



Atribacteria Reproducing over Millions of Years in the Atlantic Abyssal Subseafloor

Aurèle Vuillemin,^a Sergio Vargas,^a Ömer K. Coskun,^a Robert Pockalny,^c Richard W. Murray,^d David C. Smith,^c Steven D'Hondt,^c
 William D. Orsi^{a,b}

^aDepartment of Earth and Environmental Sciences, Paleontology and Geobiology, Ludwig-Maximilians-Universität München, Munich, Germany

^bGeoBio-Center^{LMU}, Ludwig-Maximilians-Universität München, Munich, Germany

^cGraduate School of Oceanography, University of Rhode Island, Narragansett, Rhode Island, USA

^dWoods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA

ABSTRACT How microbial metabolism is translated into cellular reproduction under energy-limited settings below the seafloor over long timescales is poorly understood. Here, we show that microbial abundance increases an order of magnitude over a 5 million-year-long sequence in anoxic subseafloor clay of the abyssal North Atlantic Ocean. This increase in biomass correlated with an increased number of transcribed protein-encoding genes that included those involved in cytokinesis, demonstrating that active microbial reproduction outpaces cell death in these ancient sediments. Metagenomes, metatranscriptomes, and 16S rRNA gene sequencing all show that the actively reproducing community was dominated by the candidate phylum “*Candidatus* Atribacteria,” which exhibited patterns of gene expression consistent with fermentative, and potentially acetogenic, metabolism. “*Ca.* Atribacteria” dominated throughout the 8 million-year-old cored sequence, despite the detection limit for gene expression being reached in 5 million-year-old sediments. The subseafloor reproducing “*Ca.* Atribacteria” also expressed genes encoding a bacterial microcompartment that has potential to assist in secondary fermentation by recycling aldehydes and, thereby, harness additional power to reduce ferredoxin and NAD⁺. Expression of genes encoding the Rnf complex for generation of chemiosmotic ATP synthesis were also detected from the subseafloor “*Ca.* Atribacteria,” as well as the Wood-Ljungdahl pathway that could potentially have an anabolic or catabolic function. The correlation of this metabolism with cytokinesis gene expression and a net increase in biomass over the million-year-old sampled interval indicates that the “*Ca.* Atribacteria” can perform the necessary catabolic and anabolic functions necessary for cellular reproduction, even under energy limitation in millions-of-years-old anoxic sediments.

IMPORTANCE The deep subseafloor sedimentary biosphere is one of the largest ecosystems on Earth, where microbes subsist under energy-limited conditions over long timescales. It remains poorly understood how mechanisms of microbial metabolism promote increased fitness in these settings. We discovered that the candidate bacterial phylum “*Candidatus* Atribacteria” dominated a deep-sea subseafloor ecosystem, where it exhibited increased transcription of genes associated with acetogenic fermentation and reproduction in million-year-old sediment. We attribute its improved fitness after burial in the seabed to its capabilities to derive energy from increasingly oxidized metabolites via a bacterial microcompartment and utilize a potentially reversible Wood-Ljungdahl pathway to help meet anabolic and catabolic requirements for growth. Our findings show that “*Ca.* Atribacteria” can perform all the necessary catabolic and anabolic functions necessary for cellular reproduction, even under energy limitation in anoxic sediments that are millions of years old.

Citation Vuillemin A, Vargas S, Coskun ÖK, Pockalny R, Murray RW, Smith DC, D'Hondt S, Orsi WD. 2020. Atribacteria reproducing over millions of years in the Atlantic abyssal subseafloor. *mBio* 11:e01937-20. <https://doi.org/10.1128/mBio.01937-20>.

Editor Nicole Dubilier, Max Planck Institute for Marine Microbiology

Copyright © 2020 Vuillemin et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to William D. Orsi, w.orsi@lrz.uni-muenchen.de.

Received 15 July 2020

Accepted 8 September 2020

Published 6 October 2020

KEYWORDS deep biosphere, energy limit to life, aribacteria, acetogenesis, metagenomics, transcriptomics, fermentation, bacterial microcompartment, clade JS1, metatranscriptomics, subseafloor life

Marine sediments contain a ubiquitous “deep biosphere” (1) extending at least as far as 2,500 m below the seafloor (mbsf) (2), which consists of active and dormant cells (3–5) with measurable impacts on subseafloor biogeochemical processes (6). At abyssal water depths in the deep sea, the subseafloor communities generally are less sampled (7) than those in continental shelf sediments that have higher activities and rates of microbial sulfate reduction (8–10). At abyssal depths under the oligotrophic ocean gyres, the sedimentation rates are low, ranging from 1 to 5 m of sediment deposited per million years (11, 12). These abyssal subseafloor communities have extremely low metabolic activity (13) and live near the energy limit to life (6, 14). As a result of this, the deep biosphere of marine sediment is generally characterized by net death, and it remains poorly understood to what extent microbial activity is translated into cellular reproduction in the deep subseafloor (15, 16).

Many subseafloor microbes exhibit viability, since they actively take up carbon and nitrogen in incubation experiments (4, 17), indicating potential for microbial growth in energy-limited anoxic subseafloor sediments. Microbial activities can also be stimulated at redox interfaces deep below the seafloor over geological timescales (3, 18). However, the capacity of microbes to reproduce in abyssal subseafloor ecosystems close to the energy limit to life (6, 14) is particularly unconstrained, given the extreme scarcity of organic substrates in these settings (19). There is reason to suspect that cellular reproduction in the abyssal subseafloor is minimal, since microbial biomass tends to decrease an order of magnitude over the top 10 m of sediment in all abyssal locations yet sampled, reaching the detection limit for life at relatively shallow subseafloor depths of ca. 15 mbsf (12). This follows the global trend whereby subseafloor microbes tend to die faster than they grow (1), particularly in the top 10 m of marine sediment.

Here, we report an exception to this global trend in anoxic deep-sea clay recovered from an abyssal water depth of >5,500 m in the North Atlantic, characterized by an ultraslow sedimentation rate of ca. 3 m per million years. In contrast to oxic abyssal red clay where microbial abundance decreases several orders of magnitude over the top 10 mbsf (12, 20), we show here that microbial abundance in the anoxic abyssal clay increases an order of magnitude from the seafloor down to 15 mbsf (spanning ca. 5 million years). We then proceeded to use metatranscriptomics to further investigate the anaerobic metabolic mechanisms that explain this net growth in the size of the subseafloor microbial ecosystem over multimillion-year timescales.

RESULTS AND DISCUSSION

Sediment biogeochemistry. We obtained deep-sea clay sediment from a 5,515-m water depth in the ultraoligotrophic open ocean of the North Atlantic. This coring site (KN223-15) is characterized by a mean sedimentation rate of ca. 3 m per million years (11). Samples ranged from 0.1 to 30 m below seafloor (mbsf). Given the mean sedimentation rate, the deepest sample has an approximate age of 9 to 10 million years. Oxygen and nitrate penetration into the sediment is restricted to the top millimeter of sediment, as they were detectable in the bottom water but below detection in the uppermost portion of the core at 0.02 and 0.03 mbsf, respectively (see Fig. S1 in the supplemental material). The abyssal sediments of the North Atlantic are typically oxic red clay that tend to have O₂ penetrating many meters into the seafloor (11, 21), but the subseafloor microbial ecosystem sampled here is unique in the sense that the sediments are anoxic despite having ultraslow sedimentation rates. Moreover, while the sediment of our abyssal subseafloor core displays sulfate (SO₄²⁻) concentrations of approximately 29 mM at 0.02 mbsf and is fully anoxic downward, the rates of anaerobic microbial SO₄²⁻ reduction over the top 10 m of sampled sediment are low (-3.8×10^{-4} mol SO₄²⁻ m³ yr⁻¹) and below detection underneath, resulting in pore

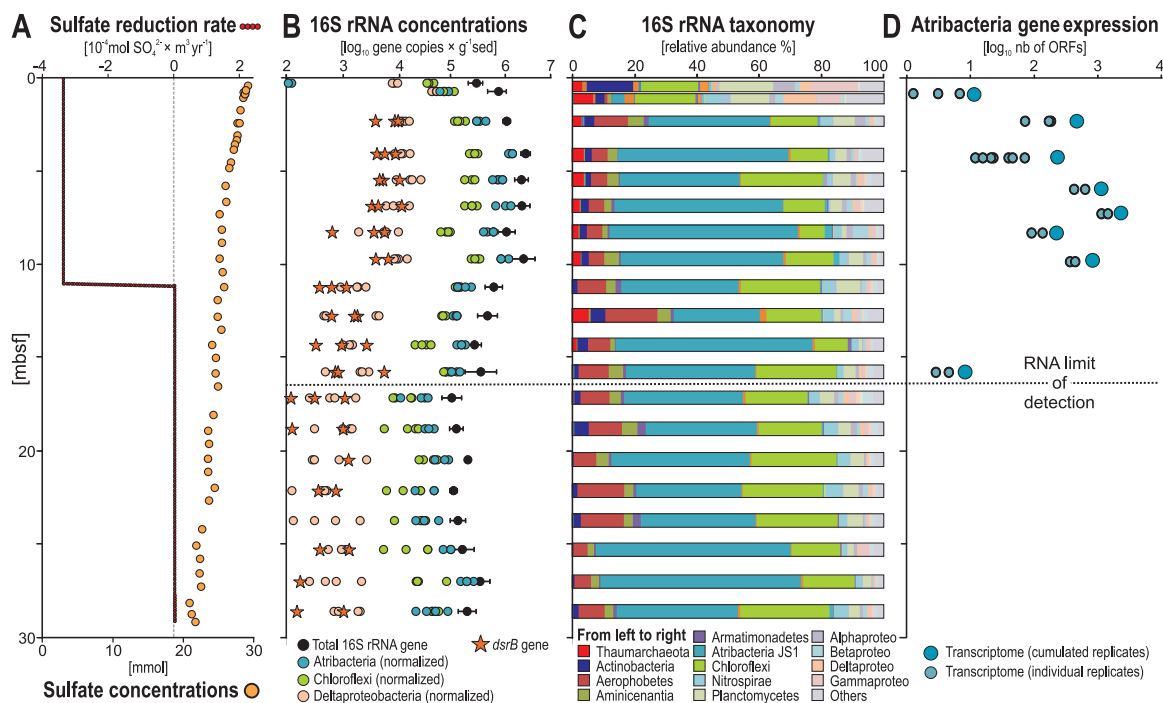


FIG 1 Biogeochemistry, microbial diversity, and abundance in the anoxic sediment of North Atlantic site KN223-15. (A) Profiles of mean net SO₄²⁻ reduction rates and dissolved SO₄²⁻ concentrations. (B) Quantitative PCR (qPCR) of 16S rRNA (black dots) and *dsrB* (orange stars) genes, and the summed qPCR-normalized abundance of 16S rRNA gene OTUs affiliated with “*Ca. Atribacteria*” (blue), *Chloroflexi* (green), and *Deltaproteobacteria* (light orange). Error bars correspond to standard deviations (2 σ) based on four biological replicates. (C) Diversity of 16S rRNA genes based on three to four biological replicates. (D) Number of ORFs attributed to “*Ca. Atribacteria*” in the metatranscriptomes as a function of depth. Light blue circles are the number of ORFs detected in individual metatranscriptome replicates, and the larger dark blue circles are the total number of unique ORFs detected when summing across all replicates. The top two samples in panels B and C (0.1 to 0.3 mbsf) were recovered by gravity coring, the deeper samples (0.5 to 29 mbsf) were recovered via long-piston coring. Note in panels B and C that an increase in abundance of “*Ca. Atribacteria*” in the upper 1 mbsf coincides with their higher level of gene expression in panel D.

water SO₄²⁻ remaining >20 mM throughout the core (Fig. 1A). This profile is very similar to profiles observed previously in anoxic sediments from other oligotrophic regions, such as the Eastern Equatorial Pacific, where sediments are anoxic but community metabolic activity is too slow to consume all of the available SO₄²⁻ (6, 22).

