



HAL
open science

Genomes of Thaumarchaeota from deep sea sediments reveal specific adaptations of three independently evolved lineages

Melina Kerou, Rafael Ponce-Toledo, Rui Zhao, Sophie S. Abby, Miho Hirai, Hidetaka Nomaki, Yoshihiro Takaki, Takuro Nunoura, Steffen Jørgensen, Christa Schleper

► To cite this version:

Melina Kerou, Rafael Ponce-Toledo, Rui Zhao, Sophie S. Abby, Miho Hirai, et al.. Genomes of Thaumarchaeota from deep sea sediments reveal specific adaptations of three independently evolved lineages. *The International Society of Microbiological Ecology Journal*, 2021, 15 (9), pp.2792-2808. 10.1038/s41396-021-00962-6 . hal-03365290

HAL Id: hal-03365290

<https://hal.science/hal-03365290>







Submitted on 4 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Genomes of Thaumarchaeota from deep sea sediments reveal specific adaptations of three independently evolved lineages

Melina Kerou ¹ · Rafael I. Ponce-Toledo¹ · Rui Zhao^{2,5} · Sophie S. Abby ^{1,6} · Miho Hirai³ · Hidetaka Nomaki³ · Yoshihiro Takaki ³ · Takuro Nunoura ⁴ · Steffen L. Jørgensen ² · Christa Schleper ¹

Received: 12 June 2020 / Revised: 12 January 2021 / Accepted: 11 March 2021
© The Author(s) 2021. This article is published with open access

Abstract

Marine sediments represent a vast habitat for complex microbiomes. Among these, ammonia oxidizing archaea (AOA) of the phylum Thaumarchaeota are one of the most common, yet little explored, inhabitants, which seem extraordinarily well adapted to the harsh conditions of the subsurface biosphere. We present 11 metagenome-assembled genomes of the most abundant AOA clades from sediment cores obtained from the Atlantic Mid-Ocean ridge flanks and Pacific abyssal plains. Their phylogenomic placement reveals three independently evolved clades within the order *Nitrosopumilales*, of which no cultured representative is known yet. In addition to the gene sets for ammonia oxidation and carbon fixation known from other AOA, all genomes encode an extended capacity for the conversion of fermentation products that can be channeled into the central carbon metabolism, as well as uptake of amino acids probably for protein maintenance or as an ammonia source. Two lineages encode an additional (V-type) ATPase and a large repertoire of DNA repair systems that may allow to overcome the challenges of high hydrostatic pressure. We suggest that the adaptive radiation of AOA into marine sediments occurred more than once in evolution and resulted in three distinct lineages with particular adaptations to this extremely energy-limiting and high-pressure environment.

These authors contributed equally: Melina Kerou, Rafael I. Ponce-Toledo

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41396-021-00962-6>.

✉ Christa Schleper
christa.schleper@univie.ac.at

- ¹ Department of Functional and Evolutionary Ecology, Archaea Biology and Ecogenomics Unit, University of Vienna, Vienna, Austria
- ² Department of Earth Science, K.G. Jebsen Centre for Deep Sea Research, University of Bergen, Bergen, Norway
- ³ Super-cutting-edge Grand and Advanced Research (SUGAR) Program, X-star, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan
- ⁴ Research Center for Bioscience and Nanoscience (CeBN), Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan
- ⁵ Present address: School of Marine Science and Policy, University of Delaware, Lewes, DE, USA
- ⁶ Present address: University Grenoble Alpes, CNRS, Grenoble INP, TIMC-IMAG, Grenoble, France

Introduction

Ammonia oxidizing archaea (AOA) comprise one of the most successful archaeal phyla having colonized almost every imaginable oxic environment of the planet where they emerge as key players in the nitrogen cycle [1–6]. This includes the marine environment where they dominate archaeal communities associated with oxic sediments ranging from shallow estuaries to the open ocean [7–12], and from the surface layers all the way into the deep oceanic crust [13–15]. In these ecosystems, they seem to play a critical role in the transformation of nitrogen compounds and control their partitioning into the bottom ocean and the underlying oceanic crust [12, 14–19].

Studies from the North Atlantic and Pacific show that the composition of the sedimentary AOA population differs drastically from that in the overlying water suggesting distinct ecophysiological potential to colonize sedimentary environments, albeit all were found to belong to the order Nitrosopumilales (NP) [20]. Whereas the *amoA*-NP-gamma clade seems to be dominant and omnipresent in these oceans, irrespective of water depth, the *amoA*-NP-alpha clade represents the most abundant ecotype in deep ocean waters

[6, 7, 21–27] (nomenclature based on *amoA* gene classification [28]). In contrast, dominant phylotypes in deep sea sediments belong to the *amoA*-NP-theta and *amoA*-NP-delta clades [11, 12]. In addition, in cases of oligotrophic oceanic regions, these were detected throughout the sediment column and further into the underlying basaltic crust, even at depth where oxygen is below detection [7, 10–12, 14, 15]. In these sites, they exhibit peaks of abundance and diversity at oxic/anoxic transition zones where increased energy availability is suggested to sustain the higher biomass of nitrifiers [12]. The abundance and distribution of *amoA*-NP-theta and -delta in the energy-starved subsurface suggest that they have adapted and evolved differently than their pelagic counterparts. These clades represent a yet unexplored diversity within NP, and so far have no cultivated or genomic representatives [28].

It has been suggested that the AOA common ancestor arose in terrestrial habitats (probably hot springs) where the AOA lineages diversified and then occupied different biomes (e.g., soils, hot springs, and freshwater environments) before conquering estuarine and marine shallow water environments and finally, radiating into deeper waters as a result of the oxygenation of the deep ocean during the Neoproterozoic [29–31]. AOA are generally well equipped for the manifold challenges of the oxic deep sea surface and subsurface environment. They encode the most energy-efficient aerobic carbon fixation pathway [32] making them important primary producers in these environments [33, 34], and their high affinity for ammonia would enable them to utilize this scarce resource [35]. Nevertheless, deep pelagic as well as benthic AOA populations are reported to have the capability for mixotrophy as well, as indicated by uptake of labeled compounds and through the detection of uptake/assimilation genes for organic carbon and nitrogen compounds by shotgun metagenomics [10, 16, 23, 24, 34, 36–38]. Stimulation of autotrophic CO₂ fixation by organic carbon was also shown by isotope labeling studies [33]. In the absence of genomic context however, virtually nothing of the above can be extrapolated to the metabolic potential or adaptations of ecotypes that dominate deep marine sediments, nor can their ecological boundaries be interpreted.

