



Complete arsenic-based respiratory cycle in the marine microbial communities of pelagic oxygen-deficient zones

Jaclyn K. Saunders^{a,1,2}, Clara A. Fuchsman^{a,3}, Cedar McKay^a, and Gabrielle Rocap^{a,1}

^aSchool of Oceanography, University of Washington, Seattle, WA 98195

Edited by David M. Karl, University of Hawaii, Honolulu, HI, and approved April 1, 2019 (received for review December 12, 2018)

Microbial capacity to metabolize arsenic is ancient, arising in response to its pervasive presence in the environment, which was largely in the form of As(III) in the early anoxic ocean. Many biological arsenic transformations are aimed at mitigating toxicity; however, some microorganisms can respire compounds of this redox-sensitive element to reap energetic gains. In several modern anoxic marine systems concentrations of As(V) are higher relative to As(III) than what would be expected from the thermodynamic equilibrium, but the mechanism for this discrepancy has remained unknown. Here we present evidence of a complete respiratory arsenic cycle, consisting of dissimilatory As(V) reduction and chemoautotrophic As(III) oxidation, in the pelagic ocean. We identified the presence of genes encoding both subunits of the respiratory arsenite oxidase AioA and the dissimilatory arsenate reductase ArrA in the Eastern Tropical North Pacific (ETNP) oxygen-deficient zone (ODZ). The presence of the dissimilatory arsenate reductase gene *arrA* was enriched on large particles (>30 μm), similar to the forward bacterial *dsrA* gene of sulfate-reducing bacteria, which is involved in the cryptic cycling of sulfur in ODZs. Arsenic respiratory genes were expressed in metatranscriptomic libraries from the ETNP and the Eastern Tropical South Pacific (ETSP) ODZ, indicating arsenotrophy is a metabolic pathway actively utilized in anoxic marine water columns. Together these results suggest arsenic-based metabolisms support organic matter production and impact nitrogen biogeochemical cycling in modern oceans. In early anoxic oceans, especially during periods of high marine arsenic concentrations, they may have played a much larger role.

oxygen deficient zones | arsenic | chemoautotrophy | dissimilatory arsenate reduction | marine metagenome

For most of Earth's history the oceans were characterized by a lack of available oxygen and as a result ancient microbes employed a variety of alternate electron acceptors (1). Proxies for these early environments are found in marine pelagic oxygen-deficient zones (ODZs) where redox gradients through the water column impact the associated microbial communities (2, 3). Naturally occurring marine ODZs are created by a combination of organic matter respiration and slow midlayer ventilation (4) and are functionally anoxic (<10 nmol L^{-1}) (2, 5). Denitrifying metabolisms, which reduce nitrate and nitrite, are prominent (3) and ultimately result in the production of N_2O and N_2 gas and thus loss of N from the marine system. There is also evidence for a cryptic sulfur cycle in marine ODZs, where dissimilatory sulfate reduction produces sulfide, which is consumed by autotrophic sulfur oxidizers (6).

The oxidized inorganic arsenic compound arsenate is the thermodynamically stable form in oxygenated aqueous environments [As(V) as H_2AsO_4^- and HAsO_4^{2-}] and arsenite is the thermodynamically stable form in anoxic environments [As(III) as H_3AsO_3^0 and H_2AsO_3^-] (7). Microbes utilizing the arsenic redox potential for energetic gains have been studied and isolated in a range of systems with moderate to high arsenic concentrations, including polluted and unpolluted soils, sediments,

hot springs, lakes, wastewater, and mine drainage (8). In some arsenic-rich aquatic systems, like the alkaline Mono Lake in California (USA), arsenic supports microbial communities through a cycle of dissimilatory heterotrophic reduction of As(V) and autotrophic oxidation of As(III), analogous to the cryptic sulfur cycle in ODZs (7). A complete metabolic arsenic cycle is ancient, existing since the Archaean (9, 10), as suggested by fossilized arsenic-rich organic globules from 2.72 billion years ago (11). Arsenic concentrations have fluctuated over geologic time, with early anoxic oceans likely experiencing periods of greater arsenic abundance than today (12, 13).

While the modern ocean is not considered an arsenic-loaded system, inorganic arsenic is ubiquitous in marine waters, ranging from 15 to 20 nmol/L in the open ocean (14, 15). Measurements of arsenic in the anoxic waters of the Eastern Tropical South Pacific (16), Black Sea (17), and the anoxic fjord Saanich Inlet, British Columbia (18) show thermodynamic disequilibrium of the arsenic species in the water column which may be due, in part, to the microbial biotransformations of arsenic species. In radioisotope arsenic tracer experiments carried out on the anoxic waters from Saanich Inlet, rates of arsenite oxidation greatly diminished when antibiotics were added to the seawater (18),

Significance

Marine oxygen-deficient zones (ODZs) are naturally occurring midlayer oxygen-poor regions of the ocean, sandwiched between oxygenated surface and deep layers. In the absence of oxygen, microorganisms in ODZs use other compounds, such as oxidized forms of nitrogen and sulfur, as terminal electron acceptors. We identified the presence and expression of genes for both arsenic reduction and oxidation in marine ODZs, suggesting the microbial community in these waters is also cycling arsenic for respiratory gain. The existence of an arsenic respiratory cycle in pelagic waters suggests microbial arsenic metabolisms may be underestimated in the modern ocean and were likely an even more significant contributor to biogeochemical cycles in the anoxic ancient oceans when arsenic concentrations were higher.

Author contributions: J.K.S. and G.R. designed research; J.K.S., C.A.F., and C.M. performed research; J.K.S. analyzed data; and J.K.S. and G.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (GenBank Bio Project PRJNA350692).

¹To whom correspondence may be addressed. Email: jsaunders@whoi.edu or rocap@uw.edu.

²Present address: Marine Chemistry & Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA 02543.

³Present address: Horn Point Laboratory, University of Maryland, Cambridge, MD 21613.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1818349116/-DCSupplemental.

Published online April 29, 2019.