Abundance and diversity of the microbial communities. The density of 16S rRNA gene copies per gram of wet sediment exhibits a subsurface peak in abundance, increasing 1 order of magnitude (2.5×10^5 to 2.3×10^6 copies) from 0.2 to 3.5 mbsf and remaining >10⁶ between 3 and 10 mbsf. Thereafter, microbial 16S rRNA gene abundances decrease gradually to a minimum of 9.9×10^4 copies at 17 mbsf (Fig. 1B). In the uppermost sediment samples (0.1 to 0.4 mbsf), *Actinobacteria*, *Planctomycetes*, *Deltaproteobacteria*, and *Gammaproteobacteria* dominate the community (Fig. 1C). Below this depth, starting at 0.5 mbsf, the relative abundance of “*Candidatus Atribacteria*” rapidly increased from <5% to >40% (Fig. 1B and C) in all four biological replicates sampled (see Fig. S2 and S3). The “*Ca. Atribacteria*” is only represented by 3 operational taxonomic units (OTUs), with one single OTU accounting for up to 40% of the whole community throughout the record (Fig. 2A and B), which was consistent across three to four biological replicates (Fig. S2). This dominant OTU (Fig. 2A) was most closely affiliated with unpublished 16S rRNA gene sequences from deep subsurface sediments from the Nankai Trough (23) and a potential gas hydrate region of southwest Taiwan subsurface sediments (Fig. 2B).

According to a protocol for quantitative normalization of barcoded 16S gene sequence data (24), we normalized group-specific 16S rRNA gene abundances by dividing the total number of 16S rRNA gene copies determined via quantitative PCR (qPCR) by their affiliated 16S rRNA gene relative abundances (fractional percentage of

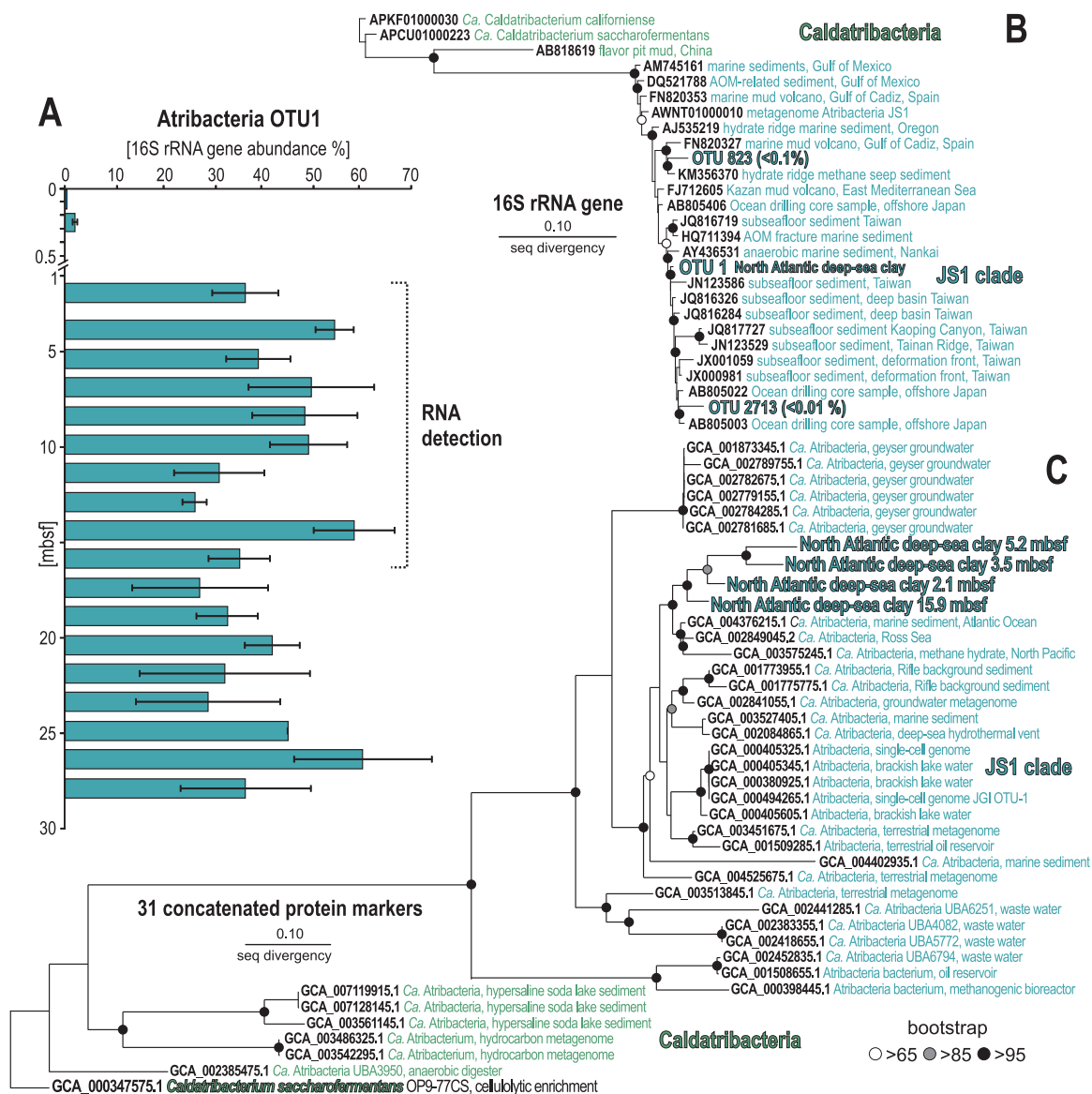


FIG 2 Relative abundance and phylogenetic analyses of the most abundant OTU (“*Ca. Atribacteria*”) and its phylogenomic protein markers. (A) The histogram shows the relative abundance of the most abundant OTU (“*Ca. Atribacteria*”), error bars are standard deviations from 3 to 4 biological replicates (see Fig. S2 in the supplemental material). (B) Phylogenetic analysis of 16S rRNA genes (V4 hypervariable region) showing the affiliation of all three “*Ca. Atribacteria*” OTUs to the JS1 clade. (C) Phylogenetic tree based on 31 concatenated protein markers identified with AMPHORA2 (33). The monophyletic clade including our samples in the tree has >95% similarity to the sister clade with taxa from the Atlantic Ocean, Ross Sea, and North Pacific methane hydrate subseafloor sediment.

total sequences per sample). This showed that “*Ca. Atribacteria*” dominates the community (Fig. 1B), which is attributed to a single 16S OTU (Fig. 2). We considered the potential influence of multiple copies of 16S rRNA genes (25) on the result and searched for all “*Ca. Atribacteria*” genomes in the JGI database sequenced to date, which revealed that all sequenced genomes from this group have only one 16S rRNA gene copy, whereas *Deltaproteobacteria* and *Chloroflexi* have a median of two 16S rRNA gene copies (up to four) per genome (25). Thus, despite having an average lower 16S rRNA gene copy number than *Chloroflexi* and *Deltaproteobacteria*, “*Ca. Atribacteria*” relative abundance was high in our 16S rRNA gene sequence data set, even suggesting that our data actually underestimate the abundance of “*Ca. Atribacteria*.” These profiles show exponential net increases of “*Ca. Atribacteria*” and *Chloroflexi* within the upper 10 m of sediment, whereas those of *Deltaproteobacteria* increase at 0.4 mbsf but recede rapidly

over time from 10^4 to 10^3 copies or less, reaching our limit of detection (10^2 gene copies g^{-1} [wet weight] sediment) at the bottom of the core. The vertical profile of the beta subunit of the dissimilatory sulfate reductase (*dsrB*) gene runs parallel to that of *Deltaproteobacteria* 16S rRNA genes, which points to *Deltaproteobacteria* as the main sulfate-reducing bacteria (SRB) in this anoxic sediment. This inference is supported by the abrupt 10-fold drop in *dsrB* and *Deltaproteobacteria* 16S rRNA gene sequences at 10 mbsf (Fig. 1B), the same depth at which net SO_4^{2-} reduction decreases to below detection levels (Fig. 1A). The abundance of SRB is lower than in organic-rich shelf sediments (26, 27), which may be related to the relatively low sedimentation rate at our sampling location, since SO_4^{2-} reduction rates are correlated with the square of the sedimentation rate (9). The relative slowness of SO_4^{2-} reduction at our sampling location is readily evident in the profile of dissolved SO_4^{2-} concentrations, which remained >20 mM throughout the core (Fig. 1A).

Metagenomic analysis. Metagenomes from five depths were sequenced at an average depth of 8.4 million reads (± 2.5 million) (see Table S3). Because the DNA was amplified by multiple displacement amplification (MDA), we only considered presence/absence of genes and did not interpret the relative abundance of genes in the MDA-amplified metagenomes to be indicative of taxon abundance. *De novo* “binning”-based methods for creating metagenome-assembled genomes (MAGs) are useful for discovering new taxa when distantly related genomes in databases preclude similarity-based searches (28–31). Because of the ultralow DNA concentrations extractable from these abyssal clay sediments, we were only able to obtain metagenomic data after amplifying the extracted DNA using multiple displacement amplification (MDA). Our attempts at *de novo* binning of the MDA products revealed a selective amplification of short fragments that precluded binning and completion of high-quality MAGs (see Table S2). Specifically, manual curation of MaxBin results using Anvi'o (32) produced 12 bins at relatively low levels of genome completeness (17% to 44%) that were able to be assigned to “*Ca. Atribacteria*” (see Fig. S4). Thus, in order to increase the annotation of putative functions of open reading frames (ORFs) in metagenomes from the “*Ca. Atribacteria*” that could not be assembled into bins, we also applied a bioinformatics pipeline, whereby ORFs harbored by *de novo* assembled contigs were searched for similarity against a large aggregated genome database of predicted proteins from all atribacterial MAGs and single-cell genomes (SAGs) sequenced to date, including all published data from subsurface metagenomics studies (see Materials and Methods). We then extracted all the ORFs having a predicted protein from a “*Ca. Atribacteria*” genome as a top BLASTp hit and ran a phylogenomic analysis based on 31 phylogenomic markers using AMPHORA (33). This phylogenomic analysis demonstrates that our previously published method (34, 35) recovers ORFs with high similarity ($>95\%$ amino acid similarity) to predicted proteins in previously sequenced “*Ca. Atribacteria*” genomes (Fig. 2C). The phylogenomic analysis shows that the atribacterial ORFs from all sampled depths form a monophyletic clade sister to those from deep-sea sediments of the Atlantic Ocean, Ross Sea, and Pacific Ocean methane hydrates within the JS1 clade (Fig. 2C). This high level of similarity to existing “*Ca. Atribacteria*” genomes enabled similarity-based assignment of “*Ca. Atribacteria*” ORFs in our samples (see Fig. S5). As further evidence of this, the median similarity of ORFs to their top hit in the database in all metatranscriptome and metagenome data sets of $>60\%$ indicates that the ORFs in our samples had relatively high similarity to existing predicted proteins in the database (see Fig. S6).

We confirmed the accuracy of this previously published approach (34, 35) by an *in silico* test for true- and false-positive annotations based on 151 randomly selected peptide fragments extracted from “*Ca. Atribacteria*” predicted proteomes as well as other bacterial and archaeal genomes (see Fig. S7). In this analysis, 50 randomly selected predicted proteins were randomly cut into peptide fragments ranging from 20 to 140 amino acid residues in length in order to replicate partial ORFs typically recovered in metagenomes. The random peptide fragments were then searched

against the large aggregate database (34, 35) for their top hits with BLASTp. This *in silico* test showed that 100% of all randomly cut peptide fragments from atribacteria were true positives; they had the same atribacterial genome as a top BLASTp hit. This shows that our use of a similarity-based approach for ORF annotations (34–36) is adequate for assigning ORFs harbored by *de novo* assembled contigs to groups at high taxonomic levels, including those derived from the candidate phylum Atribacteria.