In this study, we address the question of what adaptations enabled specific AOA clades to inhabit bathyal and abyssal (i.e., deep sea) marine sediments, and the significance of this in the context of thaumarchaeal evolution. To this end, we obtained the first high-quality metagenome-assembled genomes (MAGs) belonging to the so far uncharacterized *amoA*-NP-theta and *amoA*-NP-delta clades from sediment cores collected from the Mid-Atlantic Ridge flanks and the oligotrophic Pacific Ocean. We also describe two MAGs associated with a novel, deep-branching clade within the NP, which we designate *amoA*-NP-iota (previously NP—*incertae sedis* [28]). The pivotal phylogenetic position of the latter and the distribution of all three clades in

phylogenomic trees enables us to shed light on the evolutionary diversification of AOA into marine sediments, which seems much more complex than previously assumed and reveals unique, similar, and also overlapping adaptive strategies in all three clades.

Materials and methods

Sampling of Atlantic and Pacific sediments

Oligotrophic sediment cores were retrieved from mid-ocean ridge flanks in the Atlantic Ocean: Hole U1383E (22°48.1' N, 46°03.2'W, 4425 m water depth) from North Pond by advanced piston coring during the International Ocean Drilling Program (IODP) Expedition 336 (2011), and GS14-GC08 (71°58.0'N, 0°6.1'E, 2476 m water depth) by gravity coring from the east flank of the central Mohns Ridge (2014) (Fig. 1a). Genomic DNA from four sediment horizons in each core was selected for metagenome sequencing, based on the published porewater geochemical data and 16S rRNA gene profiles (Fig. S1) [12, 39]. In particular, sediments of 0.1 m below the sea floor (mbsf) (oxic), 10.0 mbsf (oxic), 22.0 mbsf (oxic–anoxic transition zone; OATZ), and 29.5 mbsf (anoxic–oxic transition zone) were selected from U1383E. Sediments of 0.1 mbsf (oxic), 1.0 mbsf (OATZ), 1.6 mbsf (nitrate–ammonium transition zone), and 2.5 mbsf (Mn-reduction zone) were selected from GS14-GC08 (Fig. 1b). Detailed information about sampling sites, sampling procedure, 16S rRNA gene profiles, and porewater analysis was published in refs. [12, 39].

Sediment cores (YK1309-1N and YK1312-12N) from the Pacific abyssal plain were collected using a push corer with a manned submersible Shinkai 6500 during the JAMSTEC cruises YK1309 (September 2013: 01°15.0'N, 163°14.9'E, 4277 m water depth) and YK1312 (November 2013: 11°59.9'N, 153°59.9'E, 5920 m water depth) of the *R/V Yokosuka*, respectively. Two sections from each core were selected for shotgun metagenomic sequencing: YK1309-1N-S000 (0–0.01 mbsf), YK1309-1N-S300 (0.3–0.35 mbsf), YK1312-12N-S010 (0.01–0.02 mbsf), and YK1312-12N-S200 (0.2–0.25 mbsf) (Figs. 1b and S1).

At the Pacific sites, bottom seawater temperatures were measured by a conductivity, temperature, and density (CTD) system (SEA-BIRD SBE19; Bellevue, WA, USA) settled in the ROV Shinkai 6500 and ranged from 1.3 to 1.4 °C (Table 1). At the Atlantic sites, bottom seawater temperatures of ~2.5 °C were reported at North Pond [40] and –0.7 °C measured by CTD in the vicinity of the site at Mohns Ridge GC08 (Table 1). For samples collected in the Pacific and at Mohns ridge, the lack of obvious thermal anomalies in the vicinity of the samples' collection sites and the relative proximity to the surface sediments makes us suggest that the

