



# Metagenome-assembled genomes reveal unique metabolic adaptations of a basal marine Thaumarchaeota lineage

Linta Reji<sup>1</sup> · Christopher A. Francis<sup>1</sup> <sup>1</sup>

Received: 21 December 2019 / Revised: 29 April 2020 / Accepted: 29 April 2020 / Published online: 13 May 2020  
© The Author(s), under exclusive licence to International Society for Microbial Ecology 2020

## Abstract

Thaumarchaeota constitute an abundant and ubiquitous phylum of Archaea that play critical roles in the global nitrogen and carbon cycles. Most well-characterized members of the phylum are chemolithoautotrophic ammonia-oxidizing archaea (AOA), which comprise up to 5 and 20% of the total single-celled life in soil and marine systems, respectively. Using two high-quality metagenome-assembled genomes (MAGs), here we describe a divergent marine thaumarchaeal clade that is devoid of the ammonia-oxidation machinery and the AOA-specific carbon-fixation pathway. Phylogenomic analyses placed these genomes within the uncultivated and largely understudied marine pSL12-like thaumarchaeal clade. The predominant mode of nutrient acquisition appears to be aerobic heterotrophy, evidenced by the presence of respiratory complexes and various organic carbon degradation pathways. Both genomes encoded several pyrroloquinoline quinone (PQQ)-dependent alcohol dehydrogenases, as well as a form III RuBisCO. Metabolic reconstructions suggest anaplerotic CO<sub>2</sub> assimilation mediated by RuBisCO, which may be linked to the central carbon metabolism. We conclude that these genomes represent a hitherto unrecognized evolutionary link between predominantly anaerobic basal thaumarchaeal lineages and mesophilic marine AOA, with important implications for diversification within the phylum Thaumarchaeota.

## Introduction

Archaea of the phylum Thaumarchaeota are among the most abundant microorganisms on the planet, constituting up to 20% of single-celled life in marine systems alone [1]. Although most characterized members of Thaumarchaeota are ammonia-oxidizing archaea (AOA), the phylum also encompasses several archaeal clades for which ammonia oxidation has not yet been demonstrated (e.g., Group 1.1c, and Group 1.3 [2]). These basal, non-AOA members of the phylum have primarily been described in terrestrial systems such as anoxic peat soils [3], subsurface aquifer sediments [4], geothermal springs [5, 6], and acidic forest soil [7]. Availability of molecular oxygen on Earth is hypothesized

to have influenced the evolution and habitat expansion of AOA from the basal anaerobic guilds [8].

A deeply branching marine thaumarchaeal clade that has eluded cultivation and genomic analysis efforts is the pSL12-like group, also referred to as Group 1A or ALOHA group. First detected by DeLong et al. [9] in the North Pacific Subtropical Gyre at station ALOHA, this clade appeared to be divergent from Marine Group 1 Archaea, clustering with a hot spring-associated crenarchaeal 16S rRNA sequence designated pSL12 [10]. Mincer et al. [11] suggested that at least some members of the clade may harbor the ammonia-oxidation machinery, based on correlating abundances of the 16S rRNA gene and the *amoA* gene in oceanic water column samples (*amoA* encodes the alpha-subunit of ammonia monooxygenase (AMO); conventionally used as the functional marker for AOA). The only available genomic information for the pSL12-like lineage comes from a fosmid clone library generated from the Mediterranean Sea [12]. One of the three pSL12-like fosmid sequences recovered by Martin-Cuadrado et al. [12] contained genes putatively involved in nitrogen fixation; however, there has been no genomic or biogeochemical evidence supporting this observation since. Several SSU rRNA gene surveys have detected the pSL12-like group in

---

**Supplementary information** The online version of this article (<https://doi.org/10.1038/s41396-020-0675-6>) contains supplementary material, which is available to authorized users.

✉ Christopher A. Francis  
caf@stanford.edu

<sup>1</sup> Earth System Science, Stanford University, Stanford, CA, USA

various marine systems such as the Atlantic Ocean [13], Mediterranean Sea [14], multiple Pacific Ocean transects [15], the Northern Gulf of Mexico [16], and Monterey Bay [17]. Despite their suggested roles in N-cycle transformations, the metabolic adaptations of the pSL12-like lineage remain an open question.

Here we analyze the genomic repertoire and metabolic strategies of the pSL12-like lineage, based on two metagenome-assembled genomes (MAGs) obtained from seawater incubation metagenomes derived from Monterey Bay. Metabolic reconstructions point to a heterotrophic lifestyle. Intriguingly, both genomes also encoded a form III ribulose-bisphosphate carboxylase (RuBisCO), which may participate in a CO<sub>2</sub> incorporation pathway linked to nucleoside salvage reactions. The high degree of phylogenetic and metabolic separation between these MAGs and typical marine thaumarchaeal clades suggests that the pSL12-like lineage represents an evolutionary link between anaerobic basal clades of Thaumarchaeota and aerobic marine AOA.

## Materials and methods

### Sample collection, incubation, and DNA extraction

Water column samples for AOA enrichment incubations were collected from Monterey Bay, CA, in May 2010. ASW2 was collected from 150 m at station M1 (36.747 N, -122.022 W), and ASW8 was collected from 200 m at station M2 further offshore (36.697 N, -122.378 W). After 8 years of incubation at 12°C (seawater samples were unamended; the long incubation period was to facilitate natural enrichment of AOA), 925 and 1000 mL each of the samples (for ASW2 and ASW8, respectively) were filtered using a 0.22-µm filter (Supor, Pall Inc, New York, USA). DNA was extracted using the DNeasy kit (Qiagen, Valencia, CA, USA), following the manufacturer's protocol. To maximize DNA yield, DNeasy capture columns were eluted twice with 50 µL each of elution buffer, resulting in 100 µL total elution volume for each sample. DNA concentration was measured using Qubit Fluorometer (Invitrogen, NY, USA); 1.41 and 1.88 µg/mL DNA was obtained from ASW2 and ASW8, respectively.

### Metagenome sequencing, assembly, and binning

Metagenome sequencing was performed as a part of a Community Science Program (CSP) project with the DOE Joint Genome Institute (JGI); the samples were sequenced (2 × 151 bp) using the HiSeq 2000 1TB platform. Read quality-filtering was carried out using the custom JGI script `jgi_mga_meta_rqc.py` (v2.0.0). Briefly, trimmed paired-end

reads filtered using BBDuk [18] (v37.50; BBTools software package, <http://bbtools.jgi.doe.gov>) were read-corrected using BFC (v.r181 [19]). Reads without a mate pair were removed.

Quality-filtered reads were assembled using MEGAHIT (v1.1.3 [20, 21]), using a range of k-mers ( $k = 21, 33, 55, 77, 99, 127$ ). Contigs longer than 2000 bp were binned using two algorithms: MetaBAT2 (v2.12.1 [22]) and MaxBin2 (v2.2.6 [23, 24]). Resulting bins were refined using the bin refinement module in metaWRAP (v1.2.2 [25]), and subsequently re-assembled using SPAdes (v3.13.0 [26]) to improve assembly quality. CheckM (v1.0.12 [27]) was used to assess bin completion. Taxonomic classifications were obtained using the GTDB-tk toolkit (v0.3.2 [28]). Dereplication based on average nucleotide identity (ANI) was performed using dRep (v2.3.2 [29]). Only bins with estimated completeness  $\geq 70\%$  and contamination  $< 10\%$  were retained for downstream analysis.