Gene expression analysis. Metatranscriptomes from eight depths were produced in biological replicates (two to seven replicates per depth) (Fig. 1D) and sequenced at an average depth of 4.0 million reads (± 1.5 million) with 91,199 contigs across all samples sequenced (Table S3). Similar to that for the 16S rRNA gene abundances, there was a subsurface peak in the number of unique expressed ORFs assigned to “*Ca. Atribacteria*” that increases exponentially between 3.5 and 10 mbsf, which was consistent across multiple replicate metatranscriptomes from each depth, and at 16 m, our RNA limit of detection was reached (Fig. 1D). We defined this depth as our RNA detection limit, because the number of unique ORFs assigned to “*Ca. Atribacteria*” was no longer detectable and, below this depth, the only ORFs annotated were those assigned to groups of known contaminants from molecular kits, including those from human skin and soil (37). Many of these same groups are common laboratory contaminants found in dust samples from our lab in 16S rRNA gene surveys (38) and include *Pseudomonas*, *Rhizobium*, *Acinetobacter*, and *Staphylococcus*. Presumably, the detection limit was reached because a smaller amount of extracted RNA from the *in situ* active community becomes overprinted by background “noise” from contaminating DNA (or RNA) derived from the kits, aerosols, or other laboratory contaminants.

There was a statistically significant correlation (r values = 0.59 and 0.61, P values = 0.016 and 0.014) between the abundance of atribacterial 16S rRNA genes and expressed ORFs (see Table S1). The number of expressed ORFs correlated with 16S rRNA gene quantities, which was consistent for the entire data set (total bacteria) and when comparing the number of ORFs expressed per group to the qPCR normalized abundance of the same taxonomic group for the four groups with the highest numbers of detected ORFs in the metatranscriptomes (*atribacteria*, *Deltaproteobacteria*, *Chloroflexi*, and *Archaea*) (Fig. 3A).

Evidence for subsurface reproduction. The steadily increasing abundance of “*Ca. Atribacteria*” over time (since sediment deposition) is apparently due to cells undergoing cytokinesis, as the 16S rRNA gene abundances correlated with the transcription of genes encoding proteins involved in cell division and cell shape determination such as *FtsAEKQWZ*, *MreBC*, and *RodA* (Fig. 3B) (39, 40). The expressed genes encoding the Fts proteins form the “divisome” (41), which includes the FtsZ ring, a cytokinetic protein ring that localizes at the cell division site prior to cytokinesis in dividing bacterial cells and pumps the replicated chromosome into the daughter cell (42). These genes are expressed during cellular division in “*Ca. Atribacteria*,” indicating the exponential growth phase (43). Thus, while we do not report measurements of biomass turnover such as amino acid racemization (44, 45) or stable isotope probing (4, 17), the steadily increasing microbial biomass of “*Ca. Atribacteria*” with increasing sediment depth together with the detected correlation between their abundance and transcriptional activity (Fig. 3A) and the transcription of ORFs encoding FtsZ ring proteins (Fig. 3B) strongly suggest that the higher abundances of “*Ca. Atribacteria*” between 0.5 and 10 mbsf are due to actively dividing cells. In this zone of apparent increased reproduction, the number of unique expressed ORFs correlated significantly with the number of 16S rRNA gene copies from the dominant groups (Fig. 1 and 3). The most parsimonious explanation for this is that more expressed ORFs from atribacteria are detected at these depths because there are higher numbers of metabolically active atribacteria producing mRNA transcripts.

The sediment ages in this interval span several million years; thus, this higher number of metabolically active microbes originated from “*Ca. Atribacteria*” that slowly reproduced through binary fission and cytokinesis over multimillion-year timescales.

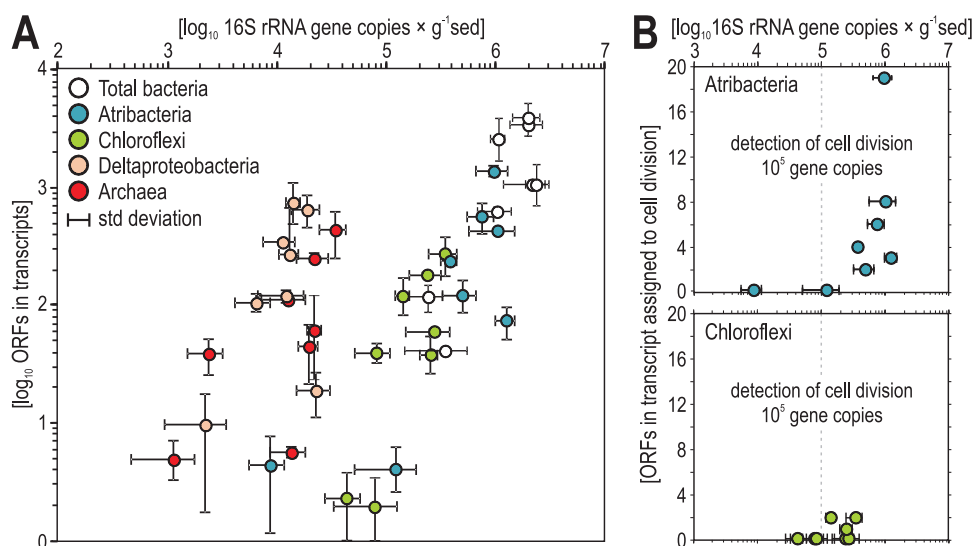


FIG 3 Group-specific transcriptional activity correlates with abundance. (A) Abundance of 16S rRNA genes (x axis) normalized to total bacteria (white), *Ca. Atribacteria* (blue), *Chloroflexi* (green), *Deltaproteobacteria* (light orange), and *Archaea* (red) 16S rRNA sequences plotted against their respective number of ORFs assigned to the same groups in the metatranscriptomes (y axis). Error bars represent standard deviations across biological replicates (in the case of three or more) or ranges (in the case of two replicates). (B) Abundance of 16S rRNA genes normalized to *Ca. Atribacteria* (blue) and *Chloroflexi* (green) plotted against their respective number of ORFs in the transcripts assigned to cell division (*FtsAEKQWZ*, *MreBC*, and *RodA*), showing that cell division is only detectable at 16S rRNA gene densities of $>10^5$ copies.

Because genes involved in the formation of the divisome, FtsZ ring, and cytokinesis were only expressed at depths where the highest 16S rRNA gene copies were found (Fig. 4B), the higher number of 16S rRNA gene copies at those depths is likely due to higher numbers of actively reproducing atribacterial cells. Active cells have higher numbers of ribosomes and thus higher copies of 16S rRNA per cell (46), but our qPCR assay targeted DNA (the bacterial chromosome), not RNA (expressed genes), thus discarding active rRNA synthesis as a confounding factor in our bacterial abundance estimations. Since *Ca. Atribacteria* has only one 16S rRNA gene copy, it is possible to conclude that the observed 10-fold increase in atribacterial 16S rRNA gene copy numbers over the top 10 mbsf requires chromosomal (genome) replication. In bacteria, genome replication occurs in actively dividing cells prior to cytokinesis and binary fission (46), assuming steady-state microbial input in sediment over millions of years. The only alternative explanation to our observations would require an increase in 16S rRNA copy number in the chromosome of *Ca. Atribacteria* over the top 10 mbsf followed by its subsequent decrease in the same chromosome below 10 mbsf. Such an incredibly high rate of genome evolution affecting the highly conserved 16S gene and then acting in a reversible way after 10 mbsf is inconceivable, especially considering this gene evolves at a rate of 1% every 100 million years (47) and our deepest sampled depth is roughly 10 million years old. Thus, the most likely explanation is that the detection of more 16S rRNA gene copies from the dominant OTU of *Ca. Atribacteria* (Fig. 2A) is due to more cells, each containing one chromosome with one copy of the 16S rRNA gene. The correlation of these 16S rRNA genes with higher numbers of expressed genes from *Ca. Atribacteria* over the top 15 mbsf (Fig. 1D and 3) can be attributed to a single clade (Fig. 2) that has been slowly reproducing and increasing in abundance over millions of years.

Predicted metabolism for *Ca. Atribacteria*. Our finding that *Ca. Atribacteria* expressed the highest number of protein-encoding genes among all other groups (Fig. 3) is furthermore consistent with their high levels of gene transcription in deep seafloor sediments of the Baltic Sea (48). Below 15 mbsf, the microbial ecosystem transitions to net death, since microbial abundances decrease by 2 orders of magni-

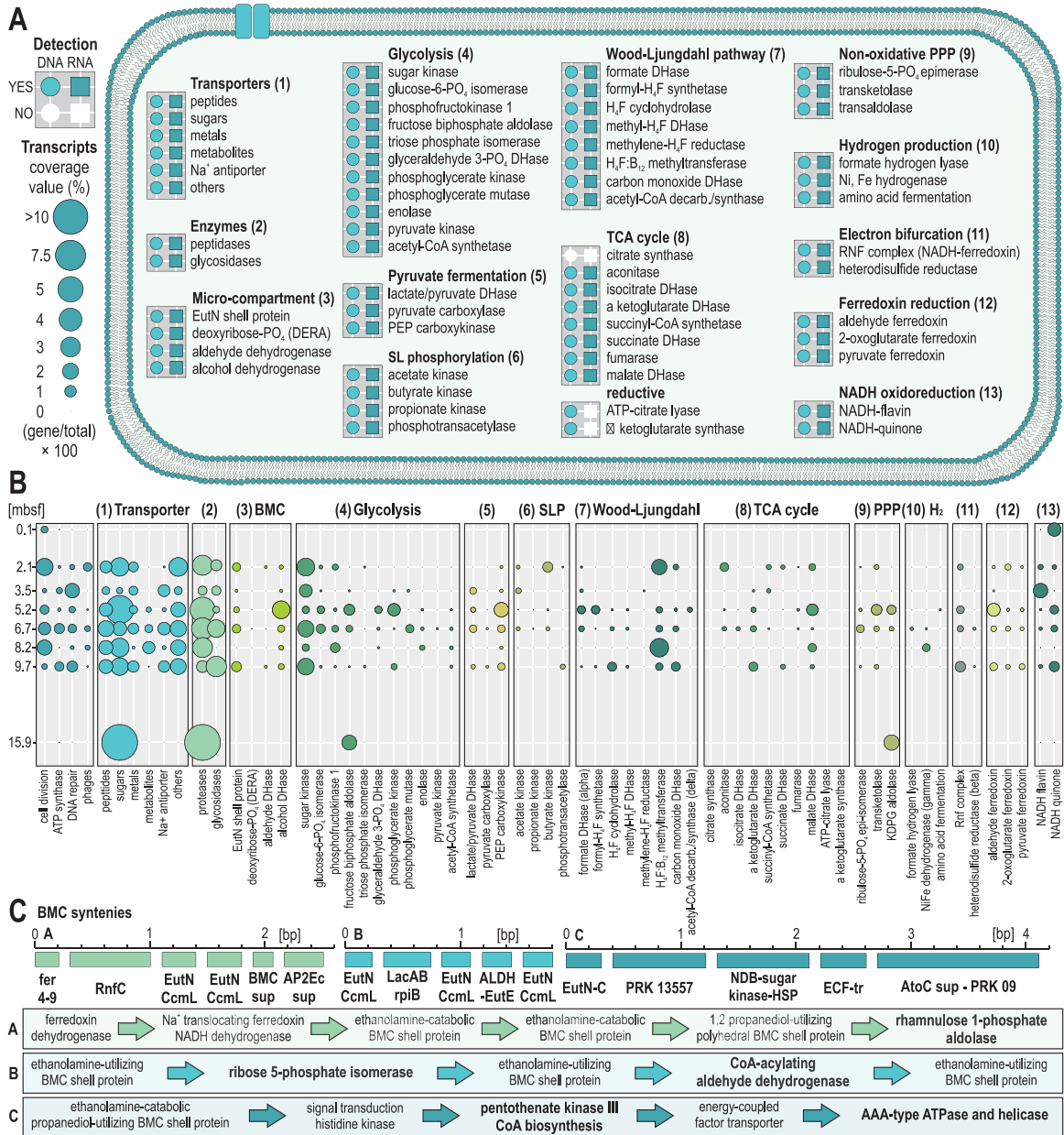


FIG 4 Metabolic potential and transcriptional activity of “*Ca. Atribacteria*” and its microcompartment. (A) Presence (colored) or absence (white) of ORFs assigned to “*Ca. Atribacteria*” encoding predicted proteins in metagenomes (circles) and metatranscriptomes (squares). (B) Bubble plots showing the coverage values (percent total reads) of expressed genes identified for “*Ca. Atribacteria*” at eight different depths. The numbering of the groups in panel A corresponds to the same numbering in panel B. DERA, 2-deoxy-D-ribose 5-phosphate aldolase. (C) Sequences of gene synteny related to the BMC present in the cell of “*Ca. Atribacteria*.” These three synteny provide evidence for metabolic use of aldehydes and alcohols via dehydrogenases with biosynthesis of acetyl-coenzyme A (acetyl-CoA) and regeneration of NAD⁺.