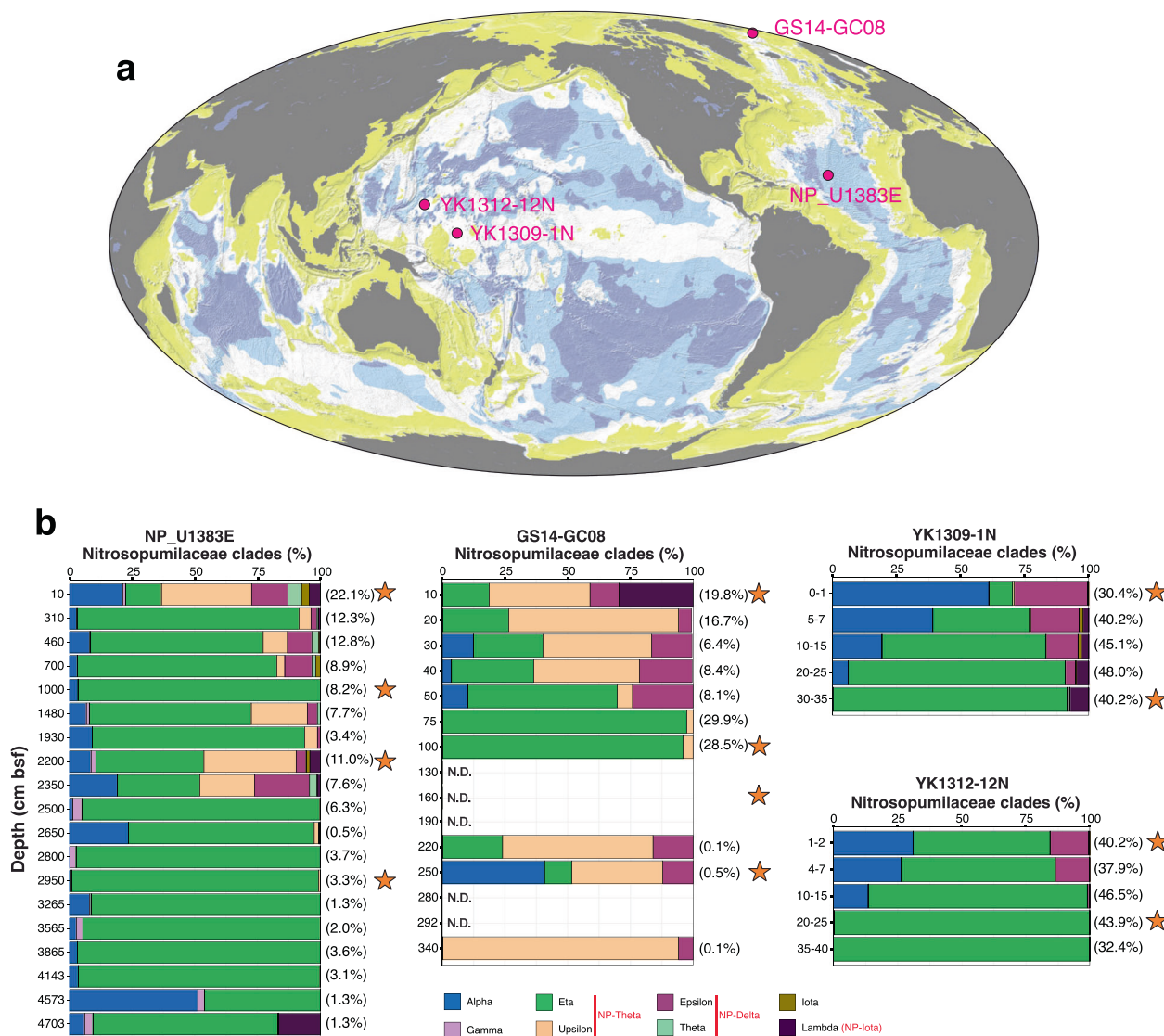


Fig. 1 Study sites and the community structure of AOA. a Global bathymetric map showing the coring locations of the sediment cores used in this study, modified from reference [19]. The dark blue and light blue regions represent the minimum and maximum areas over which dissolved O₂ is expected to penetrate throughout the sediment from sea floor to basement. **b** AOA community structure based on the 16S rRNA gene phylogeny. *Nitrosopumilaceae* 16S rRNA gene OTUs

were classified based on their placements in the phylogenetic tree as described in Zhao et al. [14]. In the figure key, the corresponding clades of the AOA *amoA* gene are shown in red. They are based on ref. [58] and our phylogenetic analyses (unpublished data). The sediment horizons selected for metagenome sequencing are highlighted by orange stars. Data for NP_U1383E and GS14-GC08 were retrieved from [12].

in situ temperatures are close to bottom surface temperatures (Table 1, [41]). However, the depth from where the samples in the Atlantic were taken and the thermal gradient at this location suggest that these are 2 °C higher than the bottom seawater (conductivity 0.99 and heat flux 90 [42]).

DNA of all samples was prepared using standard techniques and was sequenced on Illumina HiSeq2500; 16S rRNA gene amplicons were generated and sequenced using standard procedures. Detailed information about sampling sites, sampling procedure, and geochemical analyses are also shown in Supplementary Material.

Assembly and comparative genomics

All sequencing data were processed to remove illumina adapters and low-quality reads using Trimmomatic [43] before de novo assembly using MEGAHIT [44] (k-mer length of 27–117). Binning of contigs of Pacific metagenomes was performed applying a contig dereplication and binning optimization tool [45] based on the binning output of CONCOCT [45], MetaBAT [46], and MaxBin2 [47], while contigs of Atlantic samples were binned with MaxBin2 [47] followed by a sequence of refinement steps for

Table 1 Statistics of deep marine sediment-derived MAGs.

Genome bin	AOA clade	Depth (m)	Depth (m, below sea floor)	Contigs (No.)	Bottom seawater temperature (°C)	Estimated sediment temperature (°C)	Genome size (bp)	Proteins (No.)	Longest contig (bp)	N50	GC (%)	Completeness (%)	Contamination (%)	Presence/absence of 16S rRNA gene
NPMR_NP_delta_1*	NP-delta	4425	0.1	83	2.5	2.51	1,173,997	1467	79,876	28,788	34.0	86.2	5.4	-
NPMR_NP_delta_2	NP-delta	4425	0.1	270	2.5	2.51	610,368	870	12,327	3054	33.9	67.9	0.7	-
NPMR_NP_delta_3	NP-delta	4425	0.1	143	2.5	2.51	937,620	1222	58,541	15,484	34.2	66.3	3.4	-
NPMR_NP_theta_1*	NP-theta	4425	1	176	2.5	2.59	1,524,500	2000	40,454	14,847	35	99.5	2.9	-
NPMR_NP_theta_2*	NP-theta	4425	22	298	2.5	4.48	1,201,755	1671	26,499	5123	35	93.7	4	-
YK1312_12N_NP_theta*	NP-theta	5920	0.2	299	1.4	1.41	1,221,321	1695	20,455	5469	34	88.8	5.3	-
NPMR_NP_theta_3*	NP-theta	2476	0.1	552	-0.7	-0.69	1,069,514	1706	14,642	2738	34.4	83.6	3.9	+
NPMR_NP_theta_4	NP-theta	2476	0.1	232	-0.7	-0.69	1,040,901	1408	21,324	6332	34	76.2	6.3	-
NPMR_NP_theta_5	NP-theta	4425	0.1	274	2.5	2.51	697,980	970	14,641	3916	35.1	75	1.2	-
NPMR_NP_iota_1*	NP-iota	2476	0.1	68	-0.7	-0.69	1,121,898	1402	75,887	30,338	35.8	98.5	1	-
YK1309_IN_NP_iota*	NP-iota	4277	0.3	227	1.3	1.32	1,298,549	1700	40,725	9447	35.8	96.6	3.2	-

High-quality MAGs as defined for this dataset (>80% completeness and ~5% contamination or below, see “Results and discussion”) are marked with an asterisk symbol. Sediment in situ temperatures were estimated based on the bottom seawater temperatures and an assumed constant thermal gradient. At North Pond, a gradient of 0.09 °C/m was used [42]. For the other sites, the median (53 °C/km) of global sea floor temperature thermal gradients was used (compiled in Malinverno and Martinez [41]).

the thaumarchaeal bins (see Supplementary Methods). Completeness and contamination of Pacific and Atlantic bins were evaluated with CheckM (“lineage_wf” parameter) [48]. Assemblies are available on DDBJ/ENA/GenBank (accession numbers JADNRU000000000, JAEHKP000000000, JAEHKP000000000, JAEHKU000000000, JAEHKV000000000, JAEHKQ000000000, JAEHKR000000000 JAEHKS000000000, JAEHKT000000000) and in the Microbial Genome Annotation & Analysis Platform Microscope.