The assembled genome sequences can be accessed under the BioSample IDs SAMN14765629 and SAMN14765628, respectively, for ASW2\_bin45 and ASW8\_bin1 (corresponding BioProject accessions are PRJNA621967 and PRJNA539366, respectively).

### MAG annotation and metabolic reconstruction

Prodigal (v2.6.3 [30]) was used for gene prediction, and functional annotations were obtained using Prokka (v1.12 [31]). In addition, the BlastKOALA and GhostKOALA tool servers [32] were used to obtain KO annotations for genes predicted by Prodigal. KEGG-decoder [33] was used to estimate pathway completeness based on KO annotations, and the results were plotted in R [34]. SEED annotations were obtained from the online Rapid Annotation using Subsystem Technology server [35]. Metabolic reconstructions were carried out using the 'Reconstruct Pathway' tool in KEGG mapper (<https://www.genome.jp/kegg/mapper.html>). TransportDB (v2.0 [36]) was used to predict membrane transporters; these annotations were further confirmed by BLASTp searches against the NCBI nonredundant protein database. SignalP-5.0 Server was used for signal peptide prediction (<http://www.cbs.dtu.dk/services/SignalP-5.0/>).

### Phylogenetic analyses

Reference genomes for Thaumarchaeota and Aigarchaeota were downloaded from NCBI or the Integrated Microbial Genomes system. The phylogenomics module in Anvi'o (v5.4 [37]) was used to retrieve ribosomal sequences from the MAGs and the reference genomes. The 'anvi-get-sequences-for-hmm-hits' command was used to search for and retrieve 30 ribosomal proteins from each genome (these

included ribosomal proteins L1, L10, L11, L11\_N, L13, L14, L16, L18p, L2, L22, L23, L29, L2\_C, L3, L4, L5, L5\_C, L6, S11, S13, S15, S17, S19, S2, S3\_C, S5, S5\_C, S7, S8, and S9). Amino acid sequences for the retrieved proteins were aligned using MUSCLE [38] and concatenated. The alignment was trimmed using trimal [39], with the parameters: `-gapthreshold 0.75 -simthreshold 0.001`. Further manual refinement was carried out in Geneious (v10.2; Biomatters Ltd, New Zealand). Since some of the genomes included in the analysis were assembled from metagenome or single-cell genome data, not all ribosomal proteins were universally identified across genomes. In the final alignment, only those genes identified in all genomes were retained, and this amounted to a total of 11 genes across 23 genomes. A maximum-likelihood tree was computed using FastTree [40] with 100 bootstrap replicates.

We used BLASTp [41] to search the MAGs for proteins of interest—both to confirm automatic annotations and to search for specific pathways/genes. Barnap (v0.9; <https://github.com/tseemann/barnap>) was used to identify ribosomal features. 16S rRNA sequences were aligned with reference sequences using MAFFT [42], and a maximum-likelihood phylogenetic tree was computed in FastTree [40] with 1000 bootstrap replicates. RuBisCO reference sequences were obtained from Jaffe et al. [43]; MAFFT and FastTree, respectively, were used for generating an alignment and a phylogenetic tree.

FastANI [44] was used to compute ANI between the MAGs. GTDB-tk identified a moderate-quality (62% estimated completeness) MAG assembled from a deep hydrothermal plume [45] as a close relative of the MAGs assembled here; this genome (UBA57) was also included in FastANI and function comparison analyses.

### Assessing environmental distribution of MAGs

As part of the time-series microbiome survey in Monterey Bay, we previously obtained a depth-resolved dataset of 16S rRNA V4-V5 amplicon sequences [46], as well as metagenomes and metatranscriptomes [47]. We were able to match one of the MAG-derived 16S rRNA sequences to an operational taxonomic unit (OTU) obtained in the 16S rRNA time-series dataset. We estimated the relative abundance of this OTU as well as another that shared 96–97% sequence identity with the MAG-derived sequences.

We used three metagenome sets for read recruitment: (i) a depth- and time-resolved metagenome dataset from Monterey Bay; (ii) a North Atlantic Ocean depth profile from the TARA Oceans dataset; and (iii) a North Pacific Ocean depth profile from the TARA Oceans dataset. Note that the TARA oceans datasets do not represent a continuous depth profile (Table S1). Bowtie2 (v2.3.5 [48]) was used to recruit metagenomic and metatranscriptomic

reads against the MAGs. Read abundances were normalized as the number of reads recruited per kilobase of MAG and gigabase of metagenome (RPKG). The RPKG values allowed the direct comparison of genome abundances (measured as coverage) between metagenomes of different sizes.

## Results and discussion

### Genomes recovered from reduced-diversity metagenomes

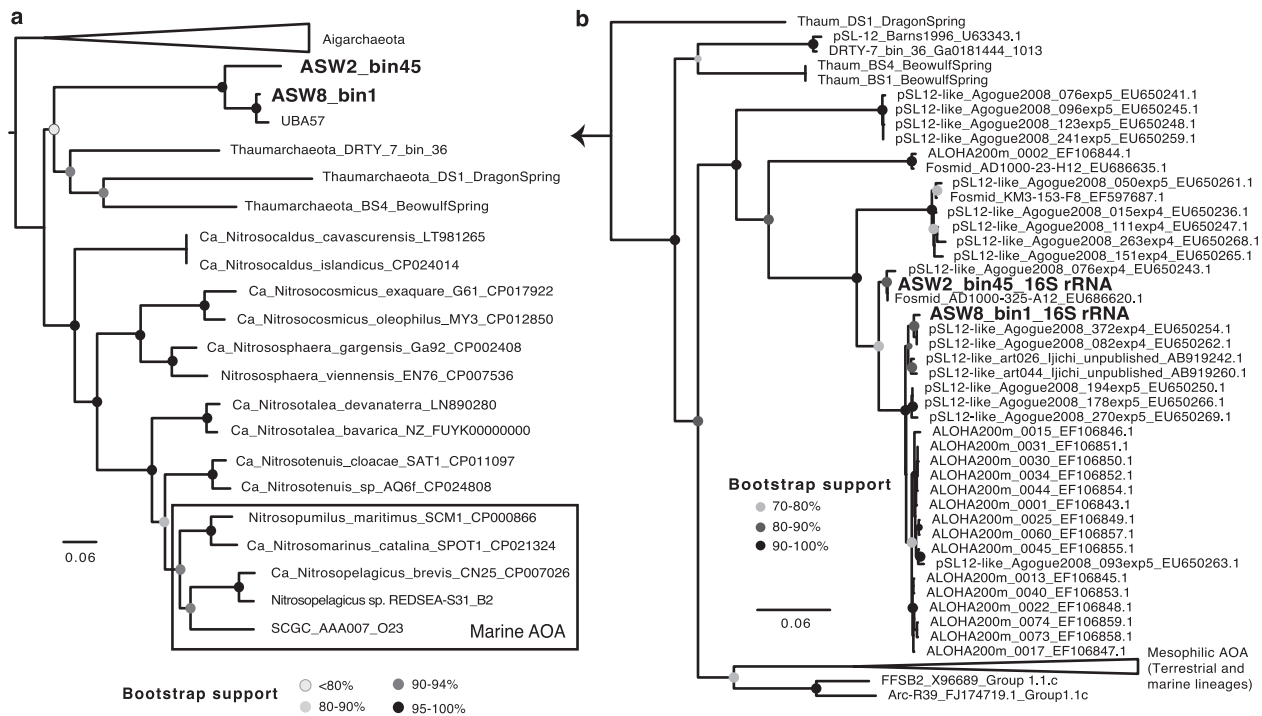
Unamended seawater incubations were started in 2010, using water collected from various depths in Monterey Bay (see Materials and methods). Prior to metagenome sequencing, 16S rRNA gene amplicon libraries were generated to examine the community composition in each incubation. This suggested an enrichment of Thaumarchaeota in both samples presented here (Fig. S1), and these samples were further examined via metagenome sequencing. Assembly and binning resulted in three genomic bins of the pSL12-like lineage, which were further dereplicated into two MAGs (see Materials and methods).

