechnol Biochem **77:**1522–1528.
13. **Allen EE, Facciotti D, Bartlett DH.** 1999. Monounsaturated but not polyunsaturated fatty acids are required for growth of the deep-sea bacterium *Photobacterium profundum* SS9 at high pressure and low temperature. Appl Environ Microbiol **65:**1710–1720.
14. **Allen EE, Bartlett DH.** 2000. FabF is required for piezoregulation of cis-vaccenic acid levels and piezophilic growth of the deep-sea bacterium *Photobacterium profundum* strain SS9. J Bacteriol **182:**1264–1271.
15. **Kawamoto J, Sato T, Nakasone K, Kato C, Mihara H, Esaki N, Kurihara T.** 2011. Favourable effects of eicosapentaenoic acid on the late step of the cell division in a piezophilic bacterium, *Shewanella violacea* DSS12, at high-hydrostatic pressures. Environ Microbiol **13:**2293–2298.

16. **Eloe E, Fadrosh D, Novotny M, Zeigler Allen L, Kim M, Lombardo M, Yee-Greenbaum J, Yooseph S, Allen EE, Lasken R, Williamson SJ, Bartlett DH.** 2011. Going deeper: metagenome of a hadopelagic microbial community. PLoS One **6:**e20388. http://dx.doi.org/10.1371/journal.pone.0020388.
17. **DeLong E, Preston C, Mincer T, Rich V, Hallam S, Frigaard N, Martinez A, Sullivan MB, Edwards R, Rodriguez-Brito B, Chisholm SW, Karl DM.** 2006. Community genomics among stratified microbial assemblages in the ocean's interior. Science **311:**496–503. http://dx.doi.org/10.1126/science.1120250.
18. **Martín-Cuadrado A-B, Lopez-Garcia P, Alba J, Moreira D, Monticelli L, Strittmatter A, Gottschalk G, Rodriguez-Valera F.** 2007. Metagenomics of the Deep Mediterranean, a warm bathypelagic habitat. PLoS One **2:**e0000914. http://dx.doi.org/10.1371/journal.pone.0000914.
19. **Konstantinidis KT, Braff J, Karl DM, DeLong EF.** 2009. Comparative metagenomic analysis of a microbial community residing at a depth of 4,000 meters at station ALOHA in the North Pacific subtropical gyre. Appl Environ Microbiol **75:**5345–5355. http://dx.doi.org/10.1128/AEM.00473-09.
20. **Smedile F, Messina E, La Cono V, Tsoy O, Monticelli L, Borghini M, Giuliano L, Golyshin PN, Mushegian A, Yakimov MM.** 2013. Metagenomic analysis of hadopelagic microbial assemblages thriving at the deepest part of Mediterranean Sea, Matapan-Vavilov Deep. Environ Microbiol **15:**167–182. http://dx.doi.org/10.1111/j.1462-2920.2012.02827.x.
21. **Lasken RS.** 2012. Genomic sequencing of uncultured microorganisms from single cells. Nat Rev Microbiol **10:**631–640.
22. **Lasken RS, McLean JS.** 2014. Recent advances in genomic DNA sequencing of microbial species from single cells. Nat Rev Genet **15:**577–584.
23. **McLean JS, Lombardo M-J, Badger JH, Edlund A, Novotny M, Yee-Greenbaum J, Vyahhi N, Hall AP, Yang Y, Dupont CL, Ziegler MG, Chitsaz H, Allen AE, Yoosheph S, Tesler G, Pevzner PA, Friedman RM, Nealson KH, Venter JC, Lasken RS.** 2013. Candidate phylum TM6 genome recovered from a hospital sink biofilm provides genomic insights into this uncultivated phylum. Proc Natl Acad Sci U S A **110:**E2390–E2399. http://dx.doi.org/10.1073/pnas.1219809110.
24. **Altschul S, Gish W, Miller W, Myers E, Lipman D.** 1990. Basic Local Alignment Search Tool. J Mol Biol **215:**403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.