A dataset consisting of 163,852 predicted proteins from 85 genomes (11 MAGs reported here, 31 complete and near-complete AOA genomes and 43 MAGs or single-amplified genomes (SAGs) from NCBI or IMG) was collected for this study (Table S1). Annotation of the MAGs assembled in this study was performed automatically using the Microscope annotation platform from Genoscope [49], followed by extensive manual curation. NCBI annotations were supplemented with arCOG assignments from the archaeal Clusters of Orthologous Genes database (2018 release) [50] using COGsoft [51] (*e*-value of 10^{-10}). We clustered the protein dataset into protein families based on sequence identity (35%) and alignment coverage (70%) using CD-Hit V4.8.1 [52] (“-c 0.35 -aL 0.7 -aS 0.7”) (Table S2).

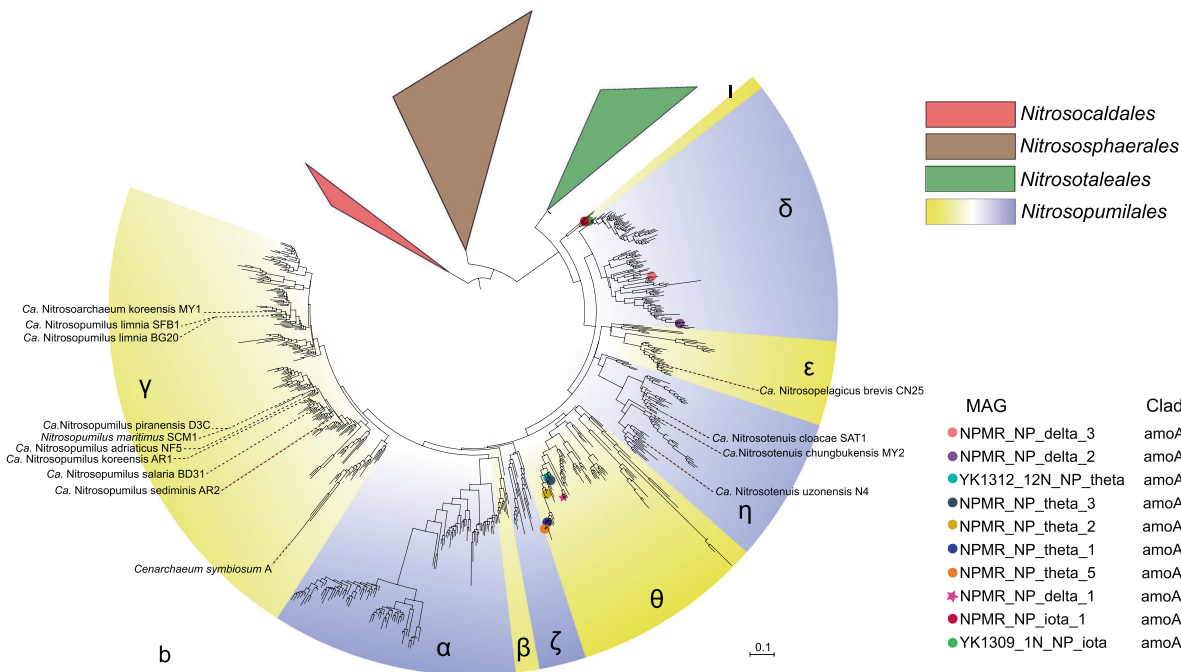
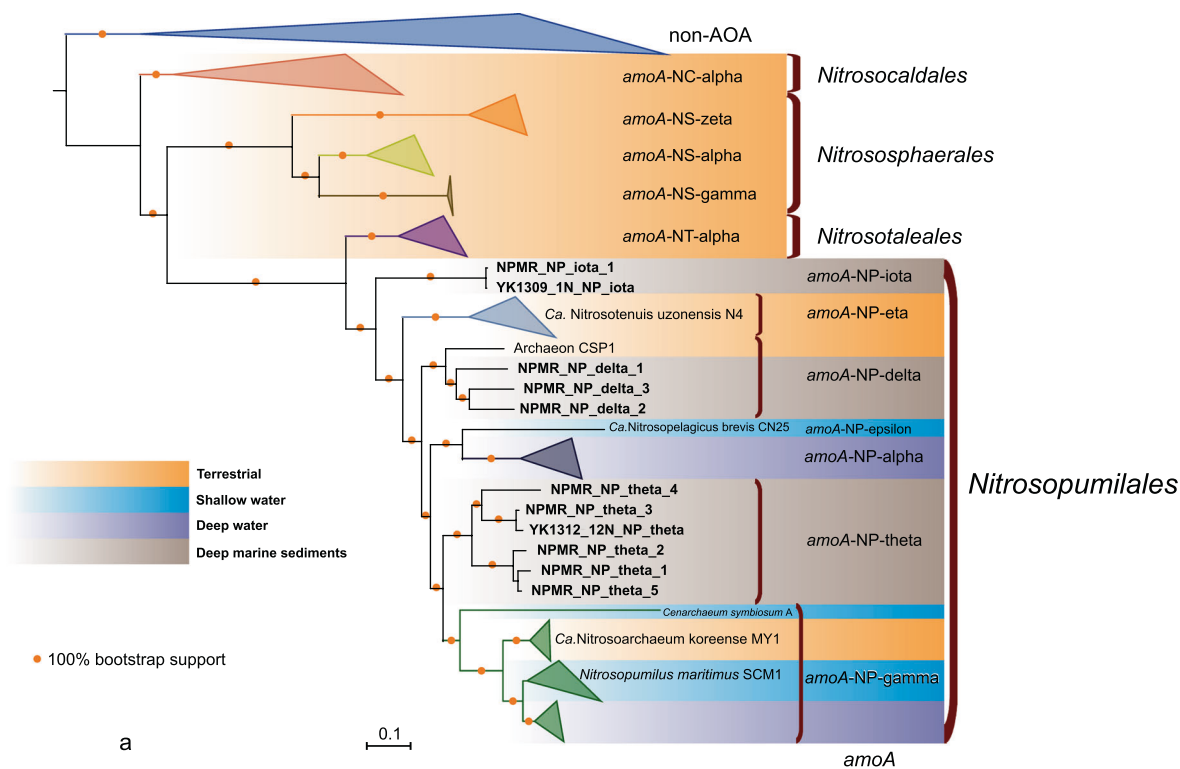
Selection of markers and phylogenomic tree

The identification of markers to perform the phylogenomic tree reconstruction was based on the phylogenomic workflow proposed by [53] (*e*-value 10^{-10}) using the archaeal single-copy gene collection [54]. We selected 79 markers (Table S3), present in at least 70 of the 85 genomes used in this study. Each protein family was aligned using MAFFT v7 (“-maxiterate 1000 -localpair”) [55] and trimmed with BMGE [56]. The concatenated alignment was used to reconstruct a maximum likelihood (ML) phylogenomic tree in IQTREE (v2.0-rc1) [57] under the LG+C20+F+G model with 1000 ultrafast bootstrap replicates. For *amoA* phylogeny and detailed methodological procedures, see Supplementary Material.