The MAGs assembled here represent the first high-quality genomes reported for the pSL12-like lineage (completion estimates for the two MAGs are 88.8 and 97.08%, with <3% contamination; Table 1). Their relative placement within the phylum Thaumarchaeota was confirmed by both phylogenomic and single-gene phylogenetic analyses (Fig. 1). Both MAGs contained two partial copies each of the 16S ribosomal rRNA gene. On two separate maximum-likelihood trees computed on nucleotide alignments that included reference sequences from all major thaumarchaeal lineages, the MAG-derived 16S rRNA sequences clustered with Group 1A clone fragments generated from various ocean regions in prior studies (Figs. 1a and S2; [11–13]). The 3'-truncated 16S rRNA gene fragments within the MAGs shared 93.85% nucleotide identity along the aligned region (910 aligned positions), while the 5'-truncated fragments shared 92.32% nucleotide identity (664 aligned positions). The original primer pairs developed by Mincer et al. [11] to target the pSL12-like lineage aligned without any mismatches to the longer 3'-truncated 16S rRNA gene fragments from both genomes. Similarly, the widely used universal primers targeting the V4-V5 region of the 16S rRNA gene [49] also aligned with the MAG-derived sequences, again without any mismatches. Thus, microbial community surveys employing either of these primer sets should pick up the pSL12-like/Group 1A lineage. We were able to verify this in a high-resolution 16S rRNA gene dataset generated from Monterey Bay targeting the V4-V5 region (see discussion below).

**Table 1** Metagenome-assembled genome (MAG) statistics.

MAG ID	Completion	Contamination	Number of contigs	N50	Number of bases	ANI with UBA57
ASW8_bin1	97.08%	2.91%	91	16957	996,535	95.00%
ASW2_bin45	88.83%	0.97%	46	35482	918,577	80.21%

Estimates of genome completeness and contamination. Also presented are the average nucleotide identity (ANI) comparisons with UBA57, a closely related metagenome-assembled genome from a deep hydrothermal plume metagenome (See Materials and methods for details).



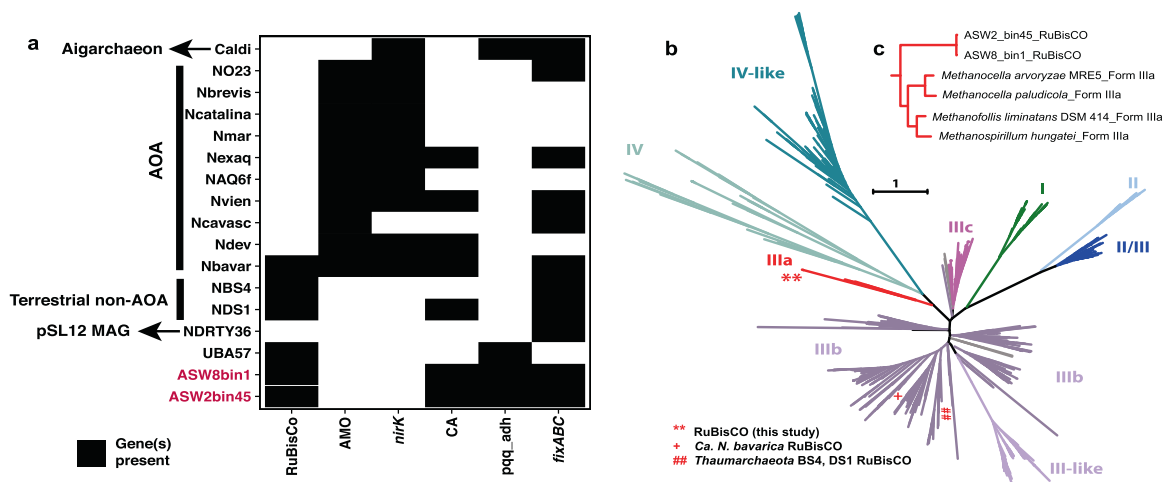
**Fig. 1** The assembled genomes cluster within the marine pSL12-like thaumarchaeal lineage. **a** Maximum-likelihood phylogenomic tree computed using a concatenated alignment of 11 ribosomal proteins. Bootstrap values are indicated on nodes. See Materials and

methods for details on alignment and tree computation. **b** Phylogeny of MAG-derived 16S rRNA gene sequences with genomic as well as environmental reference sequences. Node shading indicates bootstrap support.

The closest genomic relative in the database was a MAG obtained from a hydrothermal vent plume metagenome (from 4900 m depth on the Mid-Cayman Spreading Center [45]), which potentially represents a species-level relative [50] of ASW8\_bin1 (Table 1). Within a maximum-likelihood tree computed using a concatenated alignment of 11 core ribosomal proteins, the two MAGs were placed as a sister-clade to all known ammonia-oxidizing Thaumarchaeota of Group 1.1a (marine AOA) and 1.1b (soil AOA) (Fig. 1a). Similarly, based on 16S rRNA gene phylogeny, the MAGs clustered with environmental clone sequences of the pSL12-like clade (Fig. 1b). The original hot spring pSL12 lineage (including the only available MAG for this lineage, DRTY-7 bin\_36, assembled from a hot spring metagenome [6]) comprised a distant sister clade to the marine pSL12-like group.