25. **Markowitz V, Chen I, Palaniappan K, Chu K, Szeto E, Pillay M, Ratner A, Huang J, Woyke T, Huntemann M, Anderson I, Billis K, Varghese N, Mavromatis K, Pati A, Ivanova NN, Kyrpides NC.** 2014. IMG 4 version of the integrated microbial genomes comparative analysis system. Nucleic Acids Res **42:**D560–D567. http://dx.doi.org/10.1093/nar/gkt963.
26. **Pruesse E, Peplies J, Glöckner F.** 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics **28:**1823–1829. http://dx.doi.org/10.1093/bioinformatics/bts252.
27. **Price M, Dehal P, Arkin A.** 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol **26:**1641–1650. http://dx.doi.org/10.1093/molbev/msp077.
28. **Podell S, Gaasterland T.** 2007. DarkHorse: a method for genome-wide prediction of horizontal gene transfer. Genome Biol **8:**R16.
29. **Jones AC, Monroe EA, Podell S, Hess WR, Klages S, Esquenazi E, Niessen S, Hoover H, Rothmann M, Lasken RS, Yates JR, III, Reinhard R, Kube M, Burkart MD, Allen EE, Dorrestein PC, Gerwick WH, Gerwick L.** 2011. Genomic insights into the physiology and ecology of the marine filamentous cyanobacterium *Lyngbya majuscula*. Proc Natl Acad Sci U S A **108:**8815–8820.
30. **Podell S, Ugalde J, Narasingarao P, Banfield J, Heidelberg K, Allen E.** 2013. Assembly-driven community genomics of a hypersaline microbial ecosystem. PLoS One **8:**e61692. http://dx.doi.org/10.1371/journal.pone.0061692.
31. **Niu B, Zhu Z, Fu L, Wu S, Li W.** 2011. FR-HIT, a very fast program to recruit metagenomic reads to homologous reference genomes. Bioinformatics **27:**1704–1705. http://dx.doi.org/10.1093/bioinformatics/btr252.
32. **Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers YH, Smith HO.** 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science **304:**66–74. http://dx.doi.org/10.1126/science.1093857.
33. **Rusch D, Halpern A, Sutton G, Heidelberg K, Williamson S, Yooseph S, Wu D, Eisen JA, Hoffman JM, Remington K, Beeson K, Tran B, Smith H, Baden-Tillson H, Stewart C, Thorpe J, Freeman J, Andrews-Pfannkoch C, Venter JE, Li K, Kravitz S, Heidelberg JF, Utterback T, Rogers YH, Falcón LI, Souza V, Bonilla-Rosso G, Eguiarte LE, Karl DM, Sathyendranath S, Platt T, Bermingham E, Gallardo V, Tamayo-Castillo G, Ferrari MR, Strausberg RL, Nealson K, Friedman R, Frazier M, Venter JC.** 2007. The Sorcerer II Global Ocean sampling expedition: Northwest Atlantic through Eastern Tropical Pacific. PLoS Biol **5:**e77.
34. **Yayanos AA, Dietz AS, Van Boxtel R.** 1981. Obligately barophilic bacterium from the Mariana Trench. Proc Natl Acad Sci U S A **78:**5212–5215.
35. **Lauro FM, Chastain RA, Blankenship LE, Yayanos AA, Bartlett DH.** 2007. The unique 16S rRNA genes of piezophiles reflect both phylogeny and adaptation. Appl Environ Microbiol **73:**838–845.
36. **Könneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA.** 2005. Isolation of an autotrophic ammonia-oxidizing marine archaeon. Nature **437:**543–546.
37. **Pester M, Schleper C, Wagner M.** 2011. The *Thaumarchaeota*: an emerging view of their phylogeny and ecophysiology. Curr Opin Microbiol **14:**300–306. http://dx.doi.org/10.1016/j.mib.2011.04.007.