tude, RNA levels decline to below detection, and the net SO₄²⁻ reduction rate is below our detection limit (Fig. 1). However, “*Ca. Atribacteria*” remains the dominant group throughout the core, even below 10 mbsf. Its continued dominance is consistent with its abundance in other deep subseafloor settings (49–52) and the findings that dominant taxa in the subseafloor community do not necessarily require cellular reproduction to outcompete other taxa but can reach higher relative abundances from lower mortality than their competitors over long timescales (53, 54). This is supported by the lack of expression of ORFs by atribacteria harboring cellular division proteins below 10 mbsf, where atribacterial 16S rRNA gene abundances are <10⁵ copies per g (Fig. 3B).

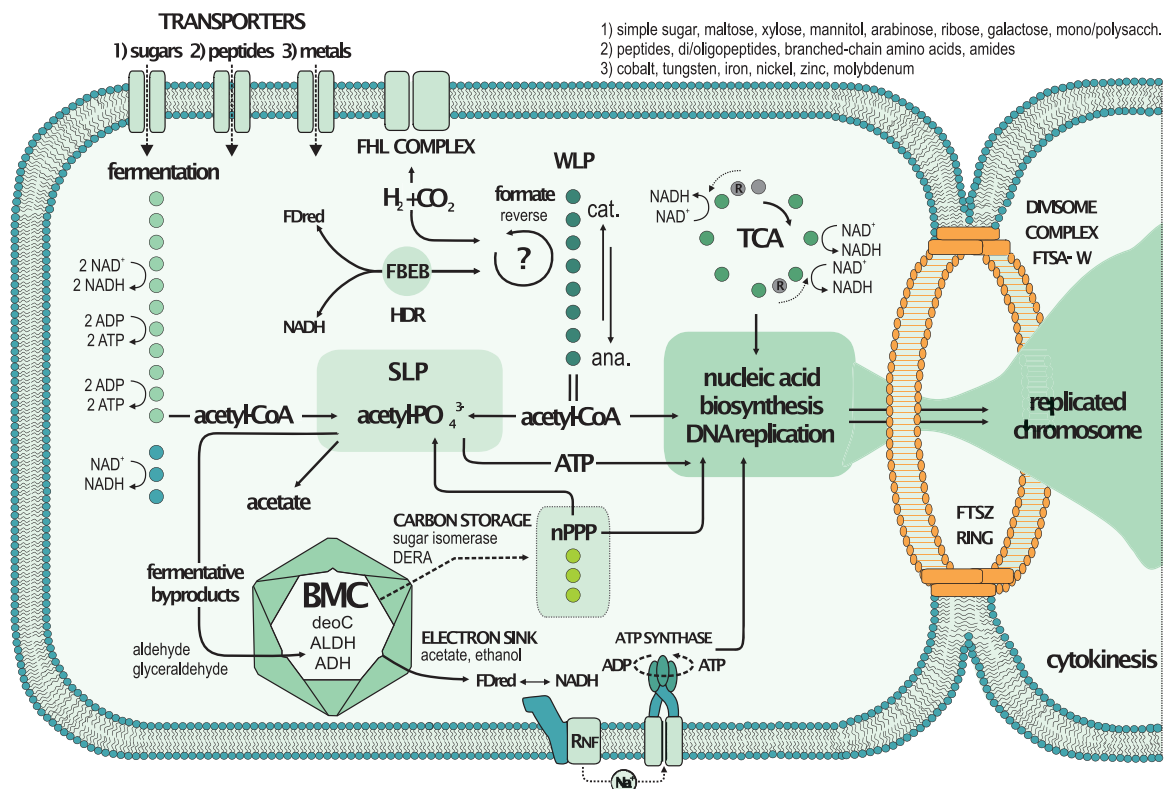


FIG 5 Potential “*Ca. Atribacteria*” physiology based on metagenomes and transcriptomes. The figure is based on the known interlinking of energy metabolism pathways for (homo)acetogenic bacteria (61) and BMC metabolic functions (57), as well as the function of the FtsZ ring and divisome (41). Colored dots correspond to the successive enzymatic steps (as listed in Fig. 4), e.g., fermentation corresponds to glycolysis (11 dots) and lactate fermentation (3 dots). Actively expressed genes encoding the divisome complex (i.e., *FtsAEKQWZ*) are indicated at the top right. nPPP, nonoxidative pentose phosphate pathway; SLP, substrate-level phosphorylation; DERA, 2-deoxy-D-ribose 5-phosphate aldolase; ALDH, aldehyde dehydrogenase; ADH, alcohol dehydrogenase; FBEB, flavin-based electron bifurcation; FHL, formate hydrogen lyase; WLP, Wood-Ljungdahl pathway.

Similar to what has been predicted from genomic studies (55–58), the gene transcription data strongly indicate that the “*Ca. Atribacteria*” dominating throughout the core actively utilizes a sugar-based acetogenic metabolism (Fig. 4A). This is also consistent with gene transcription from “*Ca. Atribacteria*” in deep subsurface Baltic Sea sediments that showed active utilization of trehalose (48).

The Wood-Ljungdahl pathway (WLP) has been found to be a widespread and potentially important metabolic pathway in subsurface microbial life (57, 59, 60). The WLP can be used catabolically to achieve redox balance and regenerate NAD⁺ and oxidized ferredoxin to thereby increase anaerobic metabolic efficiency (61). In acetogenic bacteria, this is coupled to the Rnf complex at the membrane that utilizes either ferredoxin or NAD⁺ as the terminal electron acceptor, driving a Na⁺ ion gradient and ATP synthesis (Fig. 5) (62). Several lines of evidence in our gene transcription data indicate that “*Ca. Atribacteria*” utilizes a similar metabolism.

Specifically, “*Ca. Atribacteria*” expressed transcripts encoding ORFs with similarity to proteins involved in sugar fermentation and the WLP (Fig. 4B). Genes encoding enzymes of the WLP were all present (Fig. 4A) and differently expressed by “*Ca. Atribacteria*” as were those encoding proteins involved in glycolysis, fermentation, and electron bifurcation (Fig. 4B). Transcription of ORFs by “*Ca. Atribacteria*” with similarity to enzymes known to be involved in electron bifurcation that results in production of molecular hydrogen included the beta subunit of the heterodisulfide reductase, formate hydrogen lyase (63), gamma subunit of the NiFe hydrogenase (64), and subunit alpha of the reversible formate dehydrogenase (65). The formate hydrogen lyase complex combines NiFe hydrogenase and soluble formate dehydrogenase to couple

formate and/or CO-dependent hydrogen production to the generation of the Na^+ motive force to generate ATP. This enzymatic complex allows growth on formate disproportionation (66), during which two electrons are passed to hydrogenase and two protons are neutralized (67). The gene expression data indicate that the “*Ca. Atribacteria*” produces the chemiosmotic Na^+ gradient using the Rnf complex for ATP synthesis (61). Atribacterial ORFs encoding the Rnf complex were expressed at the majority of depths (Fig. 4B), indicating this is indeed an important mechanism of anaerobic ATP synthesis.

We also detected active transcription of genes encoding pyruvate-formate lyase, acetate kinase, and formate dehydrogenase, which is reversible (65); therefore, the WLP could be used for either catabolic or anabolic purposes (Fig. 5). However, the expression of an $\text{H}_4\text{F}:\text{B}_{12}$ methyltransferase from “*Ca. Atribacteria*” is a direct indication that the WLP is functioning as an acetogenic pathway as opposed to one for methanogenesis, because acetogens use 5-methyltetrahydrofolate:corrinoid ($\text{H}_4\text{F}:\text{B}_{12}$) methyltransferase, whereas methanogens use tetrahydromethanopterin methyltransferase in the last step of methyl synthesis prior to acetyl synthesis (68). Expression of the $\text{H}_4\text{F}:\text{B}_{12}$ methyltransferase from “*Ca. Atribacteria*” was relatively high in the deeper samples at 8.2 and 9.7 mbsf (Fig. 4B), where the steadily increasing abundance of 16S rRNA genes with depth reached their peak values (Fig. 1B).

“*Ca. Atribacteria*” expression of transporter-encoding genes for multiple types of sugar (i.e., hexose, hexulose, pentose, pentulose, and ribose) and glycosidases (Fig. 4B) indicates a potential preference for sugar-based substrates. Sugars are in anoxic sediments as an energy substrate from bacterial necromass (e.g., ribose-containing nucleic acids DNA and RNA) and could be used as a fermentation substrate (69). The only detectable gene expression from “*Ca. Atribacteria*” observed at a 15.9-m depth indicates sugar transport, peptidases, fructose biphosphate, and phosphogluconate aldolase activities (Fig. 4B), again pointing that utilization of necromass (sugars and peptides from dead cells) is an important activity for long-term survival.

Fermentation products and other oxidized organic substrates (i.e., alcohols, ketones, aldehydes, and carboxylic acids) become toxic if they accumulate in the cell (70, 71). The gene expression from “*Ca. Atribacteria*” shows a metabolic potential for aldehydes to be reoxidized in secondary fermentations via a bacterial microcompartment (BMC), regenerating additional reduced ferredoxin and NADH in the process (72, 73). We obtained three contigs from “*Ca. Atribacteria*” with ORFs encoding BMC shell proteins syntenous with Rnf, aldehyde dehydrogenase, and 2-deoxy-D-ribose 5-phosphate aldolase (DERA) (Fig. 4C), similar to the already reported genome assembled from “*Ca. Atribacteria*” (57). These genes were coexpressed at multiple depths (Fig. 4B), suggesting that the BMC and Rnf-based establishment of the Na^+ ion gradient for ATP production are connected (Fig. 5). Collectively, these findings indicate that “*Ca. Atribacteria*” may use a BMC to recycle toxic intermediates produced during secondary fermentations for additional energy (Fig. 4 and 5). The ability to use a BMC to sustain secondary fermentations, in addition to primary fermentation of sugars (Fig. 4 and 5), may optimize energy conservation (e.g., increase ATP yields per mole substrate oxidized) under the extreme energy-limiting seafloor conditions. Similar to our results, a dominance of “*Ca. Atribacteria*” in organic lean sediments has been observed previously during the IODP expedition 313 to the New Jersey shelf (27). However, we acknowledge that biases inherent to all methods for determining the abundance of *Bacteria* and *Archaea* make quantitative comparisons difficult (5) and thus interpret the qPCR data showing a relatively lower abundance of *Archaea* (Fig. 1B) with caution.