## Metabolic potential distinct from typical marine Thaumarchaeota

Capacity for ammonia oxidation was not detected in either MAG, as we could not retrieve homologs of the AMO or nitrite reductase (*nirK*) genes. Moreover, the carbon-fixation pathway uniquely found in chemolithoautotrophic Thaumarchaeota—a modified version of the 3-hydroxypropionate/4-hydroxybutyrate (HP/HB) cycle [51]—appeared to be missing in both genomes. The myriad of multicopper oxidases characteristic of mesophilic AOA genomes [52] were also missing; although manual BLASTp searches did identify copper-binding proteins of the plastocyanin/azurin family in both genomes. These genes were located in the vicinity of cytochrome or ATP synthase proteins, suggesting a role in electron transfer. Since the



**Fig. 2 Metabolic capabilities of pSL12-like clade distinct from typical AOA.** **a** Comparison of selected metabolic features across thaumarchaeal genomes. pSL12-like MAGs are highlighted in red. *Caldiarchaeum subterraneum* belonging to the closely related candidate phylum Aigarchaeota, is also included for comparison. Gene abbreviations: AMO, ammonia monoxygenase; *nirK*, nitrite reductase; CA, carbonic anhydrase; pqq-adh, PQQ-dependent alcohol dehydrogenase; *fixABC*, electron transferring flavoprotein subunits A, B, and C. Taxa abbreviations: Caldi, *Ca. Caldiarchaeum subterraneum*; NO23, SCGC AAA007 O23; Nbrevis, *Ca. Nitrosopelagicus brevis* CN25; Ncatalina, *Ca. Nitrosomarinus catalina* SPOT01; Nexaq, *Ca. Nitrosocosmicus exaquare*; NAQ6f, *Ca. Nitrosotenuis aquarius* AQ6f; Nvien, *Nitrososphaera viennensis*; Ncavasc, *Ca.*

genomes are not closed, our failure to detect the ‘expected’ pathways/genes does not definitively indicate their absence. However, there were striking differences in the overall genomic repertoire of typical AOA genomes and the MAGs recovered here (Fig. 2a), which cannot be explained by the lack of genome completeness alone.

None of the six canonical carbon-fixation pathways were complete in the MAGs. It is possible that these Thaumarchaeota may use the recently described reverse oxidative TCA cycle for CO<sub>2</sub> fixation [53], since the genomes contained fumarate reductases, and 2-oxoglutarate/2-oxoacid ferredoxin oxidoreductases. In this pathway, a reversible citrate synthase catalyzes the production of citrate from acetyl-CoA. Recently, metabolic reconstructions were used to predict the existence of the roTCA cycle in Aigarchaeota [6]. However, we take caution in asserting roTCA CO<sub>2</sub> fixation in pSL12-like Thaumarchaeota, since genomic inference alone is not sufficient evidence for this pathway (many of the enzymes are bifunctional and common with the anabolic TCA cycle).

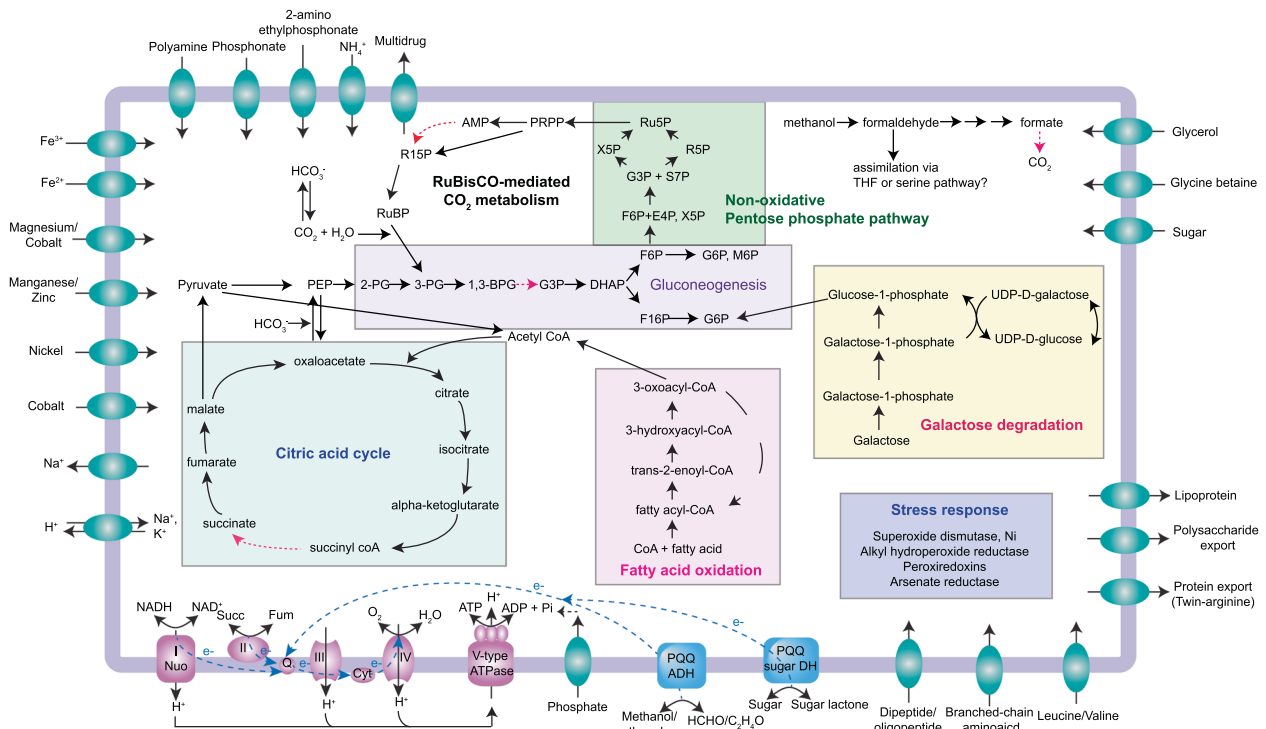
### Metabolic reconstructions indicate aerobic heterotrophy

The presence of respiratory complexes and various organic carbon-assimilating metabolic pathways (e.g., fatty acid oxidation, sugar metabolism, amino acid degradation, and

*Nitrosocaldus cavascurensis*; Ndev, *Ca. Nitrosotalea devanattera*; Nbavar, *Ca. Nitrosotalea bavarica*; BS4, Thaumarchaeota archaeon BS4 (MAG); DS1, Thaumarchaeota archaeon DS1 (MAG); and DRTY36, DRTY-7 bin\_36 (MAG). **b** Phylogenetic tree of RuBisCO sequences computed in FastTree using a MAFFT alignment of amino acid sequences. The MAG-derived RuBisCO sequences are highlighted. Previously reported thaumarchaeal RuBisCO sequences are also highlighted. Forms I, II, and III exhibit carboxylation activity, whereas the form IV RuBisCO does not. **c** Subtree highlighting the relative placements of the MAG-derived RuBisCO sequences with respect to Form III-a sequences from methanogens. This subtree was extracted from the original maximum-likelihood tree presented in (b).

potential methylotrophy; Fig. 3) suggest a predominantly heterotrophic lifestyle for these Thaumarchaeota. No external inorganic electron donors were identified based on the genome annotations. In addition to the aerobic respiratory chain, both genomes contained electron transfer flavoprotein (*fixABC*) homologs. These proteins are involved in electron transfer to nitrogenase in diazotrophic bacteria [54]. Homologs of *fixABC* have previously been reported in non-diazotrophic archaea, including terrestrial AOA [55]; yet their functional role in non-diazotrophs remains unclear. The *fix* operon has not been reported in marine AOA, but it appears that the deep marine AOA clade (water column B (WCB) group found predominantly at depths >200 m [56]) may also contain the *fix* genes (Fig. 2a; SCGC AAA007 O23 is a representative WCB genome). As discussed in a later section, the pSL12-like lineage appears to be particularly abundant deeper in the water column, resembling the distribution of the WCB lineage (also observed in a recent survey of Thaumarchaeota communities in Monterey Bay [17]). The presence of *fixABC* genes in these two clades might be a reflection of their niche adaptation, and will need to be investigated further.