38. **Durbin AM, Teske A.** 2010. Sediment-associated microdiversity within the marine group I *Crenarchaeota*. Environ Microbiol Rep **2:**693–703. http://dx.doi.org/10.1111/j.1758-2229.2010.00163.x.
39. **Thrash JC, Temperton B, Swan BK, Landry ZC, Woyke T, DeLong EF. Stepanauskas R, Giovannoni ST.** 2014. Single-cell enabled comparative genomics of a deep ocean SAR11 bathytype. ISME J **8:**1440–1451.
40. **Vergin KL, Beszteri B, Monier A, Thrash JC, Temperton B, Treusch AH, Kilpert F, Worden AZ, Giovannoni SJ.** 2013. High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic time-series study site by phylogenetic placement of pyrosequences. ISME J **7:**1322–1332. http://dx.doi.org/10.1038/ismej.2013.32.
41. **Shaw AK, Halpern AL, Beeson K, Tran B, Venter JC, Martiny JB.** 2008. It's all relative: ranking the diversity of aquatic bacterial communities. Environ Microbiol **10:**2200–2210. http://dx.doi.org/10.1111/j.1462-2920.2008.01626.x.
42. **Hugler M, Gartner A, Imhoff J.** 2010. Functional genes as markers for sulfur cycling and CO_2 fixation in microbial communities of hydrothermal vents of the Logatchev field. FEMS Microbiol Ecol **73:**526–537.
43. **Eloe E, Malfatti F, Gutierrez J, Hardy K, Schmidt W, Pogliano K, Pogliano J, Azam F, Bartlett DH.** 2011. Isolation and characterization of a psychropiezophilic alphaproteobacterium. Appl Environ Microbiol **77:**8145–8153.
44. **Goffredi S, Orphan V, Rouse G, Jahnke L, Embaye T, Turk K, Lee R, Vrijenhoek RC.** 2005. Evolutionary innovation: a bone-eating marine symbiosis. Environ Microbiol **7:**1369–1378. http://dx.doi.org/10.1111/j.1462-2920.2005.00824.x.
45. **DeLong E, Franks D, Yayanos A.** 1997. Evolutionary relationships of cultivated psychrophilic and barophilic deep-sea bacteria. Appl Environ Microbiol **63:**2105–2108.
46. **Lauro FM, Stratton TK, Chastain RA, Ferriera S, Johnson J, Goldberg SM, Yayanos AA, Bartlett DH.** 2013. Complete genome sequence of the deep-sea bacterium *Psychromonas* strain CNPT3. Genome Announc **1**(3):e00304-13. http://dx.doi.org/10.1128/genomeA.00304-13.
47. **Kryazhimskiy S, Plotkin JB.** 2008. The population genetics of *dN/dS*. PLoS Genet **4:**e1000304. http://dx.doi.org/10.1371/journal.pgen.1000304.
48. **Swan BK, Tupper B, Sczyrba A, Lauro FM, Martinez-Garcia M, González JM, Luo H, Wright JJ, Landry ZC, Hanson NW, Thompson BP, Poulton NJ, Schwientek P, Acinas SG, Giovannoni SJ, Moran MA, Hallam SJ, Cavicchioli R, Woyke T, Stepanauskas P.** 2013. Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. Proc Natl Acad Sci U S A **110:**11463–11468. http://dx.doi.org/10.1073/pnas.1304246110.
49. **Richardson MD, Briggs KB, Bowles FA, Tietjen JH.** 1995. A depauperate benthic assemblage from the nutrient-poor sediments of the Puerto Rico Trench. Deep-Sea Res **42:**351–364.
50. **Rasmussen JJ, Vegge CS, Frøkiaer H, Howlett RM, Krogfelt KA, Kelly DJ, Ingmer KH.** 2013. *Campylobacter jejuni* carbon starvation protein A (CstA) is involved in peptide utilization, motility and agglutination, and has a role in stimulation of dendritic cells. J Med Microbiol **62:**1135–1143. http://dx.doi.org/10.1099/jmm.0.059345-0.