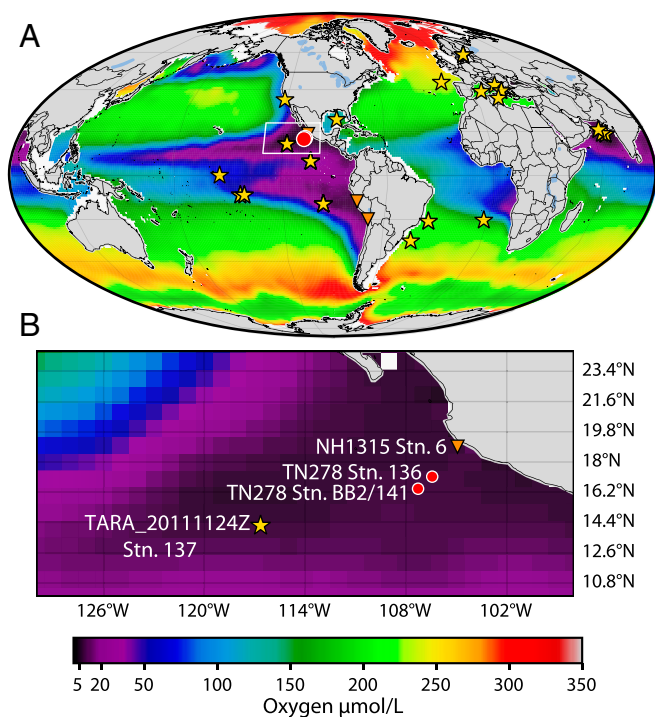


Fig. 1. Maps displaying interpolated oxygen concentrations from 300-m depth from the World Ocean Atlas (19). (A) Global map with highlighted points. Red circles designate metagenomic sampling sites from locations described here. Orange triangles represent publicly available metatranscriptomes (34–36). Stars represent locations where arsenotrophy genes were identified in metagenomes from NCBI WGS database (*SI Appendix, Table S2*). The white box highlights the location of origin of metagenomes described here with a zoomed-in view of the region in B.

suggesting microbially mediated impacts on arsenic transformations. We hypothesized that the apparent redox disequilibrium of arsenic species in anoxic marine waters may be due to the presence of arsenotrophic microbes.

Results and Discussion

Full-Length Arsenic Respiratory Genes in ODZ Metagenomes. Samples of the marine microbial community were collected on an oceanographic research cruise through the Eastern Tropical North Pacific (ETNP) ODZ in April of 2012. Filtered samples were collected at two oxygen-deficient stations (Fig. 1 and *SI Appendix, Fig. S1*) (19) including multiple depths through the anoxic core of the ODZ water column at station 136, and both free-living and particle-attached (>30 μm) fractions at 120 m at station BB2/141 (20). Metagenomic reads were assembled into contigs and searched for complete gene sequences for enzymes capable of arsenite oxidation (*aioAB*) and dissimilatory arsenate reduction (*arrAB*). The alpha subunits of these metabolic arsenic cycling enzymes, *aioA* and *arrA*, are found within the DMSO reductase superfamily, also known as the complex iron sulfur molybdoenzymes. This superfamily also contains other enzymes critical in respiratory redox transformations like the nitrate reductases Nap and Nar, formate dehydrogenases Fdn and Fdh, polysulfide reductase Psr, and tetrathionate reductase Ttr, among others (21). Potential arsenotrophic genes were confirmed through phylogenetic inference within the larger DMSO reductase superfamily.

A complete dissimilatory arsenate reductase alpha subunit (*arrA*) sequence was identified in the ETNP ODZ metagenome assemblies (Fig. 2A); the beta subunit was also identified downstream of *arrA* on the same contig (Fig. 2B) and is phylogenetically

closely related to the beta subunits of known *ArrB* sequences (*SI Appendix, Fig. S2*) (22). This genomic arrangement is also found in known dissimilatory arsenate-reducing prokaryotes (DARPs) (23), a phylogenetically diverse group of heterotrophs which obtain metabolic energy from the reduction of arsenate, typically using organic compounds as electron donors, although there are examples of hydrogen or reduced sulfur compounds used as electron donors (8). This *ArrA* sequence contains conserved residues for integral binding motifs found in the active site of functional respiratory arsenate reductase enzymes (*SI Appendix, Fig. S3*) (24). Notably, no reverse respiratory arsenate reductase (*arxA*) sequences were assembled—*ArxA* resembles the respiratory arsenate reductase of the DARPs (*ArrA*) in sequence, but functions in a reverse direction by oxidizing arsenite to arsenate (25).

Genes for the complementary oxidation reaction in the arsenic respiratory cycle, arsenite oxidase, were also found in the ETNP ODZ. The large subunit of arsenite oxidase, *AioA*, is evolutionarily distinct from the *ArrA/ArxA* group and is found elsewhere on the DMSO reductase superfamily tree (9) (*SI Appendix, Fig. S4*). The *Aio* enzyme can be used in both dissimilatory arsenite oxidation, believed to be a detoxification strategy (8), and in chemoautotrophic arsenite oxidation, where metabolic gains are made from reduction of the respiratory chain via electron conservation during the oxidation of As(III) to As(V) (26). Most known chemoautotrophic arsenite oxidizers are strict mesophilic aerobes (8, 26); however, there are a few representatives which can utilize nitrate as the terminal electron acceptor (27–29), with some strains capable of complete denitrification by reduction of NO_3 to N_2 (29). Phylogenetic analysis of assembled sequences from the ETNP ODZ identified a complete *aioA* sequence that branched basally to a clade of sequences from Alphaproteobacteria in the Rhizobiales order with known chemoautotrophic arsenite oxidation capabilities (Fig. 3A). Genes encoding a beta subunit of arsenite oxidase (*aioB*)