Although fermenters and acetogens are often outcompeted by SRB in organic-rich seafloor sediments, the poor fitness of SRB in the sampled environment (Fig. 1B and C), compared to that of “*Ca. Atribacteria*,” may be related to the relatively low SO_4^{2-} reduction rates (Fig. 1A) and relatively low rates of organic matter burial at our sampled site (11, 19). In contrast, *Chloroflexi* related to *Dehalococcoidia*, which are predicted (homo)acetogens with metabolic potential to degrade complex organic compounds (59, 60), are relatively abundant in these deep-sea anoxic clays (Fig. 1B and C) and

transcribe genes involved in cell division (Fig. 3B). Our results lead us to speculate that organic-lean, anoxic abyssal seafloor sediment thus represents a niche that favors the reproduction of *Chloroflexi* and “*Ca. Atribacteria*” over that of other microbes, including SRB. Moreover, in culture experiments from the first cultivated representative of the atribacteria, the presence of an H₂-consuming methanogenic partner increased the growth rate of atribacteria by more than 100-fold (43). Based on this experimental result, we speculate that the relatively high abundance of the atribacteria in our sampled ancient anoxic sediment could be attributed to an as-of-yet-unidentified syntrophic or semisymbiotic partner organism, from the *Chloroflexi*, for example, *Dehalococcoidia* (74, 75).

Our data show that reproduction can occur in the seafloor over multimillion-year timescales. Although RNA from the “*Ca. Atribacteria*” was no longer detectable below 16 mbsf, metagenomes and 16S sequencing show that they subsist and remain the dominant group down to at least 29 mbsf. The extractable DNA with our protocol does not target extracellular DNA (20); thus, the DNA in these deeper sediments likely does not derive from “dead” DNA preserved from once-living cells. Therefore, we assume that “*Ca. Atribacteria*” remains viable, and potentially active, in the deeper sediments but at cell abundances and transcriptional activities that are too low for our current RNA-based methods. Thus, the survival of the “*Ca. Atribacteria*,” associated with gene expression at relatively low levels, supports the hypothesis that reduced metabolic activity is a fitness advantage (16) in the energy-starved seafloor.

MATERIALS AND METHODS

Sampling. All samples were taken by cruise KN223 of the R/V *Knorr* in the North Atlantic, from 26 October to 3 December 2014 (Woods Hole, MA, to Woods Hole, MA). At site KN223-15 (33°29.0'N, 54°10.0'W; water depth, 5,515 m), successively longer sediment cores were retrieved using a multicorer (~0.4 m), gravity corer (~3 m), and the Woods Hole Oceanographic Institution (WHOI) piston-coring device (~28 m). Additional details of sampling are published (11). Dissolved oxygen concentrations in the core sections (see Fig. S1 in the supplemental material) were measured with optical O₂ sensors as described previously (76). Concentrations of dissolved SO₄²⁻ were measured as published previously (6), and SO₄²⁻ reduction rates were calculated as previously described (14). Sediment subcores were retrieved on the ship aseptically using end-cut sterile syringes and kept frozen at -80°C until extraction in the home laboratory.

DNA extraction, quantitative PCR, and 16S rRNA gene libraries. For each sampled depth, we performed four biological replicates of DNA extraction. Total DNA was extracted from 0.7 g of sediment as previously described (20). DNA templates were diluted 10 times in ultrapure PCR water (Roche) and used in qPCR amplifications with updated 16S rRNA gene primer pair 515F (5'-GTG YCA GCM GCC GCG GTA A-3') and 806R (5'-GGA CTA CNV GGG TWT CTA AT-3') to increase our coverage of *Archaea* and marine clades (77) and run as previously described (38). All qPCRs were set up in 20- μ l volumes with 4 μ l of DNA template, 20 μ l SsoAdvanced SYBR green Supermix (Bio-Rad), 4.8 μ l nuclease-free H₂O (Roche), 0.4 μ l primers (10 μ M; biomers.net), and 0.4 μ l MgCl₂ and carried out on a CFX-Connect qPCR machine for gene quantification. For 16S rRNA genes, we ran 40 PCR cycles of two steps corresponding to denaturation at 95°C for 15 s and annealing and extension at 55°C for 30 s. To measure the abundance of *dsrB* genes, we used a previously described assay (78) with the primer pair dsrB4-R (5'-GTG TAG CAG TTA CCG CA-3') and dsrB2060F (5'-CAA CAT CGT YCA YAC CCA GGG-3'). All qPCRs were set up in 20- μ l volumes with 4 μ l of DNA template and performed as previously described (79). Gel purified amplicons of the 16S rRNA, *dsrB*, and *mcrA* genes were quantified in triplicates using QuantiT dsDNA reagent (Life Technologies) and used as a standard. An EpMotion 5070 automated liquid handler (Eppendorf) was used to set up all qPCRs and prepare the standard curve dilution series spanning from 10⁷ to 10¹ gene copies. Reaction efficiency values in all qPCR assays were between 90% and 110% with R² values of >0.95% for the standards.

For 16S rRNA gene library preparation, qPCR runs were performed with barcoded primer pair 515F and 806R. All 16S rRNA gene amplicons were purified from 1.5% agarose gels using the QIAquick gel extraction kit (Qiagen), quantified with the Qubit dsDNA HS assay kit (Thermo Fisher Scientific), normalized to 1 nM solutions, and pooled (38). Library preparation was carried out according to the MiniSeq System Denature and Dilute Libraries guide (Illumina). Sequencing was performed on all four biological replicates (Fig. S2) on the Illumina MiniSeq platform at the Geo-Bio LMU Center. We used USEARCH version 10.0.240 for MiniSeq read trimming and assembly, OTU picking, and 97% sequence identity clustering (80), which we showed previously captures an accurate diversity represented within mock communities sequenced on the same platform (38). OTU representative sequences were identified by BLASTn searches against SILVA database version 132 (81). To identify contaminants, 16S rRNA genes from extraction blanks and dust samples from the lab were also sequenced in triplicates (38). These 16S rRNA gene sequences were used to identify any contaminating bacteria (e.g., *Acinetobacter*, *Bacillus*, and *Staphylococcus*) and selectively curate the OTU table of our anoxic abyssal clay samples prior to downstream analysis. The 30 most abundant OTUs were aligned with SINA online v.1.2.11 (82) and

plotted in a maximum likelihood RAXML phylogenetic tree (83) (Fig. S3) against the SILVA 16S rRNA SSU NR99 reference database version 132 (81) using ARB (84). Closest environmental sequences with nearly full-length sequences (>1,400 bp) were selected as taxonomic references and used to calculate trees implemented with the bacterial and archaeal filter and advanced bootstrap refinement selecting the best tree among 300 replicates (84). Partial OTU sequences were added to the tree using the maximum parsimony algorithm without allowing changes of tree topology (Fig. 2 and S3).

Metagenome libraries. Whole-genome amplifications were performed on DNA extracts at $10\times$ dilutions through a multiple displacement amplification (MDA) step of 6 to 7 h, using the REPLI-g Midi kit (Qiagen) and according to the manufacturer's instructions. Dilution of the extracted DNA was necessary, because the MDA reaction was inhibited in undiluted DNA extracts (presumably from coextracted inhibitory chemicals). We added SYBR green I (Invitrogen) at $1,000\times$ concentration to visualize the amplification in real time on the CFX-Connect qPCR machine with a fluorescence reading taken every 10 min. Amplification was stopped after reaching the exponential increase by heating to 65°C for 3 min. Metagenomic libraries were prepared using the Nextera XT DNA Library Prep kit (Illumina), quantified on an Agilent 2100 bioanalyzer system (Agilent Genomics), and normalized with the Select-a-Size DNA Clean and Concentrator MagBead kit (Zymo Research) as previously described (20).

RNA extractions and metatranscriptome libraries. Total RNA extractions were obtained according to a previously published protocol (85). In brief, RNA was extracted from 3 g of sediments using the FastRNA Pro Soil-Direct kit (MP Biomedicals) according to the manufacturer's instructions, with the addition of $4\ \mu\text{l}$ glycogen ($0.1\ \text{g}\ \text{ml}^{-1}$) to increase yield during precipitation of the RNA pellet and final elution in $40\ \mu\text{l}$ PCR-grade water (Roche). Extraction blanks (no RNA added) were processed alongside to assess laboratory contamination, and sequencing of these contamination controls revealed common laboratory contaminants found in dust samples from our lab in 16S rRNA gene surveys (38), including *Pseudomonas*, *Rhizobium*, *Acinetobacter*, and *Staphylococcus*. RNA extracts were quantified using the QuBit RNA HS assay kit (Thermo Fisher Scientific). DNase treatment, synthesis of cDNA, and library construction were performed on the same day from $10\ \mu\text{l}$ of RNA templates using the Trio RNA-Seq kit protocol (NuGEN Technologies). Libraries were quantified as described above. All libraries were diluted to 1 nM and pooled for further sequencing on the MiniSeq platform (Illumina).

Gene identification and normalization in metagenomes. The SqueezeMeta (86) metagenomic analysis pipeline was used for downstream analysis of metagenomic reads in coassembly mode. For adapter removing, trimming, and quality filtering, we used Trimmomatic set to the following: leading, 8; trailing, 8; sliding window, 10:15; and minimum length, 30 (87). Contigs were assembled to minimum length of 200 bp using Megahit assembler (88). ORFs for genes and rRNAs were called using Prodigal (89), and rRNA genes were determined by barrnap (90). RDP classifier were used for the classification of 16S rRNA genes (91). Diamond software (92) was deployed for taxonomic assignment of retrieved gene homologies against the GenBank, eggNOG v. 4.5 (93), and KEGG (94) databases. Cutoff values for assigning hits to specific taxa were performed at an E value of $1 \times e^{-3}$, minimum amino acid similarity of 40 for taxa and 30 for functional assignment, using SqueezeMeta with default settings. Reads were mapped onto contigs and genes by using Bowtie2 (95). Coverage and transcripts per million (TPM) values were calculated using SqueezeMeta. For binning, we used MaxBin 2.0 (96) and MetaBAT (97), and bins generated by the two different algorithms were merged into one single data set using DAS Tool (98). Bin completeness and contamination were checked using CheckM (99). Further analysis of metagenome-assembled genomes (MAGs) based on results from SqueezeMeta was achieved using Anvi'o v. 6.2 (32) and MAGs selected by DAS tool further refined manually based on hierarchical clustering of contigs. The complete workflow using DAS Tool separates 14 bins whose taxonomic affiliations are uncertain. Manual curation of these results points to "*Candidatus* Aerophobetes," *Chloroflexi*, and "*Ca.* Atribacteria" as potential bin affiliations (Table S2). In comparison, from the manual curation of MaxBin results, we produced 31 bins (17% to 59%), among which, 12 bins are clearly assigned to "*Ca.* Atribacteria" with different levels of genome completeness (Fig. S4).