51. **Herndl GJ, Reinthaler T, Teira E, van Aken H, Veth C, Pernthaler A, Pernthaler J.** 2005. Contribution of *Archaea* to total prokaryotic production in the Deep Atlantic Ocean. Appl Environ Microbiol **71:**2303–2309. http://dx.doi.org/10.1128/AEM.71.5.2303-2309.2005.
52. **Posner MG, Upadhyay A, Crennell SJ, Watson AJA, Dorus S, Danson MJ, Bagby S.** 2013. Post-translational modification in the archaea: struc-

tural characterization of multi-enzyme complex lipoylation. Biochem J **449:**415–425. http://dx.doi.org/10.1042/BJ20121150.
53. **Borziak K, Posner M, Upadhyay A, Danson M, Bagby S, Dorus S.** 2014. Comparative genomic analysis reveals 2-oxoacid dehydrogenase complex lipoylation correlation with aerobiosis in archaea. PLoS One **9:**e87063. http://dx.doi.org/10.1371/journal.pone.0087063.
54. **Lokanath N, Kuroishi C, Okazaki N, Kunishima N.** 2004. Purification, crystallization and preliminary crystallographic analysis of the glycine-cleavage system component T-protein from *Pyrococcus horikoshii* OT3. Acta Crystallogr D Biol Crystallogr **60:**1450–1452. http://dx.doi.org/10.1107/S0907444904012910.
55. **Fischer S, Maier LK, Stoll B, Brendel J, Fischer E, Pfeiffer F, Dyall-Smith M, Marchfelder A.** 2012. An archaeal immune system can detect multiple protospacer adjacent motifs (PAMs) to target invader DNA. J Biol Chem **287:**33351–33363. http://dx.doi.org/10.1074/jbc.M112.377002.
56. **Lu L, Han W, Zhang J, Wu Y, Wang B, Lin X, Zhu J, Cai Z, Jia Z.** 2012. Nitrification of archaeal ammonia oxidizers in acid soils is supported by hydrolysis of urea. ISME J **6:**1978–1984. http://dx.doi.org/10.1038/ismej.2012.45.
57. **Leigh J, Dodsworth J.** 2007. Nitrogen regulation in bacteria and archaea. Annu Rev Microbiol **61:**349–377. http://dx.doi.org/10.1146/annurev.micro.61.080706.093409.
58. **Lombard J, López-García P, Moreira D.** 2012. An ACP-independent fatty acid synthesis pathway in archaea: implications for the origin of phospholipids. Mol Biol Evol **29:**3261–3265. http://dx.doi.org/10.1093/molbev/mss160.
59. **Sanford R, Wagner D, Wu Q, Chee-Sanford J, Thomas S, Cruz-Garcia C, Rodríguez G, Massol-Deyá A, Krishnan KK, Ritalahti KM, Nissen S, Konstantinidis KT, Löffler FE.** 2012. Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils. Proc Natl Acad Sci U S A **109:**19709–19714. http://dx.doi.org/10.1073/pnas.1211238109.
60. **Wilson ST, del Valle DA, Segura-Noguera M, Karl DM.** 2014. A role for nitrite in the production of nitrous oxide in the lower euphotic zone of the oligotrophic North Pacific Ocean. Deep-Sea Res **85:**47–55. http://dx.doi.org/10.1016/j.dsr.2013.11.008.
61. **Schwalbach MS, Tripp HJ, Steindler L, Smith DP, Giovannoni SJ.** 2010. The presence of the glycolysis operon in SAR11 genomes is positively correlated with ocean productivity. Environ Microbiol **12:**490–500.
62. **Grote J, Thrash JC, Huggett MJ, Landry ZC, Carini P, Giovannoni SJ, Rappé MS.** 2012. Streamlining and core genome conservation among highly divergent members of the SAR11 clade. mBio **3**(5):e00252-12. http://dx.doi.org/10.1128/mBio.00252-12.