Unlike other AOA, our two MAGs encoded several pyrroloquinoline quinone (PQQ)-dependent dehydrogenases containing N-terminal signal peptides (indicating extracellular localization), which can directly contribute reducing equivalents to the respiratory chain via



**Fig. 3 Overview of metabolic potential based on metabolic reconstructions of the pSL12-like MAGs.** Red dashed arrows indicate unidentified genes. The TCA cycle is presented in the anabolic direction. For detailed gene information, see Dataset 1. PEP phosphoenol pyruvate, 2-PG 2-phosphoglycerate, 3-PG 3-phosphoglycerate, 1,3-BPG 1,3-bisphosphoglycerate, G3P glyceraldehyde 3-phosphate, DHAP dihydroxyacetone phosphate, F6P fructose 6-phosphate, F16P fructose 1,6-bisphosphate, G6P glucose 6-phosphate, M6P mannose 6-phosphate, E4P erythrose 4-phosphate, X5P xylulose 5-phosphate, S7P sedoheptulose 7-phosphate, G3P glyceraldehyde 3-phosphate, PRPP phosphoribosylpyrophosphate, AMP adenosine monophosphate, R15P ribose 1,5-bisphosphate, RuBP ribulose 1,5-bisphosphate.

phosphate, F16P fructose 1,6-bisphosphate, G6P glucose 6-phosphate, M6P mannose 6-phosphate, E4P erythrose 4-phosphate, X5P xylulose 5-phosphate, S7P sedoheptulose 7-phosphate, G3P glyceraldehyde 3-phosphate, PRPP phosphoribosylpyrophosphate, AMP adenosine monophosphate, R15P ribose 1,5-bisphosphate, RuBP ribulose 1,5-bisphosphate.

extracellular sugar and/or alcohol oxidation (Fig. 3). Specific proteins were identified in both MAGs as putative PQQ-dependent methanol, ethanol, and glucose dehydrogenases (Dataset 1). Both methanol and glucose dehydrogenases that use PQQ as the prosthetic group are known to catalyze the oxidation of diverse alcohols and hexoses/pentoses, respectively [57], suggesting some degree of metabolic versatility in these archaea. PQQ synthase proteins were also identified in both genomes (Dataset 1). Up to 5 quinoprotein dehydrogenases were found to be colocalized on the same contig, along with amicyanin/plastocyanin-like small copper proteins and ATP synthase subunits (e.g., contigs ASW2bin45\_2 and ASW8bin1\_21; Dataset 1), indicating their combined involvement in an electron transport chain coupled to energy conservation.

Formaldehyde resulting from methanol oxidation is cytotoxic, and hence is promptly removed via dissimilatory or assimilatory pathways. Formaldehyde oxidation to formate likely proceeds via the tetrahydromethanopterin ( $H_4MPT$ ) pathway in these archaea, as the annotated genes included a F420-dependent methylene-tetrahydromethanopterin dehydrogenase (*mtd*) and a methylene-tetrahydromethanopterin cyclohydrolase/reductase (Dataset 1). Whether formaldehyde

oxidation proceeds all the way to  $CO_2$  is unclear based on the annotations, since neither MAG encoded a formate dehydrogenase. Alternatively, formaldehyde may also get assimilated via the tetrahydrofolate or the serine pathway (neither pathway annotations were complete).

Metabolic reconstructions suggest the use of diverse organic compounds as potential electron donors. In addition to the fatty acid oxidation pathway, multiple sugar transporters with homology to trehalose/maltose import proteins and arabinose permeases were annotated in the MAGs (Dataset 1). Both MAGs also encoded a halolysin-like protease, which may hydrolyze proteins extracellularly and the resulting peptides may be imported as nutrients. Supporting this, peptide ABC transporter permease proteins and branched-chain amino acid transporters were identified in both genomes. Protein topology modeling suggested the extracellular localization of the halolysin protease, suggesting its involvement in protein degradation externally.

Thaumarchaeal lineages previously identified as basal groups lacking the capacity to oxidize ammonia (which were obtained from nonmarine environments) are reported to possess anaerobic energy generation pathways such as sulfate or nitrate reduction [5]. The MAGs assembled here

contained no definitive evidence for anaerobic respiration, although we acknowledge this might be due to the lack of genome completeness. Moreover, many of the genomic features identified as unique/core features for the anaerobic basal thaumarchaeal lineages in a recent comparative meta-analysis [8] were also absent in these MAGs [(i.e., pyruvate: ferredoxin oxidoreductase (*porABDG*), cytochrome bd-type terminal oxidase (*cydA*), and acetyl-CoA decarboxylase/synthase (*codhAB*)]. Thus, multiple lines of evidence point to these MAGs representing a divergent, basal lineage within the aerobic, mesophilic clade of Thaumarchaeota.

### Metabolic hypothesis on a RuBisCO-mediated anaplerotic CO<sub>2</sub> assimilation pathway

Unexpectedly, both MAGs harbored an archaeal type III RuBisCO gene (463 aa long; 96.76% amino acid identity to each other). Hypothesized to be the most ancient form of RuBisCO, form III is predominantly found in Archaea [58]. Recent metagenomic surveys have revealed numerous members of the candidate phyla radiation [59, 60] and DPANN archaea [43, 61] also encoding a form III-like RuBisCO. A divergent variant categorized as form III-a is found in methanogenic archaea. Our MAG-derived sequences clustered with the methanogen III-a RuBisCO sequences (Fig. 2b–c), albeit with 30–35% amino acid identity.

Two separate studies have previously reported a form III RuBisCO in Thaumarchaeota, and in both cases the assembled genomes represented acidophilic terrestrial lineages: (i) *Ca. Nitrosotalea bavarica* SbT1 was assembled and binned from an acidic peatland metagenome [62], and (ii) the deeply branching BS4 and DS1 were assembled from acidic geothermal spring sediments in Yellowstone National Park [5]. RuBisCO sequences from these MAGs clustered within the main archaeal form III clade (Fig. 2b), and were <30% identical (in the amino acid space) to the sequences we obtained in this study.