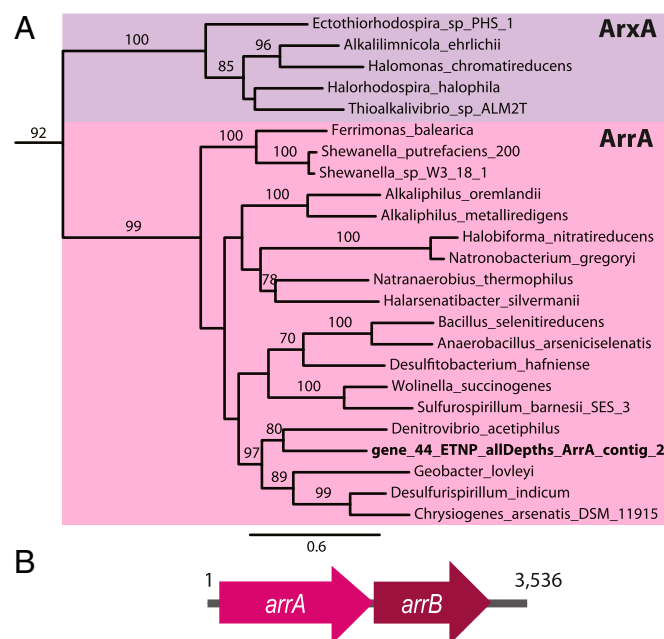


Fig. 2. Full-length dissimilatory arsenate reductase (*ArrA*) sequence identified in ETNP ODZ metagenomes. (A) A maximum-likelihood tree of *ArrA* and the closely related arsenite oxidase (*ArxA*) sequences including a full de novo assembled environmental *ArrA* sequence from the ETNP ODZ metagenome in bold. Bootstrap values ≥ 70 are shown. (B) The gene complement found along the contig *ArrA_ETNP_allDepths_contig_2* showing the presence of a downstream arsenate reductase beta subunit *arrB* sequence.

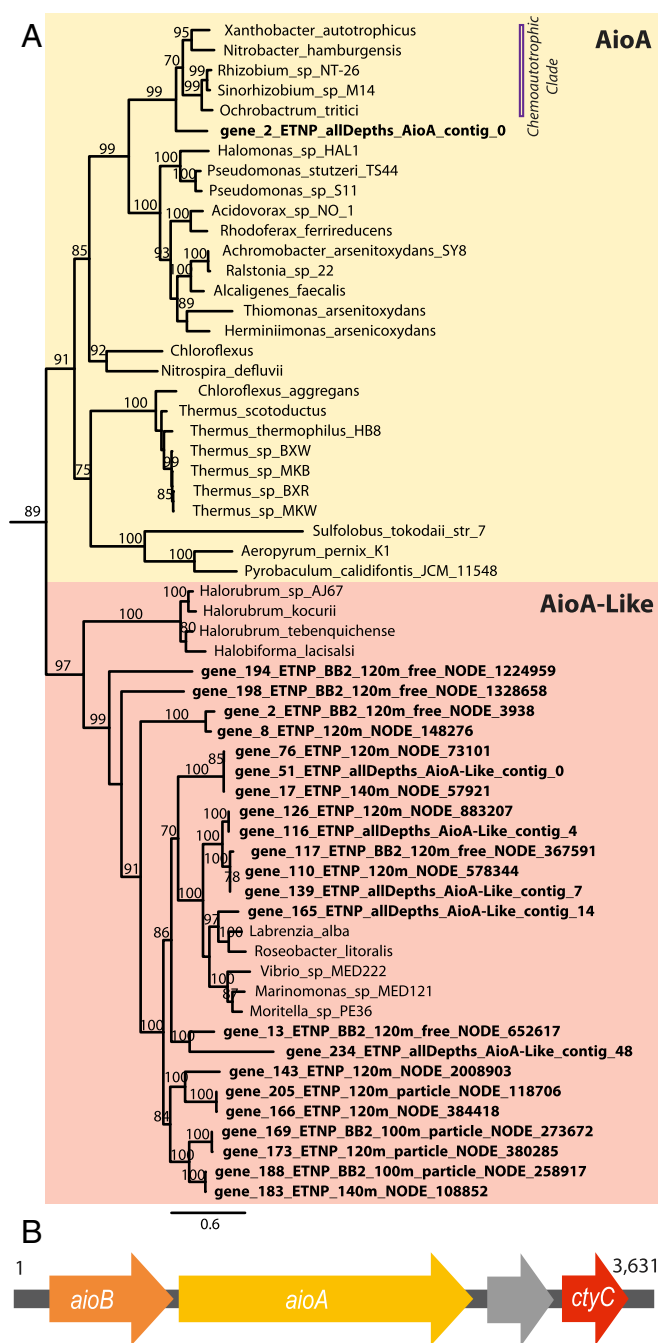


Fig. 3. Full-length respiratory arsenite oxidase (AioA) sequence identified in ETNP ODZ metagenomes. (A) A maximum-likelihood tree of AioA highlighted in yellow and the closely related enzyme arsenite oxidase-like (AioA-Like) in red-orange. Full length de novo assembled environmental sequences from the ETNP are denoted by bold text. The clade which contains the environmental AioA sequence and sequences associated with chemoautotrophic activity is designated with a bar. Bootstrap values ≥ 70 are shown. (B) The gene complement found along the contig AioA_ETNP_allDepths_contig_0 showing the presence of the beta subunit sequence *aioB* and a cytochrome-C.

upstream and a complete cytochrome-c located downstream (Fig. 3B) were also present on the *aioA*-containing contig (22), a genomic organization similar to known arsenite oxidizers (26). The *aioB* subunit is also phylogenetically more closely related to beta subunits of chemoautotrophic arsenite oxidizers (SI Appendix, Fig. S5) than to dissimilatory arsenite oxidizers. Therefore, it is

likely that the environmental assembled contig containing *aioA* is representative of a chemoautotrophic arsenite oxidation pathway and may be capable of denitrification.