Results and discussion

Distribution of AOA in deep marine sediments

We examined the overall community structure of AOA (all affiliated to the family *Nitrosopumilaceae*) in these sediments by analyzing 16S rRNA gene amplicon sequencing data generated in this study for the Pacific cores and previously described for the Atlantic cores [12]. AOA communities in sediment horizons deeper than 10 cm were all



dominated by the so-called 16S-NP-eta and/or 16S-NP-epsilon and 16S-NP-alpha clades [58], which together correspond to the *amoA*-NP-theta clade (Figs. 1b and S2a) [28]. In addition, AOA affiliated to the 16S-NP-epsilon clade (corresponding to the *amoA*-NP-delta clade, Figs. 2a

and S2a) were also repeatedly detected with percentages <25% in the upper portions of these cores (Fig. 1b). Finally, the 16S-NP-lambda clade (now renamed to *amoA*-NP-iota, see below) was also detected as a minor clade in all cores except YK1312-12N, but was notably abundant in the

◀ **Fig. 2 Phylogenomic and phylogenetic trees of AOA and non-AOA Thaumarchaeota.** **a** Phylogenomic tree of AOA and non-AOA Thaumarchaeota. The phylogenomic tree was reconstructed based on the concatenated alignment of 79 markers comprising 17,483 amino acid sites using a maximum likelihood approach (see Materials and methods). Yellow circles represent 100% bootstrap support of nodes. The metagenome-assembled genomes (MAGs) reported in this study are shown in bold. Representative genomes of NP subclades were added for subclade clarification. Information about the ecological distribution of AOA clades is provided by colored shading. The scale bar indicates the number of substitutions per amino acid site. **b** Phylogeny of *amoA* sequences. The ML phylogenetic tree was reconstructed using 584 nucleotide sites. Colored circles represent MAGs from this study. The star symbol represents the NPMR_NP_delta_1 bin. The incongruity of its phylogenetic placement based on the *amoA* gene and the phylogenomic analysis of concatenated markers is discussed in the text. Greek letters represent the *amoA*-based annotation of AOA subclades as in Alves et al. [28]. The scale bar indicates the number of substitutions per nucleotide site.

uppermost horizon of GS14-GC08 (29% of the total AOA community, Fig. 1b).

Phylogenomic analysis and taxonomic placement reveal MAGs from three independent *amoA* clades

We obtained a total of 11 AOA MAGs, 9 from Atlantic and 2 from Pacific sediment samples (sequenced horizons are marked by stars in Fig. 1b). Despite high sequencing depths and high AOA abundances in the samples, based on 16S rRNA gene reads in the metagenomes (10–16% in the Pacific cores and 6.8–18.8% in the Atlantic cores), generation of good-quality bins was extremely challenging, possibly due to high microdiversity (41 OTUs) within NP as observed earlier [12]. Eventually, we obtained four MAGs with >90% completeness and three MAGs with >80% completeness, all with contamination levels ≤5%, which we consider high-quality MAGs in this study, as well as four additional medium-quality MAGs (66–76% completeness, up to 6.3% contamination level) (Table 1). The MAGs genome sizes (0.61–1.52 Mb) and GC contents (34.66 ± 0.72%) are in accordance with previous reports of free-living NP [59].

In order to study the evolution of AOA and place our deep marine sediment-derived MAGs in a phylogenetic context, we reconstructed a ML phylogenomic tree (Figs. 2a and S2a) using 79 concatenated single-copy markers from our entire dataset of 85 complete genomes, MAGs and SAGs representing a broad diversity of habitats (Table S1).

In addition, we performed an *amoA*-based phylogeny as in ref. [28] in order to assign a taxonomical rank and a respective AOA clade to our MAGs (Fig. 2b). Both trees showed similar clustering of MAGs into NP subclades except for NPMR_NP_delta_1 (see discussion in Supplementary Information) which based on the *amoA* tree clustered within the *amoA*-NP-theta clade but the more robust phylogenomic

analysis strongly suggests that it belongs to the *amoA*-NP-delta subclade. The only 16S rRNA gene recovered in MAG NPMR_NP_theta_3 is affiliated with a subclade of 16S-NP-alpha exclusively found in marine sediments (not shown).

Our phylogenomic tree revealed that the 11 AOA MAGs reported here represent the dominant AOA observed in our study (Fig. 1b) and form three well supported monophyletic groups, of which no cultured representative has been reported yet (Fig. 2a). Six MAGs represent the first genomic assemblies from the *amoA*-NP-theta lineage, one of the most dominant AOA groups in marine sediments and also found to be abundant in the crust below [14, 28]. Three MAGs are affiliated to *amoA*-NP-delta, the second most abundant AOA clade in marine sediments, and are the first marine sediment representatives of this clade, which includes a single other MAG (archaeon CSP1) assembled from river aquifer sediments [60].

Two bins (NPMR_NP_iota_1 and YK1309_NP_iota) clustered together forming a third sediment-dwelling clade, sister to all NP, which earlier escaped taxonomic assignment as it was only identified based on singular *amoA* sequences and hence had been designated *incertae sedis* [28]. Pairwise average nucleotide identity (ANI) comparisons (Fig. S2b) indicate that these two bins share >70% ANI with the other NP-MAGs recovered in this study (*amoA*-NP-delta and *amoA*-NP-theta MAGs sharing 73–79% ANI). A comparison of conserved protein families among all NP subclades indicated that this group harbors 320 out of 336 protein families that seem to be part of the NP core proteome, as opposed to only 260 of these protein families being present in the *Ca. Nitrosotaleales*, the sister lineage to all NP (Fig. 3). Moreover, environmental *amoA* sequences suggested that this clade might be restricted to deep sea sediments, an ecological specialization only found in NP. Taken together, this early-branching clade seems to be a new NP subclade and we propose the designation *amoA*-NP-iota (as it forms the ninth NP-clade following the taxonomy of [28]).