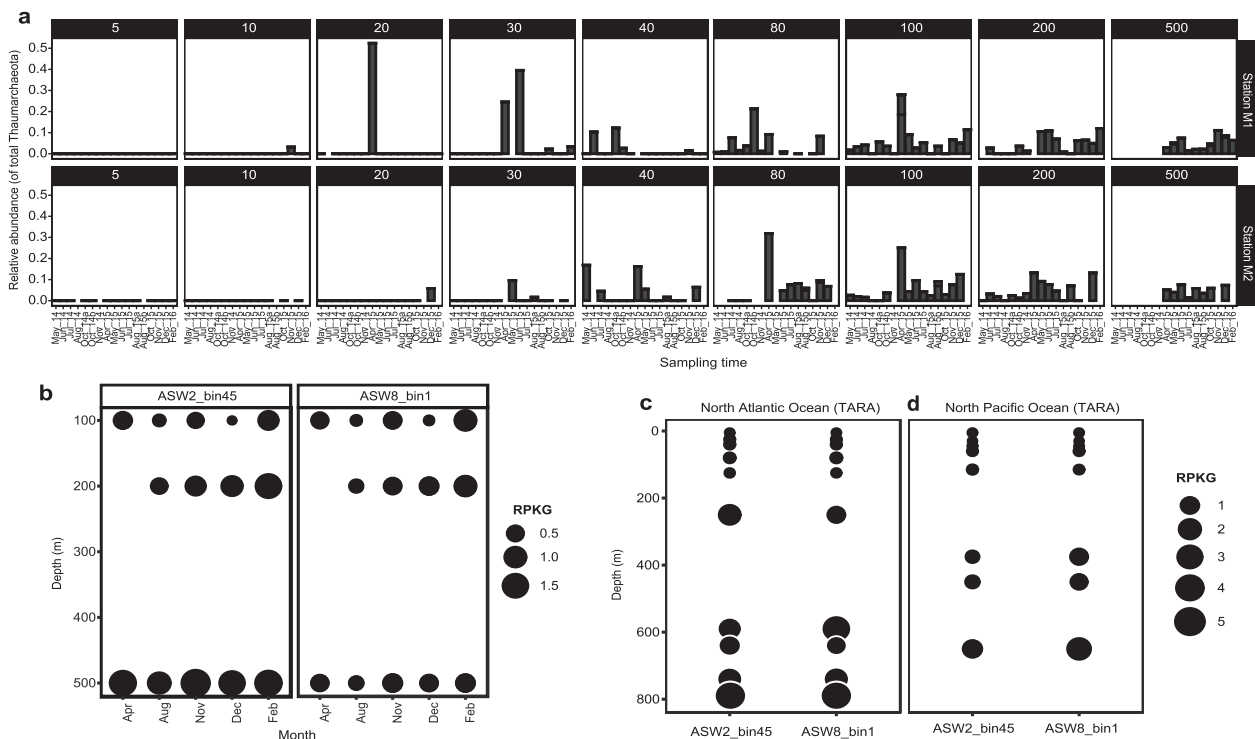
Despite exhibiting carboxylase activity, genomic and biochemical evidence suggest that form III RuBisCO is not involved in carbon fixation via the canonical Calvin–Benson–Bassham (CBB) cycle [63, 64]. In many archaea harboring RuBisCO, phosphoribulokinase (PRK) required for the regeneration of the RuBisCO substrate (RuBP) is missing [63], suggesting the absence of a functional CBB pathway. Intriguingly, methanogenic archaea harboring form III-a RuBisCO encode a PRK, yet are missing other key components of the CBB cycle [65]. Thus, RuBisCO in these methanogens is thought to be involved in carbon assimilation via the reductive-hexulose-phosphate (RHP) pathway [65]. As demonstrated in *Methanospirillum hungatei*, RuBP regeneration in the RHP pathway involves the activity of PRK, as well as the formaldehyde-

assimilating ribulose monophosphate (RuMP) pathway operating in reverse [65].

The second proposed route for form III RuBisCO-mediated carbon metabolism involves nucleoside assimilation/degradation via the archaeal AMP pathway [63, 64]. Briefly, adenosine monophosphate (AMP, retrieved from the phosphorylation of nucleosides) is converted to ribose 1,5-bisphosphate (R15P) by AMP phosphorylase. Subsequently, R15P is isomerized to ribulose 1,5-bisphosphate (RuBP) by ribose 1,5-bisphosphate isomerase (R15Pi). In an irreversible reaction, RuBisCO combines RuBP with CO<sub>2</sub> and H<sub>2</sub>O to yield 3-phosphoglycerate (3-PG), which then enters the central carbon metabolism (via glycolysis or gluconeogenesis). Sato et al. [63] proposed that the reductive pentose phosphate pathway, if present, may cyclize the above-described series of transformations, effectively rendering it a carbon-fixation pathway.

Homology comparisons revealed the conservation of key active site residues for carboxylation in our MAG-derived RuBisCO sequences (Fig. S3). However, little evidence exists to support the methanogenic RHP CO<sub>2</sub> fixation pathway—in addition to a missing PRK, many key enzymes in the methanogenic RHP and RuMP pathways could not be identified. Metabolic inferences best support an anaplerotic function for the carboxylation reaction via the AMP pathway for nucleotide salvage. A key difference from the archaeal AMP pathway, however, is the presence of a complete non-oxidative pentose phosphate pathway (nPPP) and gluconeogenesis in the pSL12-like lineage. The nPPP pathway operating in reverse to generate R5P from gluconeogenesis intermediates, combined with RuBP regeneration from PRPP and/or AMP, might constitute a cyclic CO<sub>2</sub> fixation pathway ([63, 66]; Fig. 3). Several of the genes encoding key enzymes in the proposed pathway appeared to be colocalized on the same assembled contigs in both MAGs (Fig. S4), suggesting potential co-expression. This pathway, however, likely has an anaplerotic function, potentially regulated by intracellular levels of AMP and/or PRPP. We, however, emphasize that the proposed pathway is inferred purely via bioinformatic methods, and may well be impacted by the lack of genome completeness.

A gamma-class carbonic anhydrase (CA) was present in both MAGs, which catalyzes the interconversion of CO<sub>2</sub> and HCO<sub>3</sub><sup>-</sup>. CA homologs have been identified in several terrestrial AOA, and are hypothesized to function extracellularly to facilitate CO<sub>2</sub> uptake for carbon fixation [52]. However, marine lineages do not harbor CA genes (Fig. 2a). Unlike the CAs from terrestrial AOA, the pSL12-like CAs did not contain signal peptide sequences and, therefore, are likely involved in intracellular reversible dehydration of HCO<sub>3</sub><sup>-</sup> to CO<sub>2</sub>. While CA is not exclusively



**Fig. 4** Distribution of pSL12-like lineage in oceanic water columns. **a** Relative abundances (as a percentage of total thaumarchaeal abundance) of OTUs  $\geq 96\%$  identical to the 16S rRNA gene sequences retrieved from the MAGs. The two major panels correspond to two oceanographic sampling stations, M1 and M2, in Monterey Bay. Each subpanel represents a depth profile between 5 and 500 m. **b** Read recruitments of each MAG against Monterey Bay metagenomes. Size

of the circle corresponds to normalized abundance. **c, d** Metagenome read recruitments against Atlantic Ocean and Pacific Ocean depth profiles, respectively, from the TARA Oceans dataset. Relative abundances are presented as number of reads mapped per kilobase of genome per gigabase of metagenome (RPKG). Metagenome sample accessions are provided in Table S1.

indicative of carbon fixation, its activity may facilitate  $\text{CO}_2$  incorporation by RuBisCO and/or phosphoenol pyruvate carboxykinase in the pSL12-like Thaumarchaeota.