Numerous contigs assembled from the ETNP ODZ contained sequences closely related to AioA but branching just outside the internal node defining the AioA bacterial and archaeal clade in the DMSO reductase superfamily tree (Fig. 3A and SI Appendix, Fig. S4). This clade of sequences has been previously identified as “AioA-Like” (9) and consists of a subclade made up of a handful of Alpha- and Gammaproteobacterial sequences and another subclade of sequences recently identified in the haloarchaea (*Euryarchaeota* phylum) (30). The ETNP AioA-like sequences (20, 22) grouped more closely with the bacterial subclade than with the haloarchaeal group. It has been suggested that the bacterial *aioA-like* sequences arose via horizontal gene transfer from a Euryarchaeotal ancestor, a distinct event from the origin of *aioA* where the phylogenetic radiation is more congruent with the 16S rRNA phylogeny (30). An abundance of archaeal *aioA-like* sequences was identified in an haloarchaeal red biofilm in the arsenic-loaded Diamante Lake, Argentina and a culture of *Halorubrum* sp. DM2 isolated from the same location demonstrated aerobic growth on a minimal medium enriched with As(III) (31). However, the exact substrate of the AioA-Like enzyme has yet to be verified. Bacterial AioA-Like may indeed be a functional homolog to the arsenite oxidase subunit AioA. However, without further laboratory confirmation of the explicit function of this enzyme, we cannot definitively say whether this Aio-Like enzyme functions as an As(III) oxidase or whether its function has taken another evolutionary trajectory.

The longest contig containing *aioA-like* sequences was over 82 kilobases and provided additional insight into the metabolic capabilities of the host Proteobacterium (SI Appendix, Table S1). Assembled from a single metagenomic sample, contig ETNP_120m_NODE_73101 contains a nitrate reductase sequence (*napA*) which suggests the use of nitrate as an electron acceptor, as well as a gene encoding fumarate reductase, a key enzyme in the reductive TCA carbon fixation pathway (32), suggesting the potential for a chemoautotrophic lifestyle. The cooccurrence of these genes suggests that the Aio-Like enzyme may help support autotrophic carbon fixation and nitrate reduction in the ETNP ODZ.

To identify whether or not these arsenotrophy-related genes are found in other regions of the global ocean, we searched publicly available marine metagenome assemblies deposited in the National Center for Biotechnology Information (NCBI) Whole Genome Shotgun (WGS) repository (Fig. 1A and SI Appendix, Fig. S6 and Table S2). Genes for AioA, AioA-Like, and ArrA were all found on contigs assembled from oxygen-depleted waters of the Baltic Sea (33) and the *aioA-like* gene was also found in assemblies from the Arabian Sea ODZ. Notably, dissimilatory arsenate reductase, *arrA*, was only identified in assemblies from lower oxygen regions of the ocean, whereas *aioA* and *aioA-like* were also present in waters with higher oxygen concentrations such as the Eastern Tropical Pacific (*aioA-like*) and the oxygenated Mediterranean (*aioA*) (SI Appendix, Fig. S6 and Table S2) (33–35). The analysis of WGS assemblies shows the presence of *aioA* and *aioA-like* in genomes and thus the potential for microbial arsenite oxidation is not restricted only to oxygen deficient waters. The presence of these sequences in oxygenated waters may be the signature of aerobic chemoautotrophic arsenite oxidation (36), or may be associated with arsenic detoxification processes in regions of phosphate stress (8, 37).

Niche Partitioning of Arsenotrophy and Related Genes. To determine the prevalence of arsenotrophic capabilities in microbial communities in the ETNP, we used a phylogenetically informed short-read placement approach to evaluate the frequency of metagenomic sequence reads associated with the genes *aioA*,

aioA-like, and *arrA* compared with reads associated with the single-copy core gene RNA polymerase (*rpoB*) in free-living (<30 μm) or particle-associated (>30 μm) microbial communities at 120 m (Fig. 4). For comparison with a well-studied analogous metabolic cycle in ODZs, we also identified and enumerated short reads associated with the bacterial form of the sulfur-cycling gene *dsrA* (SI Appendix, Fig. S7), the forward form used for sulfur reduction and the reverse used for sulfur oxidation (38). The genes *arrA*, *aioA*, and forward bacterial *dsrA* of sulfate-reducing bacteria (SRB) were possessed by a small portion (<2%) of the microbial community, while the genes *aioA-like* and reverse bacterial *dsrA* associated with sulfur-oxidizing bacteria (SOB) were possessed by a larger fraction of the community (4–11%). Furthermore, the genes *arrA*, *aioA-like*, and forward bacterial *dsrA* (SRB) were more abundant in the particle-associated communities compared with the free-living fraction. It makes sense to find *arrA* and *dsrA* (SRB) inhabiting similar functional environments as they are likely both exploiting similar redox-driven niches in particles, coupling the heterotrophic breakdown of organic matter with the reduction of arsenate or oxidized sulfur compounds in the absence of more energetically favorable electron acceptors, like nitrate, which are available to the free-living microbial community (3).

Active Gene Expression of Arsenotrophy and Related Genes. We analyzed publicly available metatranscriptomes in global ODZs (39–41) to identify whether arsenotrophic and arsenotrophy-related genes are transcriptionally active in these environments. Again using phylogenetically informed short read placement, we identified transcripts associated with *arrA*, *aioA*, and *aioA-like* from metatranscriptomic libraries in the ETNP (39) and two areas of the Eastern Tropical South Pacific (ETSP) (40, 41) ODZs—one ETSP library focused on the smallest free-living microbial community (0.22–1.6 μm) of a typical ODZ water