Three independent radiations of AOA into marine sediments

While MAGs and SAGs of AOA from bathypelagic (1000–4000 m), abyssopelagic (4000–6000 m), and hadopelagic (6000–11000 m) environments have been reported previously [23, 24, 61] and shotgun metagenomic analyses of deep sea sediments have been performed [10, 62, 63], the MAGs reported in this study represent to our knowledge the first high-quality AOA genomes from bathyal and abyssal sediments (>2000 m depth). Together with our taxon-enriched phylogenomic analyses they shed new light on the ecological transitions and niche differentiation undergone by AOA.

While our phylogenomic tree supports that AOA likely appeared in terrestrial habitats first (Fig. 2a), the sequence

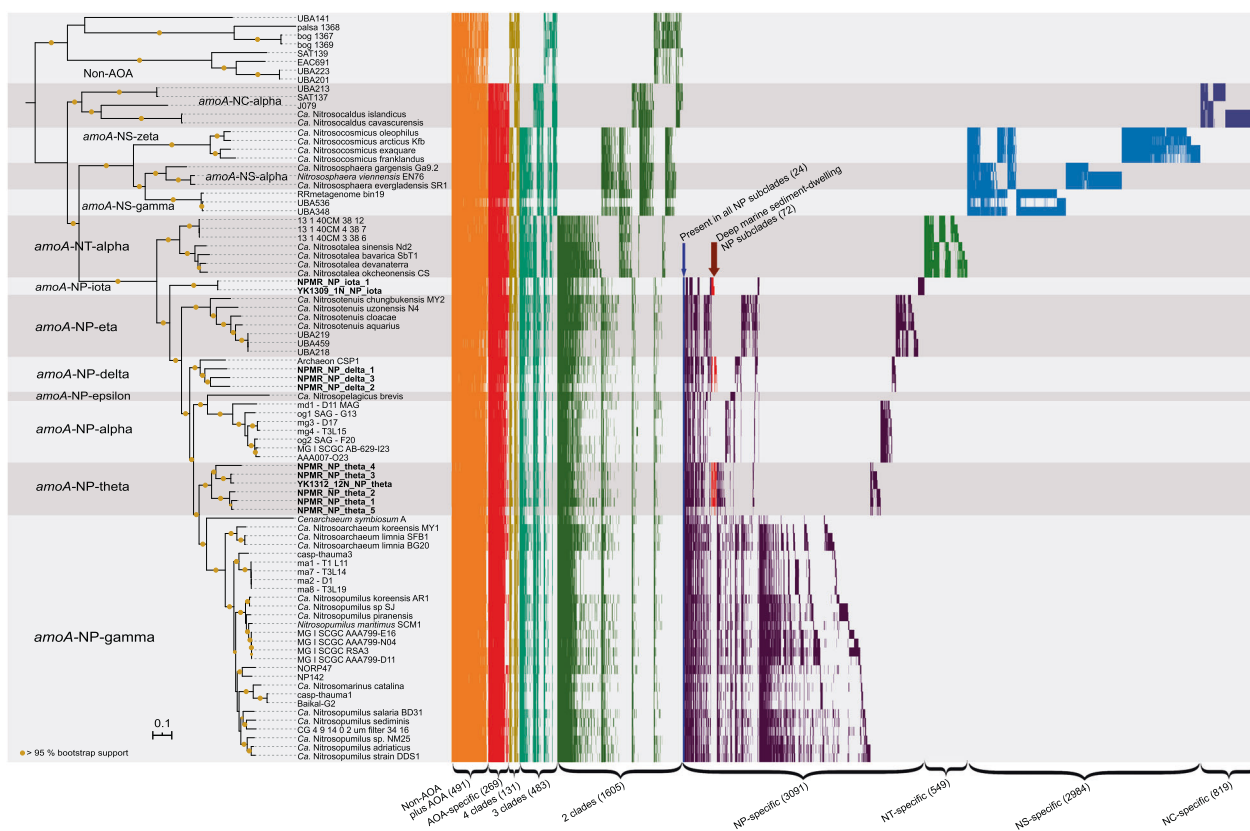


Fig. 3 Comparative analysis of the presence/absence of protein clusters among AOA. Each bar (x axis) represents the presence of a putative protein family in a genome (y axis). 10,422 clusters found in at least two different genomes are depicted. Non-AOA-specific clusters were excluded from the visualization. Yellow circles represent >95% bootstrap support of nodes. Clusters are ordered based on their

distribution pattern from the most widespread to the most uncommon: first across all lineages and then intralineage. Groups of clusters have different color codes for better visualization. NP Nitrosopumilales, NC *Ca.* Nitrosocaldades, NS *Nitrososphaerales*, NT *Ca.* Nitrosotaleales.

and number of colonization events of marine environments seem to be more complex than previously proposed [29]. Importantly, our results suggest that both the deep-water adapted AOA as well as the deep sediment-adapted AOA are polyphyletic. The colonization of deep waters, i.e., pelagic organisms, might have occurred independently at least twice in the evolution of the NP, at the origin of *amoA*-NP-alpha clade and during the diversification of *amoA*-NP-gamma (Fig. 2a). Interestingly, the *amoA*-NP-gamma clade, which is one of the most diverse NP subclades [28], has undergone particular habitat transitions and niche occupation [28]. Distinct shallow water *amoA*-NP-gamma species have established independently symbiotic associations with sponges [59, 64] while the sublineage leading to the soil isolate *Nitrosarchaeum koreense* [65] could have evolved from an estuarine or shallow water ancestor suggesting a recolonization of land (Fig. 2a).

Regarding the origin of deep sediment-dwelling AOA, the *amoA*-NP-theta lineage branches within mostly marine NP clades (Fig. 2a), suggesting that this lineage might have evolved from a pelagic marine ancestor. However, there is

no evolutionary link between any pelagic AOA and the NP-delta clade, as the latter have been mostly retrieved in estuarine and deep marine sediments [60]. The most parsimonious evolutionary scenario would be that this group underwent a direct transition from estuarine sediments to marine sediments during its diversification.