### Distribution of the pSL12-like lineage in the water column

To assess the environmental distribution of the pSL12-like lineage, we matched the MAG-derived 16S rRNA sequences to a previously generated 16S rRNA amplicon dataset from the Monterey Bay upwelling system [46]. One of the MAG-derived 16S rRNA gene sequences (from ASW8\_bin1) was an exact match to an OTU #694, which comprised  $<0.5\%$  of the total thaumarchaeal abundance at any given time in the depths sampled. The next closest match was OTU #8597, which shared 96.02% and 97.08% sequence identity with sequences from ASW2\_bin45 and ASW8\_bin1, respectively. At any given time, these two OTUs together comprised at most 0.5% of thaumarchaeal abundance in the time-series dataset (Fig. 4a). As observed in previous surveys, the pSL12-like group of Thaumarchaeota appeared to be more abundant below the euphotic zone [11, 13, 15, 16, 17], with potential seasonal variations

in relative abundances. Occasional abundance peaks were observed in the photic zone during spring at M1 (Fig. 4a), which likely reflects upwelled populations (station M1 is situated directly above the upwelling plume in Monterey Bay).

In recruiting metagenomic reads from Monterey Bay against the MAGs, we observed the highest recruitment at 500 m for ASW2\_bin45. ASW8\_bin1 recruited slightly fewer reads but appeared to have a similar abundance distribution across depths as ASW2\_bin45 (Fig. 4b). In addition, the relative abundances appeared to change with seasonal hydrologic changes in the system (Fig. 4b). Recruitment against TARA Ocean metagenomes representing Atlantic Ocean and Pacific Ocean depth profiles revealed similar depth distribution of the pSL12-like lineage, with the greatest abundance at depths well below the euphotic zone (200–800 m; Fig. 4c).

### Conclusions

In this work, we used reconstructed population genomes to infer metabolic adaptations of the elusive pSL12-like



lineage of Thaumarchaeota, widely distributed in marine systems. The high-quality genomes described here offer a first glimpse into the genomic repertoire of a marine thaumarchaeal group devoid of an exclusively chemoautotrophic energy generation strategy. Only terrestrial basal lineages of Thaumarchaeota have been described thus far; the MAGs presented here represent the first genomic description of a basal lineage inhabiting the marine oxic environment. In this context, an especially intriguing consideration is the relative positioning of the pSL12-like clade within the thaumarchaeal evolutionary trajectory. The diversification of Thaumarchaeota, from basal groups to the mesophilic AOA appears to have included multiple metabolic changes—acquiring the 3-HP/4-HB pathway for CO<sub>2</sub> fixation, ammonia oxidation, and potential differences in co-factor use, among others (Fig. 2a; also reviewed in [8]). The MAGs described here represent a basal lineage that appears to coexist with aerobic ammonia-oxidizing Thaumarchaeota in marine waters (basal lineages reported thus far have been found in terrestrial systems, as reviewed in [8]). These MAGs may thus enable a more detailed probing of the trajectory leading to marine AOA evolution from basal groups, and help constrain the relative timing of the acquisition of aerobic metabolism and ammonia oxidation within the phylum.

Overall, the divergent genomic features of the pSL12-like clade significantly alter our understanding of the metabolic diversity within this abundant archaeal phylum in the oceans. While further biochemical characterization is warranted to confirm the proposed metabolic pathways, our results suggest that obligate aerobic heterotrophy might be an overlooked metabolic strategy within pelagic Thaumarchaeota.

**Acknowledgements** Sequencing was carried out as part of a Community Science Program (CSP) grant to CAF from the DOE Joint Genome Institute. Computing for this project was performed on the Sherlock 2.0 cluster. We would like to thank Stanford University and the Stanford Research Computing Center for providing computational resources and support that contributed to the results presented here. We thank Marie Lund and Bradley B. Tolar for help with sample and data acquisition, respectively. We also thank Dr Alfred Spormann for helpful feedback and discussion on an early draft of the manuscript. This study was supported (in part) by grant OCE-1357024 from NSF Biological Oceanography (to CAF). The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- Kamer MB, DeLong EF, Karl DM. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature*. 2001;409:507–10.
- Oton EV, Quince C, Nicol GW, Prosser JI, Gubry-Rangin C. Phylogenetic congruence and ecological coherence in terrestrial Thaumarchaeota. *ISME J*. 2016;10:85–96.
- Lin X, Handley KM, Gilbert JA, Kostka JE. Metabolic potential of fatty acid oxidation and anaerobic respiration by abundant members of Thaumarchaeota and Thermoplasmata in deep anoxic peat. *ISME J*. 2015;9:2740–4.
- Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, et al. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat Commun*. 2016;7:1–11.
- Beam JP, Jay ZJ, Kozubal MA, Inskeep WP. Niche specialization of novel Thaumarchaeota to oxic and hypoxic acidic geothermal springs of Yellowstone National Park. *ISME J*. 2014;8:938–51.
- Hua Z-S, Qu Y-N, Zhu Q, Zhou E-M, Qi Y-L, Yin Y-R, et al. Genomic inference of the metabolism and evolution of the archaeal phylum Aigarchaeota. *Nat Commun*. 2018;9:208–11.
- Weber EB, Lehtovirta-Morley LE, Prosser JI, Gubry-Rangin C, Laanbroek R. Ammonia oxidation is not required for growth of Group 1.1c soil Thaumarchaeota. *FEMS Microbiol Ecol*. 2015;91:fiv001 <https://doi.org/10.1093/femsec/fiv001>.
- Ren M, Feng X, Huang Y, Wang H, Hu Z, Clingenpeel S, et al. Phylogenomics suggests oxygen availability as a driving force in Thaumarchaeota evolution. *ISME J*. 2019;13:2150–61.
- DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, Frigaard N-U, et al. Community genomics among stratified microbial assemblages in the ocean's interior. *Science*. 2006;311:496–503.
- Barns SM, Delwiche CF, Palmer JD, Pace NR. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *PNAS*. 1996;93:9188–93.
- Mincer TJ, Church MJ, Taylor LT, Preston C, Karl DM, DeLong EF. Quantitative distribution of presumptive archaeal and bacterial nitrifiers in Monterey Bay and the North Pacific Subtropical Gyre. *Environ Microbiol*. 2007;9:1162–75.
- Martin-Cuadrado A-B, Rodriguez-Valera F, Moreira D, Alba JC, Ivars-Martínez E, Henn MR, et al. Hindsight in the relative abundance, metabolic potential and genome dynamics of uncultivated marine archaea from comparative metagenomic analyses of bathypelagic plankton of different oceanic regions. *ISME J*. 2008;2:865–86.
- Agogue H, Brink M, Dinasquet J, Herndl GJ. Major gradients in putatively nitrifying and non-nitrifying Archaea in the deep North Atlantic. *Nature*. 2008;456:788–91.
- La Cono V, Smedile F, Ferrer M, Golyshin PN, Giuliano L, Yakimov MM. Genomic signatures of fifth autotrophic carbon assimilation pathway in bathypelagic Crenarchaeota. *Micro Biotechnol*. 2010;3:595–606.
- Church MJ, Wai B, Karl DM, DeLong EF. Abundances of crenarchaeal *amoA* genes and transcripts in the Pacific Ocean. *Environ Microbiol*. 2010;12:679–88.
- Tolar BB, King GM, Hollibaugh JT. An analysis of Thaumarchaeota populations from the Northern Gulf of Mexico. *Front Microbiol*. 2013;4:72 <https://doi.org/10.3389/fmicb.2013.00072>.
- Tolar BB, Reji L, Smith JM, Blum M, Pennington JT, Chavez FP, et al. Time series assessment of Thaumarchaeota ecotypes in Monterey Bay reveals the importance of water column position in predicting distribution-environment relationships. *Limnol Oceanogr*. 2020. <https://doi.org/10.1002/LNO.11436>.
- Bushnell B. BBTools software package. 2014. [sourceforge.net/projects/bbmap/](https://sourceforge.net/projects/bbmap/)