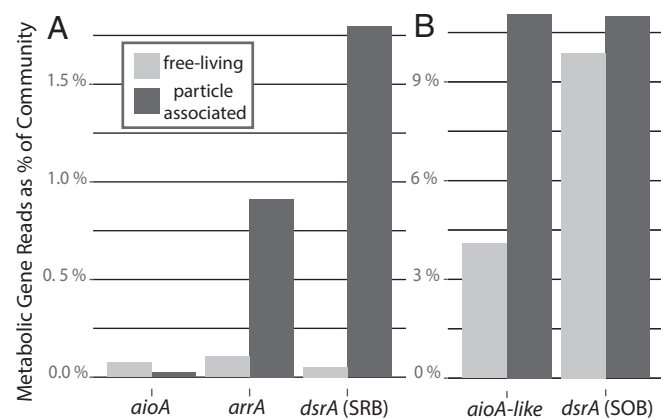


Fig. 4. Arsenic respiratory genes and other related respiratory genes show niche partitioning in genomic contribution to microbial communities among free-living (<30 μm) or particle-associated niches (>30 μm) at 120-m depth in ETNP ODZ metagenomes. Reads identified using phylogenetically informed short-read placement. Contribution of reads presented as the length-normalized number of reads associated with each target gene/the number of length-normalized RNA polymerase B, *rpoB*, reads in the metagenomes * 100% to give an estimate of the gene's % contribution to the overall microbial community metagenome (53). The relative proportion of a gene's contribution to the microbial community when the free-living and particulate fractions are compared broken up into A with a much smaller proportion of the community containing arsenotrophy genes arsenite oxidase, *aioA*, and dissimilatory arsenate reductase, *arrA*, as well as the forward version of dissimilatory sulfite reductase, *dsrA*, which is involved in SRB. B shows a larger proportion of the community possesses the sequences *aioA-like* and the reverse version of *dsrA* which is involved in SOB.

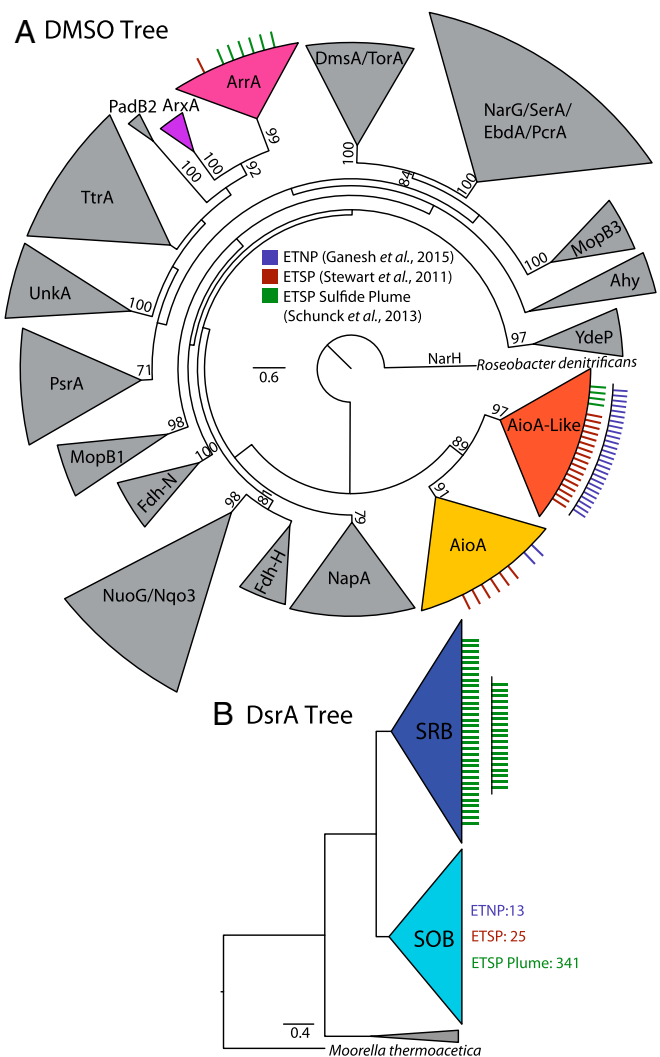


Fig. 5. (A) A collapsed view of the DMSO reductase family tree (SI Appendix, Fig. S4), with associated short reads from metatranscriptomes (34–36) identified as associated with arsenotrophic enzymes through phylogenetically informed placement represented by tick marks mapped onto their respective clades; reads associated with other enzymes are not shown. (B) A phylogenetic tree of *DsrA* showing transcripts associated with forward bacterial *DsrA* associated with dissimilatory SRB and the reverse *DsrA* associated with SOB mapped along the respective clades for comparison with the transcripts associated with arsenotrophy in these ODZ regions captured by the publicly available transcriptomes from ODZs (34–36). Specific read identifiers mapped to these clades can be found in SI Appendix, Table S3.

column (41) and the other captured the free-living and small particle associated microbial community (0.22–10 μm) within a transient sulfide plume feature (40). Again, for comparison, we also identified transcripts associated with bacterial *dsrA*. Of the arsenotrophic genes, transcripts associated with *aioA-like* were by far the most abundant (Fig. 5A) and found in both the ETNP and the ETSP ODZs. Transcript sequences associated with *aioA* were also found in both the ETNP and ETSP. Dissimilatory arsenate reductase *arrA* and forward bacterial *dsrA* (SRB) transcripts were only detected in the ETSP, with the majority of sequences identified from within the sulfidic plume (Fig. 5B) (40). Thus, arsenotrophy-related genes are being expressed at a transcriptional level comparable to the forward bacterial *dsrA* (SRB), suggesting the active use of these metabolic pathways. The gene of unknown function *aioA-like* was transcriptionally active on the same order as the sulfur-oxidizing form of *dsrA*

(SOB) in the nonsulfidic ETNP and ETSP metatranscriptomes, again suggesting that this as yet to be determined Aio-Like pathway may be of similar significance in these waters.

Conclusion and Implications. The presence of complete arsenotrophic gene pathways in the metagenome of the ETNP ODZ suggests microbial communities in these anoxic waters are capable of a complete bioenergetic arsenotrophic cycle. Gene expression of the arsenotrophy genes *arrA* and *aioA* indicates arsenotrophy is a viable metabolic pathway that is actively utilized in anoxic water columns. Chemoautotrophic arsenite oxidation may provide an additional fixed carbon source, contributing to the biochemical processes driving the reducing environment of ODZs. Nitrogen loss through the reduction of nitrogenous compounds to N_2 has been demonstrated by some chemoautotrophic arsenite oxidizers (29); if arsenite oxidizers in the ETNP are also capable of denitrification they would impact nitrogen cycling in ODZs. The enrichment of both arsenic and sulfate reduction genes in the particulate fraction identified here is an example of redox-driven niche partitioning on sinking particles, which have zones with more reducing potential than found in the open water column (3). The cryptic arsenic cycle described here is an account of arsenotrophic cycling in an environment with nanomolar levels of available arsenic (8) suggesting these arsenic respiratory pathways may be more abundant in the environment than previously estimated.