Because the MDA step produced many short fragments that did not allow high-quality binning and full genome completion, assigning a taxonomic affiliation to metagenomic and metatranscriptomic data is challenging (100). For annotating putative functions of ORFs in metagenomes and metatranscriptomes from particular "higher-level" taxonomic groups of microorganisms (34, 35), we also applied a bioinformatics pipeline whereby protein-encoding ORFs were extracted from *de novo* assembled contigs using FragGeneScan v. 1.30 (101) and functionally annotated against a large aggregated genome database (20, 34, 35) containing predicted proteins from all protist, fungal, bacterial, and archaeal genomes and MAGs in the JGI and NCBI databases using DIAMOND version 0.9.24 (92). This database, which we refer to as MetaProt, also contained all ORFs from all of the transcriptomes of microbial eukaryotes from the MMETS project (102), and we removed any hits to photosynthetic eukaryotic algae as contaminants. This custom MetaProt database that we used for this study is available as a single 32 GB amino acid fasta file on the LMU Open Data website (<https://data.uni-muenchen.de/183/>). Cutoff values for assigning hits to specific taxa were performed at a minimum bit score of 50, minimum amino acid similarity of 60, and an alignment length of 50 residues. All scripts and code used to produce the analysis have been posted on GitHub (<https://github.com/williamorsi/MetaProt-database>), and we provide a link to the MetaProt on the GitHub page as well as instructions within the scripts regarding how to conduct the workflows that we used. This approach assigns ORFs to higher-level taxonomic groups (35). As is the case in all metagenomic studies, the incomplete nature of genomes in databases, together with the lower representation of sequenced genomes from candidate clades than from cultured ones, makes it likely that our pipeline misses annotation of ORFs that are derived from as-yet-unsequenced atribacterial genomes. We

acknowledge that some genes in databases annotated as being present in “*Ca. Atribacteria*” might have been assigned to bins according to criteria that differ from study to study (Fig. S7).

Gene identification and normalization in metatranscriptomes. Paired-end reads were trimmed and assembled into contigs using CLC Genomics Workbench 9.5.4 (Qiagen, Hilden, Germany), using a word size of 20, bubble size of 50, and a minimum contig length of 300 nucleotides. Reads were then mapped to the contigs using the following parameters (mismatch penalty, 3; insertion penalty, 3; deletion penalty, 3; minimum alignment length, 50% of read length; minimum percent identity, 95%). Coverage values were obtained from the number of reads mapped to a contig divided by its length (i.e., average coverage). Only contigs with an average coverage of >5 were selected for ORF searches and downstream analysis (20, 34, 85). This protocol does not assemble rRNA; thus, results are only discussed in terms of mRNA. We then performed even further stringency controls by removing any contig that had less than $5\times$ coverage, e.g., reads per kilobase mapped (RPKM). The final resulting data set of contigs was then used for ORF searches and annotation against the MetaProt aggregate database, as described above.

For the metatranscriptomes, normalization of the relative abundance of ORFs was based on the number of unique ORFs assigned to a group (e.g., “*Ca. Atribacteria*”), a fractional percentage of total ORFs detected. We normalized expression in this manner as opposed to more conventional procedures such as RPKM because we found that the transcriptome sequencing (RNA-seq) kit we used has an amplification step (SPIA amplification, Trio RNAseq Ovation kit; NuGen) that biases the relative abundance of reads mapping to contigs when normalized using RPKM. For example, the RPKM value for the same ORF across technical replicates was found to have very large (orders of magnitude) variability in RPKM. In contrast, the total number of unique ORFs (e.g., presence/absence of an expressed ORF) assigned to specific groups (e.g., “*Ca. Atribacteria*”) was highly consistent between technical replicates. We assume that this technical variation in the RNA-seq data is associated with randomized SPIA amplification of different transcripts and/or fluctuations in the number of mRNA molecules in technical replicate tubes due to the highly labile nature of RNA during the extraction and library prep procedure. For this reason, we normalized the relative abundance of ORFs assigned to a specific group based on presence/absence of expressed ORFs, which was highly consistent between technical replicates despite the SPIA amplification. If significantly higher numbers of unique ORFs are detected from a particular group than from other groups, it can be attributed to relatively higher transcriptional activity.

COG categories were assigned by searching the ORFs against the COG database (103) using BLASTp. Metagenomic raw reads and metatranscriptomic atribacterial contigs were mapped with high stringency against a previously sequenced seafloor MAG of “*Ca. Atribacteria*” that has relatively high completeness (88%) as a reference (104), using Geneious 8.1.9. The average coverage after mapping against GenBank reference no. [NCRO00000000.1](https://ncbi.nlm.nih.gov/nucl/NCRO00000000.1) to 0.700 (104) was $224 (\pm 131)$, and 93% of ORFs in the reference MAG were detected in our seafloor metagenomes and metatranscriptomes. Consensus sequences were exported and annotated using the online tool RAST v. 2.0 (105). Taxonomic assignment of protein-encoding genes to “*Ca. Atribacteria*” clade JS1 was further confirmed in our metagenomes (Fig. S6) by selecting and aligning 31 phylogenomic markers from the corresponding metagenomic ORFs and 36 atribacterial reference genomes obtained from the NCBI database by using AMPHORA2 (33). Statistical analyses of beta-diversity were performed using RStudio v. 3.3.3 with the Bioconductor package (106).

Data availability. Data are publicly available through NCBI BioProject [PRJNA590088](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA590088). Metagenomes and metatranscriptomes have accession numbers [SAMN13317858](https://www.ncbi.nlm.nih.gov/acc/acc.cgi?acc=SAMN13317858) to [SAMN13317880](https://www.ncbi.nlm.nih.gov/acc/acc.cgi?acc=SAMN13317880). The 16S data are available in SRA BioSample accessions [SAMN10929403](https://www.ncbi.nlm.nih.gov/sra/SAMN10929403) to [SAMN10929517](https://www.ncbi.nlm.nih.gov/sra/SAMN10929517) and [SAMN13324854](https://www.ncbi.nlm.nih.gov/sra/SAMN13324854) to [SAMN13324920](https://www.ncbi.nlm.nih.gov/sra/SAMN13324920). Additional data related to this paper may be requested from the authors.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, JPG file, 0.8 MB.

FIG S2, JPG file, 0.8 MB.

FIG S3, JPG file, 1.0 MB.

FIG S4, JPG file, 0.5 MB.

FIG S5, JPG file, 0.6 MB.

FIG S6, JPG file, 0.6 MB.

FIG S7, JPG file, 0.9 MB.

TABLE S1, JPG file, 0.2 MB.

TABLE S2, JPG file, 2.2 MB.

TABLE S3, XLSX file, 0.1 MB.

ACKNOWLEDGMENTS

This work was supported primarily by the Deutsche Forschungsgemeinschaft (DFG) project OR 417/1-1 granted to W.D.O. Preliminary work was supported by the Center for Dark Energy Biosphere Investigations project OCE-0939564 also granted to W.D.O. The expedition was funded by the US National Science Foundation through grant NSF-OCE-1433150 to S.D. and R.P. R.W.M. led the expedition. Shipboard microbiology efforts

were supported by the Center for Dark Energy Biosphere Investigations (C-DEBI grant NSF-OCE-0939564). This is C-DEBI publication 545. This is a contribution of the Deep Carbon Observatory (DCO).

W.D.O. conceived the work and experimental approach. A.V., W.D.O., Ö.K.C., and S.V. contributed to the laboratory/bioinformatics analyses and experimental work. R.W.M., D.C.S., and R.P. obtained the samples during the KN223 R/V *Knorr* oceanographic expedition. W.D.O., R.W.M., D.C.S., A.V., S.V., and S.D. discussed and wrote the manuscript and commented on the paper.

We declare no competing interests.

REFERENCES

- Parkes RJ, John Parkes R, Cragg BA, Wellsbury P. 2000. Recent studies on bacterial populations and processes in subseafloor sediments: a review. *Hydrogeol J* 8:11–28. <https://doi.org/10.1007/PL00010971>.
- Inagaki F, Hinrichs K-U, Kubo Y, Bowles MW, Heuer VB, Hong W-L, Hoshino T, Ijiri A, Imachi H, Ito M, Kaneko M, Lever MA, Lin Y-S, Methe BA, Morita S, Morono Y, Tanikawa W, Bihan M, Bowden SA, Elvert M, Glombitza C, Gross D, Harrington GJ, Hori T, Li K, Limmer D, Liu C-H, Murayama M, Ohkouchi N, Ono S, Park Y-S, Phillips SC, Prieto-Mollar X, Purkey M, Riedinger N, Sanada Y, Sauvage J, Snyder G, Susilawati R, Takano Y, Tasumi E, Terada T, Tomaru H, Trembath-Reichert E, Wang DT, Yamada Y. 2015. Exploring deep microbial life in coal-bearing sediment down to ~2.5 km below the ocean floor. *Science* 349:420–424. <https://doi.org/10.1126/science.aaa6882>.
- Schippers A, Neretin LN, Kallmeyer J, Ferdelman TG, Cragg BA, John Parkes R, Jørgensen BB. 2005. Prokaryotic cells of the deep sub-seafloor biosphere identified as living bacteria. *Nature* 433:861–864. <https://doi.org/10.1038/nature03302>.
- Morono Y, Terada T, Nishizawa M, Ito M, Hillion F, Takahata N, Sano Y, Inagaki F. 2011. Carbon and nitrogen assimilation in deep subseafloor microbial cells. *Proc Natl Acad Sci U S A* 108:18295–18300. <https://doi.org/10.1073/pnas.1107763108>.
- Lloyd KG, May MK, Kevorkian RT, Steen AD. 2013. Meta-analysis of quantification methods shows that archaea and bacteria have similar abundances in the subseafloor. *Appl Environ Microbiol* 79:7790–7799. <https://doi.org/10.1128/AEM.02090-13>.
- D'Hondt S, Jørgensen BB, Miller DJ, Batzke A, Blake R, Cragg BA, Cypionka H, Dickens GR, Ferdelman T, Hinrichs K-U, Holm NG, Mitterer R, Spivack A, Wang G, Bekins B, Engelen B, Ford K, Gettemy G, Rutherford SD, Sass H, Skilbeck CG, Aiello IW, Guérin G, House CH, Inagaki F, Meister P, Naehr T, Niitsuma S, Parkes RJ, Schippers A, Smith DC, Teske A, Wiegand J, Padilla CN, Acosta JLS. 2004. Distributions of microbial activities in deep subseafloor sediments. *Science* 306:2216–2221. <https://doi.org/10.1126/science.1101155>.
- Orsi WD. 2018. Ecology and evolution of seafloor and subseafloor microbial communities. *Nat Rev Microbiol* 16:671–683. <https://doi.org/10.1038/s41579-018-0046-8>.
- Bowles MW, Mogollon JM, Kasten S, Zabel M, Hinrichs K-U. 2014. Global rates of marine sulfate reduction and implications for sub-sea-floor metabolic activities. *Science* 344:889–891. <https://doi.org/10.1126/science.1249213>.
- Jørgensen BB. 1982. Mineralization of organic matter in the sea bed—the role of sulphate reduction. *Nature* 296:643–645. <https://doi.org/10.1038/296643a0>.
- Jørgensen BB, Findlay AJ, Pellerin A. 2019. The biogeochemical sulfur cycle of marine sediments. *Front Microbiol* 10:849. <https://doi.org/10.3389/fmicb.2019.00849>.
- D'Hondt S, Inagaki F, Zarikian CA, Abrams LJ, Dubois N, Engelhardt T, Evans H, Ferdelman T, Gribsholt B, Harris RN, Hoppie BW, Hyun J-H, Kallmeyer J, Kim J, Lynch JE, McKinley CC, Mitsunobu S, Morono Y, Murray RW, Pockalny R, Sauvage J, Shimono T, Shiraishi F, Smith DC, Smith-Duque CE, Spivack AJ, Steinsbu BO, Suzuki Y, Szpak M, Toffin L, Uramoto G, Yamaguchi YT, Zhang G-L, Zhang X-H, Ziebis W. 2015. Presence of oxygen and aerobic communities from sea floor to basement in deep-sea sediments. *Nature Geosci* 8:299–304. <https://doi.org/10.1038/ngeo2387>.
- Kallmeyer J, Pockalny R, Adhikari RR, Smith DC, D'Hondt S. 2012. Global distribution of microbial abundance and biomass in subseafloor sediment. *Proc Natl Acad Sci U S A* 109:16213–16216. <https://doi.org/10.1073/pnas.1203849109>.
- D'Hondt S, Spivack AJ, Pockalny R, Ferdelman TG, Fischer JP, Kallmeyer J, Abrams LJ, Smith DC, Graham D, Hasiuk F, Schrum H, Stancin AM. 2009. Subseafloor sedimentary life in the South Pacific gyre. *Proc Natl Acad Sci U S A* 106:11651–11656. <https://doi.org/10.1073/pnas.0811793106>.
- D'Hondt S, Pockalny R, Fulfer VM, Spivack AJ. 2019. Subseafloor life and its biogeochemical impacts. *Nat Commun* 10:3519. <https://doi.org/10.1038/s41467-019-11450-z>.
- Hoehler TM, Jørgensen BB. 2013. Microbial life under extreme energy limitation. *Nat Rev Microbiol* 11:83–94. <https://doi.org/10.1038/nrmicro2939>.
- Lever MA, Rogers KL, Lloyd KG, Overmann J, Schink B, Thauer RK, Hoehler TM, Jørgensen BB. 2015. Life under extreme energy limitation: a synthesis of laboratory- and field-based investigations. *FEMS Microbiol Rev* 39:688–728. <https://doi.org/10.1093/femsre/fuv020>.
- Trembath-Reichert E, Morono Y, Ijiri A, Hoshino T, Dawson KS, Inagaki F, Orphan VJ. 2017. Methyl-compound use and slow growth characterize microbial life in 2-km-deep subseafloor coal and shale beds. *Proc Natl Acad Sci U S A* 114:E9206–E9215. <https://doi.org/10.1073/pnas.1707525114>.
- Zhao R, Hannisdal B, Mogollon JM, Jørgensen SL. 2019. Nitrifier abundance and diversity peak at deep redox transition zones. *Sci Rep* 9:8633. <https://doi.org/10.1038/s41598-019-44585-6>.
- Estes ER, Pockalny R, D'Hondt S, Inagaki F, Morono Y, Murray RW, Nordlund D, Spivack AJ, Wankel SD, Xiao N, Hansel CM. 2019. Persistent organic matter in oxic subseafloor sediment. *Nature Geosci* 12:126–131. <https://doi.org/10.1038/s41561-018-0291-5>.
- Vuillemin A, Wankel SD, Coskun ÖK, Magritsch T, Vargas S, Estes ER, Spivack AJ, Smith DC, Pockalny R, Murray RW, D'Hondt S, Orsi WD. 2019. Archaea dominate oxic subseafloor communities over multimillion-year time scales. *Sci Adv* 5:eaa4108. <https://doi.org/10.1126/sciadv.aaw4108>.
- Ziebis W, McManus J, Ferdelman T, Schmidt-Schierhorn F, Bach W, Muratli J, Edwards KJ, Villinger H. 2012. Interstitial fluid chemistry of sediments underlying the North Atlantic gyre and the influence of subsurface fluid flow. *Earth Planet Sci Lett* 323–324:79–91. <https://doi.org/10.1016/j.epsl.2012.01.018>.
- Parkes RJ, John Parkes R, Webster G, Cragg BA, Weightman AJ, Newberry CJ, Ferdelman TG, Kallmeyer J, Jørgensen BB, Aiello IW, Fry JC. 2005. Deep sub-seafloor prokaryotes stimulated at interfaces over geological time. *Nature* 436:390–394. <https://doi.org/10.1038/nature03796>.
- Newberry CJ, Webster G, Cragg BA, John Parkes R, Weightman AJ, Fry JC. 2004. Diversity of prokaryotes and methanogenesis in deep subsurface sediments from the Nankai Trough, ocean drilling program leg 190. *Environ Microbiol* 6:274–287. <https://doi.org/10.1111/j.1462-2920.2004.00568.x>.
- Lloyd KG, Bird JT, Buongiorno J, Deas E, Kevorkian R, Noordhoek T, Rosalsky J, Roy T. 25 March 2020. Evidence for a growth zone for deep subsurface microbial clades in near-surface anoxic sediments. <https://doi.org/10.1101/2020.03.24.005512>.
- Větrovský T, Baldrian P. 2013. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS One* 8:e57923. <https://doi.org/10.1371/journal.pone.0057923>.
- Schippers A, Kock D, Höft C, Köweker G, Siegert M. 2012. Quantification of microbial communities in subsurface marine sediments of the Black