Similarly, the newly proposed clade NP-iota, the earliest branching NP which has so far exclusively been detected in marine sediments [28], does not seem to be closely related to pelagic NP but emerges instead among terrestrial clades (*Ca.* Nitrosotaleales and NP-Eta). Although it is possible that pelagic AOA closely related to *amoA*-NP-iota may be detected in further environmental surveys or that respective pelagic lineages got extinct, the *amoA*-NP-iota clade might as well have developed from terrestrial-estuarine organisms, as discussed for *amoA*-NP-delta above.

Comparative genomics of deep sea sediment AOA

We constructed a total of 33,442 protein families from our taxon-enriched genome dataset representing a wide

variety of ecological environments (see Materials and methods and Table S1). From these, 12,137 have representatives from at least two different genomes. In our analysis, the AOA core proteome comprises 760 protein families present in at least one genome of each of the four major AOA lineages: *Ca. Nitrosocaldales*, *Nitrososphaerales*, *Ca. Nitrosotaleales*, and NP (Fig. 3, Table S2). Thus, our results are similar to previous estimations of the AOA core genome (743 gene families) [66], and slightly lower than our own earlier estimate of 860 gene families (based on only seven genomes [67]). Only 269 of the core AOA families seem to be AOA-specific (Fig. 3, Table S2). Only 123 out of these 269 families were found to be present in >50% of the genomes in each of the four AOA orders (a relatively low threshold to account for the incompleteness of MAGs), suggesting a relatively low degree of conservation within these lineages. These results imply great intraorder genomic variability and important differential gene loss among subclades and across genomes during the evolution and diversification of AOA. For instance, despite the fact that NP have 3091 specific families with proteins encoded in at least two genomes and present in one or more NP subclades, a subset of solely 24 families were conserved in all seven NP subclades (Fig. 3). Considering the very relaxed criteria used, this is a surprisingly small number of conserved families in all seven NP subclades.

To identify possible specific adaptations of AOA to deep marine sediments, we searched for families present in at least two of the three marine sediments clades represented by our 11 MAGs (i.e., *amoA*-NP-theta, -delta, and -iota), to the exclusion of all the other genomes analyzed in this study (Fig. S2, Table S1). A total of 72 families were identified (Fig. 3), of which only 25% (18 families) could be functionally annotated (Tables S2 and S4) and were classified into the following categories: information processing systems (7), metabolism (5), and cellular processes (6). Some of these 18 families had functional equivalents in most if not all AOA (e.g., RadA homologs). We additionally found 41 families shared predominantly between NP subclades with deep ocean (>1000 m) representatives (i.e., *amoA*-NP-alpha, NP-gamma sublineages recovered from the Mariana, Izu-Ogasawara Trenches and the Red Sea [23, 61], NP-theta, NP-iota, and NP-delta) to the exclusion of all other NP subclades. From these 41 families, 18 have functional annotation: information processing systems (8), metabolism (4), and cellular processes (6). Families with functional significance specific to marine sediments, such as a putative lactate racemase, or those shared with deep ocean MAGs (Figs. 3 and 4), are discussed below. Families identified in deep ocean MAGs but not found in the sediment clades are still depicted in Figs. 4 and S3 for comparative purposes. In addition, we investigated the number of clusters shared between deep sediment-derived MAGs

and the terrestrial (present in soils and sediments) lineage *Nitrososphaerales*, to the exclusion of all other AOA lineages and NP subclades. Interestingly, they share only one protein family, related to coenzyme F₄₂₀-dependent luciferase-like oxidoreductases.

Metabolic reconstruction of the *amoA*-NP-theta, *amoA*-NP-delta, and *amoA*-NP-iota clades

The full annotations for all genes and pathways discussed in the following section can be found in Table S4.

Central energy and carbon metabolism

All three sediment clades (i.e., *amoA*-NP-delta, *amoA*-NP-theta, *amoA*-NP-iota) encode complete sets of genes involved in ammonia oxidation, namely *amoA*XCB in the typical organization observed in other NP (Fig. 5, Table S4) [21, 68, 69]. Missing subunits in certain MAGs seem to be due to genome incompleteness. A nitrite reductase (NirK) homolog is present, as well as multiple blue copper domain proteins putatively functioning as electron carriers. All clades encode a single high-affinity ammonia transporter family protein (Amt), as opposed to two Amt transporters of differing affinities found in other AOA. This could represent an adaptation to an oligotrophic environment [70]. Four out of six NP-theta MAGs and all *amoA*-NP-delta MAGs encode complete or near-complete urease operons (Fig. 4, Table S4). Together with a putative nitrilase (Nit1, conserved in AOA) and a putative omega-amidase (Nit2, present in *amoA*-NP-theta, -delta, -eta, -gamma), these genes indicate expanded substrate utilization capabilities for ammonia (and CO₂) generation by cleaving urea, nitriles, and dicarboxylic acid monoamides. Utilization of organic nitrogen compounds is a feature shared with other NP clades that include deep sea lineages and previously described for subseafloor AOA (Figs. 4, 5, and S3) [10, 23, 24, 34].

All three sediment clades encode full gene sets for electron transfer to O₂ via NADH dehydrogenase (complex I), type bc₁ complex III, and a heme-copper terminal oxidase (complex IV) (Fig. 5, Table S4). No alternative complexes using a different electron acceptor were identified.

All three sediment clades encode the full repertoire conserved among AOA for autotrophic carbon fixation via the 3-hydroxypropionate/4-hydroxybutyrate (3HP/4HB) cycle and carbon metabolism through oxidative TCA and gluconeogenesis up to the formation of glucose-6P via a phosphoglucose isomerase (not present in NS) homolog in NP-theta and NP-delta (Fig. 5, Table S4) [32, 68, 71]. A malic enzyme, enabling the formation of pyruvate from malate with the concomitant generation of NAD(P)H, expands metabolic capacities in NP-theta and NP-iota (also