The identification of arsenotrophic microbes living in modern marine ODZs will enable proxies for arsenic-based metabolisms in early anoxic oceans. Marine arsenic concentrations have fluctuated over geologic time, with the anoxic Precambrian ocean experiencing periods of greater arsenic load than the modern-day ocean (13) due in part to enhanced weathering of continental arsenic sulfide minerals (12). Arsenotrophic organisms may have flourished during periods of higher arsenic concentrations in an anoxic ocean, with chemoautotrophic or possibly photosynthetic arsenite oxidation supporting carbon fixation. Thus, integrating over the time and spatial scale of early anoxic oceans, arsenic-based metabolisms may have had significant implications for the biogeochemical cycling of arsenic, carbon, and nitrogen. The work described here presents a foundation to explore these metabolic pathways and their implications for biogeochemical cycling and microbial evolution in the modern-day ocean and oceans of the past.

Methods

Samples were collected during a research cruise to the ETNP in April 2012 aboard the research vessel (RV) *Thompson* (cruise TN278) using 10-L Niskin bottles in a 24-bottle conductivity, temperature, and depth (CTD) water sampling rosette. The top of the ODZ was determined to be at 105 m based upon oxygen concentrations determined by switchable trace oxygen (STOX) sensor (5). Oxygen and nutrient measurements were collected as previously described (20, 42).

DNA samples were obtained from station 136 (106.543° W 17.043° N) in the anoxic zone at multiple depths (100, 110, 120, 160, 180 m). Two liters of Niskin water was vacuum filtered onto a 0.2- μ m SUPOR filter. At the off-shore multiday station BB2/141 (107.148° W 16.527° N), approximately four liters of water from 120 m were first filtered through a >30- μ m nylon filter and the prefiltered (<30 μ m) water was then filtered onto 0.2- μ m Supor membrane filters (PALL Corporation). Filters were immediately frozen in a -80 °C freezer and transported on dry ice from the ship to the University of Washington where they were maintained at -80 °C.

DNA library preparation and sequencing with Illumina HiSeq 2500 were conducted in 2014 and previously described in another publication (20). DNA samples were extracted using freeze thaw followed by incubation with lysozyme and proteinase K and phenol/chloroform extraction. A Rubicon THRUPLEX kit was used for library prep using 50 ng of DNA per sample. Two libraries (ETNP_100m & ETNP_110m) were sequenced on a HiSeq 2500 in rapid mode (~22 million 150-bp paired-end read pairs per sample) at Michigan State University. The other five libraries (ETNP_120m, ETNP_160m, ETNP_180m, ETNP_BB2_120m_free, and ETNP_BB2_120m_particle) were sequenced on a HiSeq 2500 high-output mode (~55–150 million 125-bp paired-end read pairs per sample) at the University of Utah. Sequences were quality checked and trimmed using Trimmomatic (43).

Metagenomic reads from each depth sample were first assembled separately, processed with *diginorm* (44) to normalize coverage, and then de novo assembled with the *VELVET* assembler (45). Contigs assembled from individual depths as previously described (20) were functionally annotated using the *prokka* annotation pipeline (46) with reads and assembled contigs deposited at NCBI GenBank under bioproject PRJNA350692 v.1 assemblies (*SI Appendix, Table S4* lists accessions for each assembly). *Prokka* alone was insufficient to identify DMSO reductase family-related sequences (*aioA*, *aioA-like*, *arrA*, and *arrX*). To identify these arsenotrophy-related enzymes, blast databases of all assembled contigs were created using *blast* version 2.2.28 (47, 48) and queried with *tblastn* using known reference queries. Contigs which had hits to arsenotrophy-related sequences with e values $\leq 1 \times 10^{-50}$ were parsed out, and submitted to *MetaGeneMark* for gene prediction (49, 50). Potential genes that overlapped with the region of best blast hit were then added to the DMSO reductase tree for further identification, with full-length sequences that grouped with arsenotrophy-related clades depicted on the DMSO reductase tree (*SI Appendix, Fig. S4*).

To assemble the longest possible arsenotrophy-related contigs, a targeted assembly approach was used to combine metagenomes from all anoxic samples. Partial and full arsenotrophy-related sequences identified from assembled contigs from individual depths and reference regions from known arsenotrophy-related sequences (gene sequence with 10-kb regions up- and downstream) were used as queries for sequence read recruitment from each individual depth metagenome using the program *FR-HIT* (51). Reads recruited with *FR-HIT* were then assembled with the de novo assembler *IDBA-UD* (52) optimized for assembling data of uneven depth coverage. Arsenotrophy-related genes from these contigs were identified as above. In addition, all genes called on contigs identified as containing full-length arsenotrophy-related enzymes were annotated with *InterProScan* (53) with general taxonomic identity on a gene-by-gene basis being inferred at the bacterial class level through a *blastn* best hit to the *nr/nt* database. The assembled *arrA* amino acid sequence was then pairwise aligned with that from *Shewanella ANA-3* using *Jalview 2* (24, 54).

Gene trees were constructed with amino acid translated sequences from metagenomic assemblies and known reference sequences aligned with the program *MUSCLE* version 3.6 (55). Phylogenetic trees were then constructed with the program *RAXML* version 8.0.23 (56). A maximum-likelihood tree was inferred from the best of 20 trees using the amino acid substitution matrix model which resulted in the best likelihood score during a trial run (DMSO: BLOSUM62F; DsrA, AioB, and ArrB: WAGF) along with empirical character frequencies and a gamma model of rate heterogeneity using the *RAXML* estimated alpha value. Bootstrap analyses were conducted on all trees at $n = 100$. The DMSO reductase family tree and *DsrA* tree were used as reference trees for phylogenetic placement of metagenomic and metatranscriptomic reads.