63. **Viklund J, Martijn J, Ettema TJG, Andersson SGE.** 2013. Comparative and phylogenomic evidence that the alphaproteobacterium HIMB59 is not a member of the oceanic SAR11 clade. PLoS One **8:**e78858. http://dx.doi.org/10.1371/journal.pone.0078858.
64. **Richardson DJ.** 2000. Bacterial respiration: a flexible process for a changing environment. Microbiology **146:**551–571. http://dx.doi.org/10.1099/00221287-146-3-551.
65. **Vignais P, Billoud B.** 2007. Occurrence, classification, and biological function of hydrogenases: an overview. Chem Rev **107:**4206–4272. http://dx.doi.org/10.1021/cr050196r.
66. **Kim JY, Jo BH, Cha HJ.** 2011. Production of biohydrogen by heterologous expression of oxygen-tolerant *Hydrogenovibrio marinus* [NiFe]-hydrogenase in *Escherichia coli*. J Biotechnol **155:**312–319. http://dx.doi.org/10.1016/j.jbiotec.2011.07.007.
67. **Tripp HJ, Kitner JB, Schwalbach MS, Dacey JWH, Wilhelm LJ, Giovannoni SJ.** 2008. SAR11 marine bacteria require exogenous reduced sulphur for growth. Nature **452:**741–744. http://dx.doi.org/10.1038/nature06776.
68. **Eichhorn E, van der Ploeg JR, Kertesz MA, Leisinger T.** 1997. Characterization of α-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. J Biol Chem **272:**23031–23036. http://dx.doi.org/10.1074/jbc.272.37.23031.
69. **Kurz M, Brünig ANS, Galinski EA.** 2006. NhaD type sodium/proton-antiporter of *Halomonas elongata*: a salt stress response mechanism in marine habitats? Saline Syst **2:**10. http://dx.doi.org/10.1186/1746-1448-2-10.
70. **Kumar M, Grzelakowski M, Zilles J, Clark M, Meier W.** 2007. Highly permeable polymeric membranes based on the incorporation of the functional water channel protein aquaporin Z. Proc Natl Acad Sci U S A **104:**20719–20724.
71. **Robinson CR, Sligar SG.** 1995. Hydrostatic and osmotic pressure as tools to study macromolecular recognition. Methods Enzymol **259:**395–427.
72. **Yancey PH, Gerringer ME, Drazen JC, Rowden AA, Jamieson A.** 2014. Marine fish may be biochemically constrained from inhabiting the deepest ocean depths. Proc Natl Acad Sci U S A **111:**4461–4465. http://dx.doi.org/10.1073/pnas.1322003111.
73. **Martin DD, Bartlett DH, Roberts MF.** 2002. Solute accumulation in the deep-sea bacterium *Photobacterium profundum*. Extremophiles **6:**507–514. http://dx.doi.org/10.1007/s00792-002-0288-1.
74. **Morris R, Rappé M, Connon S, Vergin K, Siebold W, Carlson C, Giovannoni SJ.** 2002. SAR11 clade dominates ocean surface bacterioplankton communities. Nature **420:**806–810. http://dx.doi.org/10.1038/nature01240.
75. **Schattenhofer M, Fuchs BM, Amann R, Zubkov MV, Tarran GA, Pernthaler J.** 2009. Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. Environ Microbiol **11:**2078–2093. http://dx.doi.org/10.1111/j.1462-2920.2009.01929.x.
76. **Karner MB, DeLong EF, Karl DM.** 2001. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. Nature **409:**507–510. http://dx.doi.org/10.1038/35054051.
77. **Teira E, van Aken H, Veth C, Herndl GJ.** 2006. Archaeal uptake of enantiomeric amino acids in the meso- and bathypelagic waters of the North Atlantic. Limnol Oceanogr **51:**60–69. http://dx.doi.org/10.4319/lo.2006.51.1.0060.
78. **Bankevich A, Nurk S, Antipov D, Gurevich A, Dvorkin M, Kulikov A, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA.** 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol **19:**455–477. http://dx.doi.org/10.1089/cmb.2012.0021.