19. Li H. BFC: correcting Illumina sequencing errors. *Bioinformatics*. 2015;31:2885–7.
20. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*. 2015;31:1674–6.
21. Li D, Luo R, Liu C-M, Leung C-M, Ting H-F, Sadakane K, et al. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods*. 2016;102:3–11.
22. Kang D, Li F, Kirton ES, Thomas A, Egan RS, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. 2019;7:e7359 <https://doi.org/10.7717/peerj.7359>
23. Wu Y-W, Tang Y-H, Tringe SG, Simmons BA, Singer SW. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome*. 2014;2:26 <https://doi.org/10.1186/2049-2618-2-26>
24. Wu Y-W, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Nucleic Acids Res*. 2016;32:605–7. <https://doi.org/10.1093/bioinformatics/btv638>
25. Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome*. 2018;6:1–13.
26. Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, et al. Assembling genomes and mini-metagenomes from highly chimeric reads. *J Comput Biol*. 2013;20:714–37.
27. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015;25:1043–55.
28. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol*. 2018;36:996–1004.
29. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J*. 2017;11:2864–8.
30. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinforma*. 2010;11:119.
31. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30:2068–9.
32. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol*. 2016;428:726–31.
33. Graham ED, Heidelberg JF, Tully BJ. Potential for primary productivity in a globally-distributed bacterial phototroph. *ISME J*. 2018;12:1861–6.
34. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2019. <https://www.r-project.org/>.
35. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res*. 2013;42:D206–14.
36. Elbourne LDH, Tetu SG, Hassan KA, Paulsen IT. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Res*. 2016;45:D320–4.
37. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, et al. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ*. 2015;3:e1319 <https://doi.org/10.7717/peerj.1319>
38. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32:1792–7.
39. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 2009;25:1972–3.
40. Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE*. 2010;5:e9490 <https://doi.org/10.1371/journal.pone.0009490>
41. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215:403–10.
42. Katoh K, Misawa K, Kuma KI, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 2002;30:3059–66.
43. Jaffe AL, Castelle CJ, Dupont CL, Banfield JF. Lateral gene transfer shapes the distribution of RuBisCO among Candidate Phyla Radiation Bacteria and DPANN Archaea. *Mol Biol Evol*. 2019;36:435–46.
44. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun*. 2018;9:1–8.
45. Parks DH, Rinke C, Chuvochina M, Chaumeil P-A, Woodcroft BenJ, Evans PN, et al. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat Microbiol*. 2017;2:1533–42.
46. Reji L, Tolar BB, Smith JM, Chavez FP, Francis CA. Differential co-occurrence relationships shaping ecotype diversification within Thaumarchaeota populations in the coastal ocean water column. *ISME J*. 2019;13:1144–58.
47. Reji L, Tolar BB, Smith JM, Chavez FP, Francis CA. Depth distributions of nitrite reductase (*nirK*) gene variants reveal spatial dynamics of thaumarchaeal ecotype populations in coastal Monterey Bay. *Environ Microbiol*. 2019;21:4032–45. <https://doi.org/10.1111/1462-2920.14753>
48. Langmead Ben, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357–9.
49. Parada AE, Needham DM, Fuhrman JA. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol*. 2015;18:1403–14.
50. Konstantinidis KT, Tiedje JM. Genomic insights that advance the species definition for prokaryotes. *PNAS*. 2005;102:2567–72.
51. Könneke M, Schubert DM, Brown PC, Hügler M, Standfest S, Schwander T, et al. Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO<sub>2</sub> fixation. *PNAS*. 2014;111:8239–44.
52. Kerou M, Offre P, Valledor L, Abby SS, Melcher M, Nagler M, et al. Proteomics and comparative genomics of *Nitrososphaera viennensis* reveal the core genome and adaptations of archaeal ammonia oxidizers. *PNAS*. 2016;113:E7937–46. <https://doi.org/10.1073/pnas.1601212113>
53. Mall A, Sobotta J, Huber C, Tschirner C, Kowarschik S, Bacnik K, et al. Reversibility of citrate synthase allows autotrophic growth of a thermophilic bacterium. *Science*. 2018;359:563–7.
54. Edgren T, Nordlund S. The *fixABCX* genes in *Rhodospirillum rubrum* encode a putative membrane complex participating in electron transfer to nitrogenase. *J Bacteriol*. 2004;186:2052–60.
55. Bartossek R, Spang A, Weidler G, Lanzen A, Schleper C. Metagenomic analysis of ammonia-oxidizing archaea affiliated with the soil group. *Front Microbiol*. 2012;3:208.
56. Francis CA, Roberts KJ, Beman JM, Santoro AE, Oakley BB. Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. *PNAS*. 2005;102:14683–8.

57. Anthony C. The quinoprotein dehydrogenases for methanol and glucose. *Arch Biochem Biophys.* 2004;428:2–9.
58. Erb TJ, Zarzycki JA. A short history of RubisCO: the rise and fall (?) of Nature's predominant CO<sub>2</sub> fixing enzyme. *Curr Opin Biotechnol.* 2018;49:100–7.
59. Wrighton KC, Castelle CJ, Varaljay VA, Satagopan S, Brown CT, Wilkins MJ, et al. RubisCO of a nucleoside pathway known from Archaea is found in diverse uncultivated phyla in bacteria. *ISME J.* 2016;10:2702–14.
60. Wrighton KC, Thomas BC, Sharon I, Miller CS, Castelle CJ, VerBerkmoes NC, et al. Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. *Science.* 2012;337:1661–5.
61. Castelle CJ, Wrighton KC, Thomas BC, Hug LA, Brown CT, Wilkins MJ, et al. Genomic expansion of domain Archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr Biol.* 2015;25:690–701.
62. Herbold CW, Lehtovirta-Morley LE, Jung M-Y, Jehmlich N, Hausmann B, Han P, et al. Ammonia-oxidising archaea living at low pH: Insights from comparative genomics. *Environ Microbiol.* 2017;19:4939–52.
63. Sato T, Atomi H, Imanaka T. Archaeal type III RuBisCOs function in a pathway for AMP metabolism. *Science.* 2007;315:1003–6.
64. Aono R, Sato T, Imanaka T, Atomi H. A pentose bisphosphate pathway for nucleoside degradation in Archaea. *Nat Chem Biol.* 2015;11:355–60.
65. Kono T, Mehrotra S, Endo C, Kizu N, Matusda M, Kimura H, et al. A RuBisCO-mediated carbon metabolic pathway in methanogenic archaea. *Nat Commun.* 2017;8:1–12.
66. Falb M, Müller K, Königsmaier L, Oberwinkler T, Horn P, Gronau von S, et al. Metabolism of halophilic archaea. *Extremophiles.* 2008;12:177–96.