Publicly available pelagic marine metagenomic assemblies in the NCBI WGS database were searched for the presence of *arrA*, *aioA*, and *aioA-like*. Multiple amino acid query sequences spanning the taxonomic range of these genes were used to identify potential genes in WGS using the *SRA* toolkit tool *tblastn_vdb* to remotely blast search assembled metagenomes annotated as marine metagenomes from seawater. Hits receiving an e value $\leq 1 \times 10^{-80}$ were selected and their contig sequences obtained from WGS using *Batch Entrez*. Contigs were then submitted to *MetaGeneMark* for gene prediction (49, 50). Potential full-length arsenotrophy genes that overlapped with the region of best blast hit were added to the DMSO reductase tree for further identification.

Recruitment and placement of metagenomic reads from the samples at 120-m depth from station BB2/141 comparing the >30- μ m (particulate) and <30- μ m (free-living) fractions was conducted. Reads were combined into a single blast database using *blast* 2.2.28 (47, 48) and queried with *tblastn* (e -value cutoff 1×10^{-5}) using full-length known reference sequences, including full-length identified environmental arsenotrophy sequences assembled in this work. Translation of blast recruited reads which were at least 33 amino acids in length after quality trimming were then identified using a phylogenetically informed placement approach by comparison with known reference sequences (20, 37, 57–59). To obtain the abundance of these reads relative to the overall prokaryotic community in the two size fractions, the length-normalized number of identified read pairs for each gene was compared with the length-normalized abundance of the single-copy core gene RNA polymerase (*rpoB*) in the sequenced prokaryotic community (20, 37).

A similar read recruitment and phylogenetic placement method was used to identify reads from ODZ metatranscriptomes (37, 57–59). Publicly available metatranscriptomes from oxygen deficient zones in the ETNP (39), ETSP (41), and within a sulfide plume in the ETSP (40) were downloaded locally. Reads were also recruited for identification of bacterial *dsrA* for comparison.

ACKNOWLEDGMENTS. We thank John Baross and Rika Anderson for helpful discussions and feedback on this project. We also thank the chief scientists of the research cruise, Al Devol and Bess Ward, as well as the captain and crew

of the *R/V Thomas G. Thompson*. This work was supported through a NASA Earth and Space Sciences Graduate Research Fellowship to J.K.S. and National Science Foundation Grant OCE-1138368 (to G.R.).