- Sea and off Namibia. *Front Microbiol* 3:e16. <https://doi.org/10.3389/fmicb.2012.00016>.
27. Breuker A, Stadler S, Schippers A. 2013. Microbial community analysis of deeply buried marine sediments of the New Jersey shallow shelf (IODP expedition 313). *FEMS Microbiol Ecol* 85:578–592. <https://doi.org/10.1111/1574-6941.12146>.
 28. Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A, Wilkins MJ, Wrighton KC, Williams KH, Banfield JF. 2015. Unusual biology across a group comprising more than 15% of domain *Bacteria*. *Nature* 523: 208–211. <https://doi.org/10.1038/nature14486>.
 29. Castelle CJ, Wrighton KC, Thomas BC, Hug LA, Brown CT, Wilkins MJ, Frischkorn KR, Tringe SG, Singh A, Markillie LM, Taylor RC, Williams KH, Banfield JF. 2015. Genomic expansion of domain *Archaea* highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr Biol* 25:690–701. <https://doi.org/10.1016/j.cub.2015.01.014>.
 30. Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, van Eijk R, Schleper C, Guy L, Ettema TJG. 2015. Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521:173–179. <https://doi.org/10.1038/nature14447>.
 31. Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, Thomas BC, Singh A, Wilkins MJ, Karaoz U, Brodie EL, Williams KH, Hubbard SS, Banfield JF. 2016. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat Commun* 7:13219. <https://doi.org/10.1038/ncomms13219>.
 32. Eren AM, Murat Eren A, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3:e1319. <https://doi.org/10.7717/peerj.1319>.
 33. Wu M, Scott AJ. 2012. Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* 28:1033–1034. <https://doi.org/10.1093/bioinformatics/bts079>.
 34. Orsi WD, Richards TA, Francis WR. 2018. Predicted microbial secretomes and their target substrates in marine sediment. *Nat Microbiol* 3:32–37. <https://doi.org/10.1038/s41564-017-0047-9>.
 35. Orsi WD, Vuillemin A, Rodriguez P, Coskun ÖK, Gomez-Saez GV, Lavik G, Mohrholz V, Ferdelman TG. 2020. Metabolic activity analyses demonstrate that Lokiarchaeon exhibits homoacetogenesis in sulfidic marine sediments. *Nat Microbiol* 5:248–255. <https://doi.org/10.1038/s41564-019-0630-3>.
 36. Orsi WD, Morard R, Vuillemin A, Eitel M, Wörheide G, Milucka J, Kucera M. 2020. Anaerobic metabolism of Foraminifera thriving below the seafloor. *ISME J* 14:2580–2594. <https://doi.org/10.1038/s41396-020-0708-1>.
 37. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 12:87. <https://doi.org/10.1186/s12915-014-0087-z>.
 38. Pichler M, Coskun ÖK, Ortega-Arbulú A-S, Conci N, Wörheide G, Vargas S, Orsi WD. 2018. A 16S rRNA gene sequencing and analysis protocol for the Illumina MiniSeq platform. *Microbiologyopen* 7:e00611. <https://doi.org/10.1002/mbo3.611>.
 39. Pichoff S, Lutkenhaus J. 2007. Identification of a region of FtsA required for interaction with FtsZ. *Mol Microbiol* 64:1129–1138. <https://doi.org/10.1111/j.1365-2958.2007.05735.x>.
 40. Loose M, Mitchison TJ. 2014. The bacterial cell division proteins FtsA and FtsZ self-organize into dynamic cytoskeletal patterns. *Nat Cell Biol* 16:38–46. <https://doi.org/10.1038/ncb2885>.
 41. den Blaauwen T, den Blaauwen T, Hamoen LW, Levin PA. 2017. The divisome at 25: the road ahead. *Curr Opin Microbiol* 36:85–94. <https://doi.org/10.1016/j.mib.2017.01.007>.
 42. Erickson HP, Anderson DE, Osawa M. 2010. FtsZ in bacterial cytokinesis: cytoskeleton and force generator all in one. *Microbiol Mol Biol Rev* 74:504–528. <https://doi.org/10.1128/MMBR.00021-10>.
 43. Katayama T, Nobu MK, Kusada H, Meng X-Y, Yoshioka H, Kamagata Y, Tamaki H. 8 August 2019. Membrane-bounded nucleoid discovered in a cultivated bacterium of the candidate phylum "Atribacteria." *bioRxiv* <https://doi.org/10.1101/728279>.
 44. Braun S, Mhatre SS, Jaussi M, Røy H, Kjeldsen KU, Pearce C, Seidenkranz M-S, Jørgensen BB, Lomstein BA. 2017. Microbial turnover times in the deep seabed studied by amino acid racemization modelling. *Sci Rep* 7:e5680. <https://doi.org/10.1038/s41598-017-05972-z>.
 45. Møller MH, Glombitza C, Lever MA, Deng L, Morono Y, Inagaki F, Doll M, Su C-C, Lomstein BA. 2018. D-L-Amino acid modeling reveals fast microbial turnover of days to months in the subsurface hydrothermal sediment of Guaymas Basin. *Front Microbiol* 9:e967. <https://doi.org/10.3389/fmicb.2018.00967>.
 46. Madigan MT, Martinko JM, Stahl DA, Clark DP. 2012. Chapter 5: Microbial growth, p 118–119. *In* Espinoza D (ed), *Brock biology of microorganisms*. Pearson Education, Inc., San Francisco, CA.
 47. Ochman H, Elwyn S, Moran NA. 1999. Calibrating bacterial evolution. *Proc Natl Acad Sci U S A* 96:12638–12643. <https://doi.org/10.1073/pnas.96.22.12638>.
 48. Bird JT, Tague ED, Zinke L, Schmidt JM, Steen AD, Reese B, Marshall IPG, Webster G, Weightman A, Castro HF, Campagna SR, Lloyd KG. 2019. Uncultured microbial phyla suggest mechanisms for multi-thousand-year subsistence in Baltic Sea sediments. *mBio* 10:e02376-18. <https://doi.org/10.1128/mBio.02376-18>.
 49. Inagaki F, Suzuki M, Neelson KH, Horikoshi K, D'Hondt SL, Jørgensen BB. 2003. Subseafloor microbial diversity in the Peru Margin (ODP Leg. 201). *Geochim Cosmochim Acta* 67:171.
 50. Inagaki F, Suzuki M, Takai K, Oida H, Sakamoto T, Aoki K, Neelson KH, Horikoshi K. 2003. Microbial communities associated with geological horizons in coastal subseafloor sediments from the Sea of Okhotsk. *Appl Environ Microbiol* 69:7224–7235. <https://doi.org/10.1128/aem.69.12.7224-7235.2003>.
 51. Inagaki F, Nunoura T, Nakagawa S, Teske A, Lever M, Lauer A, Suzuki M, Takai K, Delwiche M, Colwell FS, Neelson KH, Horikoshi K, D'Hondt S, Jørgensen BB. 2006. Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean margin. *Proc Natl Acad Sci U S A* 103:2815–2820. <https://doi.org/10.1073/pnas.0511033103>.
 52. Fry JC, John Parkes R, Cragg BA, Weightman AJ, Webster G. 2008. Prokaryotic biodiversity and activity in the deep subseafloor biosphere. *FEMS Microbiol Ecol* 66:181–196. <https://doi.org/10.1111/j.1574-6941.2008.00566.x>.
 53. Starnawski P, Bataillon T, Ettema TJG, Jochum LM, Schreiber L, Chen X, Lever MA, Polz MF, Jørgensen BB, Schramm A, Kjeldsen KU. 2017. Microbial community assembly and evolution in subseafloor sediment. *Proc Natl Acad Sci U S A* 114:2940–2945. <https://doi.org/10.1073/pnas.1614190114>.
 54. Kirkpatrick JB, Walsh EA, D'Hondt S. 2019. Microbial selection and survival in subseafloor sediment. *Front Microbiol* 10:956. <https://doi.org/10.3389/fmicb.2019.00956>.
 55. Dodsworth JA, Blainey PC, Murugapiran SK, Swingley WD, Ross CA, Tringe SG, Chain PSG, Scholz MB, Lo C-C, Raymond J, Quake SR, Hedlund BP. 2013. Single-cell and metagenomic analyses indicate a fermentative and saccharolytic lifestyle for members of the OP9 lineage. *Nat Commun* 4:e1854. <https://doi.org/10.1038/ncomms2884>.
 56. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng J-F, Darling A, Malfatti S, Swan BK, Gies EA, Dodsworth JA, Hedlund BP, Tsiamis G, Sievert SM, Liu W-T, Eisen JA, Hallam SJ, Kyrpides NC, Stepanauskas R, Rubin EM, Hugenholtz P, Woyke T. 2013. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499:431–437. <https://doi.org/10.1038/nature12352>.
 57. Nobu MK, Dodsworth JA, Murugapiran SK, Rinke C, Gies EA, Webster G, Schwientek P, Kille P, John Parkes R, Sass H, Jørgensen BB, Weightman AJ, Liu W-T, Hallam SJ, Tsiamis G, Woyke T, Hedlund BP. 2016. Phylogeny and physiology of candidate phylum "Atribacteria" (OP9/JS1) inferred from cultivation-independent genomics. *ISME J* 10:273–286. <https://doi.org/10.1038/ismej.2015.97>.
 58. Liu Y-F, Qi Z-Z, Shou L-B, Liu J-F, Yang S-Z, Gu J-D, Mu B-Z. 2019. Anaerobic hydrocarbon degradation in candidate phylum "Atribacteria" (JS1) inferred from genomics. *ISME J* 13:2377–2390. <https://doi.org/10.1038/s41396-019-0448-2>.
 59. Sewell HL, Kaster A-K, Spormann AM. 2017. Homoacetogenesis in deep-sea *Chloroflexi*, as inferred by single-cell genomics, provides a link to reductive dehalogenation in terrestrial *Dehalococcoidetes*. *mBio* 8:e02022-17. <https://doi.org/10.1128/mBio.02022-17>.
 60. Fincker M, Huber JA, Orphan VJ, Rappé MS, Teske A, Spormann AM. 2020. Metabolic strategies of marine subseafloor *Chloroflexi* inferred from genome reconstructions. *Environ Microbiol* 22:3188–3204. <https://doi.org/10.1111/1462-2920.15061>.
 61. Schuchmann K, Müller V. 2016. Energetics and application of heterotrophy in acetogenic bacteria. *Appl Environ Microbiol* 82:4056–4069. <https://doi.org/10.1128/AEM.00882-16>.
 62. Hess V, Schuchmann K, Müller V. 2013. The ferredoxin: NAD⁺ oxidoreductase (Rnf) from the acetogen *Acetobacterium woodii* requires