- Lyons TW, Reinhard CT, Planavsky NJ (2014) The rise of oxygen in Earth's early ocean and atmosphere. *Nature* 506:307–315.
- Ulloa O, Canfield DE, DeLong EF, Letelier RM, Stewart FJ (2012) Microbial oceanography of anoxic oxygen minimum zones. *Proc Natl Acad Sci USA* 109:15996–16003.
- Wright JJ, Konwar KM, Hallam SJ (2012) Microbial ecology of expanding oxygen minimum zones. *Nat Rev Microbiol* 10:381–394.
- Wyrki K (1962) The oxygen minima in relation to ocean circulation. *Deep-Sea Res Oceanogr Abstr* 9:11–23.
- Tiano L, et al. (2014) Oxygen distribution and aerobic respiration in the north and south eastern tropical Pacific oxygen minimum zones. *Deep Sea Res Part I Oceanogr Res Pap* 94:173–183.
- Canfield DE, et al. (2010) A cryptic sulfur cycle in oxygen-minimum-zone waters off the Chilean coast. *Science* 330:1375–1378.
- Oremland RS, Stolz JF (2003) The ecology of arsenic. *Science* 300:939–944.
- Amend JP, Saltikov C, Lu G-S, Hernandez J (2014) Microbial arsenic metabolism and reaction energetics. *Rev Mineral Geochem* 79:391–433.
- Duval S, Ducluzeau AL, Nitschke W, Schoepp-Cothenet B (2008) Enzyme phylogenies as markers for the oxidation state of the environment: The case of respiratory arsenate reductase and related enzymes. *BMC Evol Biol* 8:206.
- Lebrun E, et al. (2003) Arsenite oxidase, an ancient bioenergetic enzyme. *Mol Biol Evol* 20:686–693.
- Sforna MC, et al. (2014) Evidence for arsenic metabolism and cycling by microorganisms 2.7 billion years ago. *Nat Geosci* 7:811–815.
- Fru EC, et al. (2015) Arsenic stress after the Proterozoic glaciations. *Sci Rep* 5:17789.
- Bergman IA, Kolesov GM (2012) Arsenic, antimony, and bismuth as indicators of the genesis of ore material in Early Precambrian ferrous quartzite formations. *Geochem Int* 50:816–831.
- Ellwood MJ, Maher WA (2002) An automated hydride generation-cryogenic trapping-ICP-MS system for measuring inorganic and methylated Ge, Sb and As species in marine and fresh waters. *J Anal At Spectrom* 17:197–203.
- Cutter GA, Cutter LS (2006) Biogeochemistry of arsenic and antimony in the North Pacific Ocean. *Geochem Geophys Geosyst* 7:Q05M08.
- Cutter GA, Moffett JG, Nielsdóttir MC, Sanial V (2018) Multiple oxidation state trace elements in suboxic water off Peru: In situ redox processes and advective/diffusive horizontal transport. *Mar Chem* 201:77–89.
- Cutter GA (1992) Kinetic controls on metalloid speciation in seawater. *Mar Chem* 40: 65–80.
- Peterson ML, Carpenter R (1983) Biogeochemical process affecting total arsenic and arsenic species distribution in an intermittently anoxic fjord. *Mar Chem* 12:295–321.
- Garcia H, et al. (2014) World Ocean Atlas 2013, volume 4: Dissolved inorganic nutrients (phosphate, nitrate, silicate). *NOAA Atlas NESDIS* 76:25.
- Fuchsman CA, Devol AH, Saunders JK, McKay C, Rocap G (2017) Niche partitioning of the N cycling microbial community of an offshore oxygen deficient zone. *Front Microbiol* 8:2384.
- Rothery RA, Workun GJ, Weiner JH (2008) The prokaryotic complex iron-sulfur molybdoenzyme family. *Biochim Biophys Acta* 1778:1897–1929.
- Saunders JK, Fuchsman CA, McKay C, Rocap G (2016) Arsenic respiratory pathways in the anoxic pelagic waters of the Pacific Ocean. GenBank. Available at <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA350692>. Deposited December 22, 2016.
- Andres J, Bertin PN (2016) The microbial genomics of arsenic. *FEMS Microbiol Rev* 40: 299–322.
- Glasser NR, Oyala PH, Osborne TH, Santini JM, Newman DK (2018) Structural and mechanistic analysis of the arsenate respiratory reductase provides insight into environmental arsenic transformations. *Proc Natl Acad Sci USA* 115:E8614–E8623.
- Richey C, et al. (2009) Respiratory arsenate reductase as a bidirectional enzyme. *Biochem Biophys Res Commun* 382:298–302.
- van Lis R, Nitschke W, Duval S, Schoepp-Cothenet B (2013) Arsenics as bioenergetic substrates. *Biochim Biophys Acta* 1827:176–188.
- Oremland RS, et al. (2002) Anaerobic oxidation of arsenite in Mono Lake water and by a facultative, arsenite-oxidizing chemoautotroph, strain MLHE-1. *Appl Environ Microbiol* 68:4795–4802.
- Rhine ED, Ni Chadhain SM, Zylstra GJ, Young LY (2007) The arsenite oxidase genes (*aroAB*) in novel chemoautotrophic arsenite oxidizers. *Biochem Biophys Res Commun* 354:662–667.
- Rhine ED, Phelps CD, Young LY (2006) Anaerobic arsenite oxidation by novel denitrifying isolates. *Environ Microbiol* 8:899–908.
- Rascovan N, Maldonado J, Vazquez MP, Eugenia Farias M (2016) Metagenomic study of red biofilms from Diamante Lake reveals ancient arsenic bioenergetics in haloarchaea. *ISME J* 10:299–309.
- Ordoñez OF, Rasuk MC, Soria MN, Contreras M, Farias ME (2018) Haloarchaea from the Andean Puna: Biological role in the energy metabolism of arsenic. *Microb Ecol* 76: 695–705.
- Hügler M, Sievert SM (2011) Beyond the Calvin cycle: Autotrophic carbon fixation in the ocean. *Annu Rev Mar Sci* 3:261–289.
- Dupont CL, et al. (2014) Functional tradeoffs underpin salinity-driven divergence in microbial community composition. *PLoS One* 9:e89549.
- Zeigler Allen L, et al. (2012) Influence of nutrients and currents on the genomic composition of microbes across an upwelling mosaic. *ISME J* 6:1403–1414.
- Sunagawa S, et al.; Tara Oceans Coordinators (2015) Ocean plankton. Structure and function of the global ocean microbiome. *Science* 348:1261359.
- Santini JM, Sly LI, Schnagl RD, Macy JM (2000) A new chemolithoautotrophic arsenite-oxidizing bacterium isolated from a gold mine: Phylogenetic, physiological, and preliminary biochemical studies. *Appl Environ Microbiol* 66:92–97.
- Saunders JK, Rocap G (2016) Genomic potential for arsenic efflux and methylation varies among global *Prochlorococcus* populations. *ISME J* 10:197–209.
- Müller AL, Kjeldsen KU, Rattei T, Pester M, Loy A (2015) Phylogenetic and environmental diversity of DsrAB-type dissimilatory (bi)sulfite reductases. *ISME J* 9: 1152–1165.
- Ganesh S, et al. (2015) Size-fraction partitioning of community gene transcription and nitrogen metabolism in a marine oxygen minimum zone. *ISME J* 9:2682–2696.
- Schunck H, et al. (2013) Giant hydrogen sulfide plume in the oxygen minimum zone off Peru supports chemolithoautotrophy. *PLoS One* 8:e68661.
- Stewart FJ, Ulloa O, DeLong EF (2012) Microbial metatranscriptomics in a permanent marine oxygen minimum zone. *Environ Microbiol* 14:23–40.
- Devol AH (2013) Physical, chemical and biological CTD and bottle data from R/V Thomas G. Thompson cruise TN278 in eastern tropical North Pacific Ocean (National Centers for Environmental Information, National Oceanographic and Atmospheric Administration, Asheville, NC).
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120.
- Brown CT, Howe A, Zhang Q, Pyrkosz AB, Brom TH (2012) A reference-free algorithm for computational normalization of shotgun sequencing data. arXiv:1203.4802 v2. Preprint, posted May 21, 2012.
- Zerbino DR, Birney E (2008) Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res* 18:821–829.
- Seemann T (2014) Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* 30: 2068–2069.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410.
- Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.
- Besemer J, Borodovsky M (1999) Heuristic approach to deriving models for gene finding. *Nucleic Acids Res* 27:3911–3920.
- Zhu W, Lomsadze A, Borodovsky M (2010) *Ab initio* gene identification in metagenomic sequences. *Nucleic Acids Res* 38:e132.
- 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.
- Peng Y, Leung HC, Yiu SM, Chin FY (2012) IDBA-UD: A *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28: 1420–1428.
- Jones P, et al. (2014) InterProScan 5: Genome-scale protein function classification. *Bioinformatics* 30:1236–1240.
- Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ (2009) Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25: 1189–1191.
- Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
- Stamatakis A (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.
- Berger SA, Stamatakis A (2011) Aligning short reads to reference alignments and trees. *Bioinformatics* 27:2068–2075.
- Berger SA, Krompass D, Stamatakis A (2011) Performance, accuracy, and Web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst Biol* 60:291–302.
- Berger SA, Stamatakis A (2012) PaPaRa 2.0: A vectorized algorithm for probabilistic phylogeny-aware alignment extension. *Heidelberg Institute for Theoretical Studies*. Available at <https://cme.h-its.org/exelixis/pubs/Exelixis-RRDR-2012-5.pdf>. Accessed August 31, 2016.