- Na⁺ and is reversibly coupled to the membrane potential. *J Biol Chem* 288:31496–31502. <https://doi.org/10.1074/jbc.M113.510255>.
63. McDowall JS, Murphy BJ, Haumann M, Palmer T, Armstrong FA, Sargent F. 2014. Bacterial formate hydrogenlyase complex. *Proc Natl Acad Sci U S A* 111:E3948–E3956. <https://doi.org/10.1073/pnas.1407927111>.
 64. Vignais PM, Billoud B, Meyer J. 2001. Classification and phylogeny of hydrogenases. *FEMS Microbiol Rev* 25:455–501.
 65. Finney AJ, Sargent F. 2019. Formate hydrogenlyase: a group 4 [NiFe]-hydrogenase in tandem with a formate dehydrogenase. *Adv Microb Physiol* 74:465–486. <https://doi.org/10.1016/bs.ampbs.2019.02.004>.
 66. Hakoubyan M, Sargsyan H, Bagramyan K. 2005. Proton translocation coupled to formate oxidation in anaerobically grown fermenting *Escherichia coli*. *Biophys Chem* 115:55–61. <https://doi.org/10.1016/j.bpc.2005.01.002>.
 67. Schauer NL, Ferry JG. 1982. Properties of formate dehydrogenase in *Methanobacterium formicicum*. *J Bacteriol* 150:1–7. <https://doi.org/10.1128/JB.150.1.1-7.1982>.
 68. Sousa FL, Martin WF. 2014. Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism. *Biochim Biophys Acta* 1837:964–981. <https://doi.org/10.1016/j.bbabi.2014.02.001>.
 69. Orsi WD, Schink B, Buckel W, Martin WF. 2020. Physiological limits to life in anoxic subsurface sediment. *FEMS Microbiol Rev* 44:219–231. <https://doi.org/10.1093/femsre/fuaa004>.
 70. Wilbanks B, Trinh CT. 2017. Comprehensive characterization of toxicity of fermentative metabolites on microbial growth. *Biotechnol Biofuels* 10:262. <https://doi.org/10.1186/s13068-017-0952-4>.
 71. LaRowe DE, Van Cappellen P. 2011. Degradation of natural organic matter: a thermodynamic analysis. *Geochim Cosmochim Acta* 75:2030–2042. <https://doi.org/10.1016/j.gca.2011.01.020>.
 72. Axen SD, Erbilgin O, Kerfeld CA. 2014. A taxonomy of bacterial micro-compartment loci constructed by a novel scoring method. *PLoS Comput Biol* 10:e1003898. <https://doi.org/10.1371/journal.pcbi.1003898>.
 73. Kerfeld CA, Erbilgin O. 2015. Bacterial microcompartments and the modular construction of microbial metabolism. *Trends Microbiol* 23:22–34. <https://doi.org/10.1016/j.tim.2014.10.003>.
 74. Vuillemin A, Kerrigan Z, D'Hondt S, Orsi WD. 27 May 2020. Chloroflexi persisting for millions of years in oxic and anoxic deep-sea clay. *bioRxiv* <https://doi.org/10.1101/2020.05.26.116590>.
 75. Zhuang W-Q, Yi S, Bill M, Brisson VL, Feng X, Men Y, Conrad ME, Tang YJ, Alvarez-Cohen L. 2014. Incomplete Wood-Ljungdahl pathway facilitates one-carbon metabolism in organohalide-respiring *Dehalococcoides mccartyi*. *Proc Natl Acad Sci U S A* 111:6419–6424. <https://doi.org/10.1073/pnas.1321542111>.
 76. Fischer JP, Ferdelman TG, D'Hondt S, Røy H, Wenzhöfer F. 2009. Oxygen penetration deep into the sediment of the South Pacific gyre. *Biogeosciences* 6:1467–1478. <https://doi.org/10.5194/bg-6-1467-2009>.
 77. Parada AE, Needham DM, Fuhrman JA. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol* 18:1403–1414. <https://doi.org/10.1111/1462-2920.13023>.
 78. Jochum LM, Chen X, Lever MA, Loy A, Jørgensen BB, Schramm A, Kjeldsen KU. 2017. Depth distribution and assembly of sulfate-reducing microbial communities in marine sediments of Aarhus Bay. *Appl Environ Microbiol* 83:e01547-17. <https://doi.org/10.1128/AEM.01547-17>.
 79. Coskun ÖK, Özen V, Wankel SD, Orsi WD. 2019. Quantifying population-specific growth in benthic bacterial communities under low oxygen using H₂¹⁸O. *ISME J* 13:1546–1559. <https://doi.org/10.1038/s41396-019-0373-4>.
 80. Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 10:996–998. <https://doi.org/10.1038/nmeth.2604>.
 81. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–D596. <https://doi.org/10.1093/nar/gks1219>.
 82. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glockner FO. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35:7188–7196. <https://doi.org/10.1093/nar/gkm864>.
 83. Stamatakis A. 2006. RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690. <https://doi.org/10.1093/bioinformatics/btl446>.
 84. Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadukumar Buchner A, Lai T, Steppi S, Jobb G, Förster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, König A, Liss T, Lüßmann R, May M, Nonhoff B, Reichel B, Strehlow R, Stamatakis A, Stuckmann N, Vilbig A, Lenke M, Ludwig T, Bode A, Schleifer K-H. 2004. ARB: a software environment for sequence data. *Nucleic Acids Res* 32:1363–1371. <https://doi.org/10.1093/nar/gkh293>.
 85. Ortega-Arbulú A, Pichler M, Vuillemin A, Orsi WD. 2019. Effects of organic matter and low oxygen on the mycobenthos in a coastal lagoon. *Environ Microbiol* 21:374–388. <https://doi.org/10.1111/1462-2920.14469>.
 86. Tamames J, Puente-Sánchez F. 2019. SqueezeMeta, a highly portable, fully automatic metagenomic analysis pipeline. *Front Microbiol* 9:e3349. <https://doi.org/10.3389/fmicb.2018.03349>.
 87. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
 88. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31:1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>.
 89. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
 90. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
 91. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73:5261–5267. <https://doi.org/10.1128/AEM.00062-07>.
 92. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>.
 93. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, von Mering C, Bork P. 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. <https://doi.org/10.1093/nar/gky1085>.
 94. Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30. <https://doi.org/10.1093/nar/28.1.27>.
 95. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
 96. Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32:605–607. <https://doi.org/10.1093/bioinformatics/btv638>.
 97. Kang DD, Froula J, Egan R, Wang Z. 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* 3:e1165. <https://doi.org/10.7717/peerj.1165>.
 98. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF. 2018. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol* 3:836–843. <https://doi.org/10.1038/s41564-018-0171-1>.
 99. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
 100. Breitwieser FP, Lu J, Salzberg SL. 2019. A review of methods and databases for metagenomic classification and assembly. *Brief Bioinform* 20:1125–1136. <https://doi.org/10.1093/bib/bbx120>.
 101. Rho M, Tang H, Ye Y. 2010. FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res* 38:e191. <https://doi.org/10.1093/nar/gkq747>.
 102. Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA, Armbrust EV, Archibald JM, Bharti AK, Bell CJ, Beszteri B, Bidle KD, Cameron CT, Campbell L, Caron DA, Cattolico RA, Collier JL, Coyne K, Davy SK, Deschamps P, Dyrman ST, Edvardsen B, Gates RD, Gobler CJ, Greenwood SJ, Guida SM, Jacobi JL, Jakobsen KS, James ER, Jenkins B, John U, Johnson MD, Juhl AR, Kamp A, Katz LA, Kiene R, Kudryavtsev A, Leander BS, Lin S, Lovejoy C, Lynn D, Marchetti A, McManus G, Nedelcu AM, Menden-Deuer S, Miceli C, Mock T, Montresor M, Moran MA, Murray S, et al. 2014. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of

- eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol* 12:e1001889. <https://doi.org/10.1371/journal.pbio.1001889>.
103. Galperin MY, Makarova KS, Wolf YI, Koonin EV. 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res* 43:D261–D269. <https://doi.org/10.1093/nar/gku1223>.
104. Lee YM, Hwang K, Lee JI, Kim M, Hwang CY, Noh H-J, Choi H, Lee HK, Chun J, Hong SG, Shin SC. 2018. Genomic insight into the predominance of candidate phylum Atribacteria JS1 lineage in marine sediments. *Front Microbiol* 9:2909. <https://doi.org/10.3389/fmicb.2018.02909>.
105. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>.
106. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T, Gottardo R, Hahne F, Hansen KD, Irizarry RA, Lawrence M, Love MI, MacDonald J, Obenchain V, Oleś AK, Pagès H, Reyes A, Shannon P, Smyth GK, Tenenbaum D, Waldron L, Morgan M. 2015. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods* 12:115–121. <https://doi.org/10.1038/nmeth.3252>.