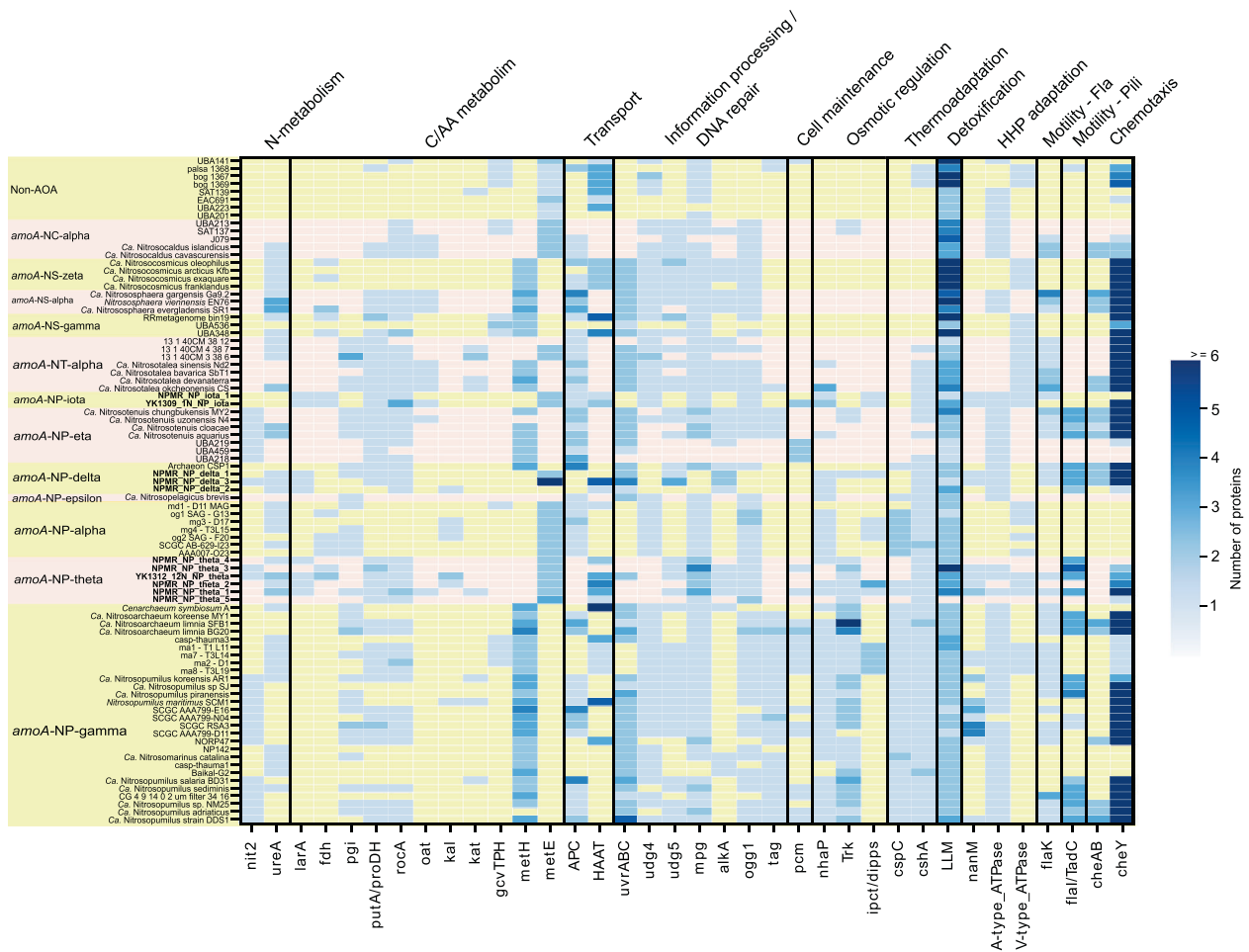


Fig. 4 Heatmap depicting the distribution and abundance of genes involved in the main functional categories discussed in the text. *nit2* nitrilase/omega-amidase, *ureA* urease subunit gamma, *fdh* formate dehydrogenase, *larA* lactate racemase, *pgi* phosphoglucose isomerase, *proDH* proline dehydrogenase, *rocA* 1-pyrroline-5-carboxylate dehydrogenase, *oat* putative ornithine--oxo-glutarate aminotransferase/class III aminotransferase, *kal* 3-aminobutyryl-CoA ammonia lyase, *kat* putative 3-aminobutyryl-CoA aminotransferase, *gvtTPH* glycine cleavage system proteins T/P/H, *metH* methionine synthase II (cobalamin-independent), *metE* methionine synthase I (cobalamin-dependent), *APC* amino acid--polyamine--organocation transporter family, *HAAT* the hydrophobic amino acid uptake transporter (HAAT) family, *uvrABC* the Uvr excision repair system endonucleases ABC, *udg4/5* uracil DNA glycosylase family 4/5, *mpg* methylpurine/

alkyladenine-DNA glycosylase, *ogg1* 8-oxoguanine DNA glycosylase, *alkA* DNA-3-methyladenine glycosylase, *tag* 3-methyladenine DNA glycosylase, *pcm* protein-L-isoaspartate carboxylmethyltransferase, *nhaP* the monovalent cation:proton antiporter-1 (CPA1) family, *Trk* the K+ transporter (Trk) family, *ipct/dipps* bifunctional CTP:inositol-1-phosphate cytidyltransferase/di-myo-inositol-1,3'-phosphate-1'-phosphate synthase, *cspC* cold-shock protein A, *csnA* cold-shock DEAD-box protein A, *LLM* luciferase-like monooxygenase family protein, *nanM* N-acetylneuraminic acid mutarotase, *flaK* archaeal preflagellin peptidase *FlaK*, *cheY* chemotaxis response regulator *CheY*, *cheAB* chemotactic sensor histidine kinase *cheA* and methyl-esterase *cheB*. All locus tags and cluster information are given in Supplementary Tables 2 and 4. An extended version of the heatmap is shown in Fig S3.

in some other AOA, Fig. 5). As with most AOA, all sediment clades have the capacity to synthesize polyhydroxybutyrate storage compounds, an obvious advantage in an oligotrophic environment [72].

Complete or near-complete amino acid biosynthesis pathways as well as vitamins (including vitamin B12) are present in all three sediment clades, as in other AOA (Table S4). As observed in NP-alpha representatives [23], the sediment clades use the B12-independent pathway for methionine biosynthesis (*metE*) (Fig. 4). Albeit this being a less catalytically efficient enzyme than the B12-dependent

metH present in all other AOA, it is nevertheless much less costly energetically [73], and would therefore be an advantage in an energy-limiting environment where maintenance rather than fast growth is the norm [74].

Utilization of exogenous organic compounds

All three sediment lineages seem to be capable of utilizing exogenous organic compounds from fermentation processes such as formate, lactate, and 3-aminobutyryl-CoA as a source of carbon, nitrogen, and reductive

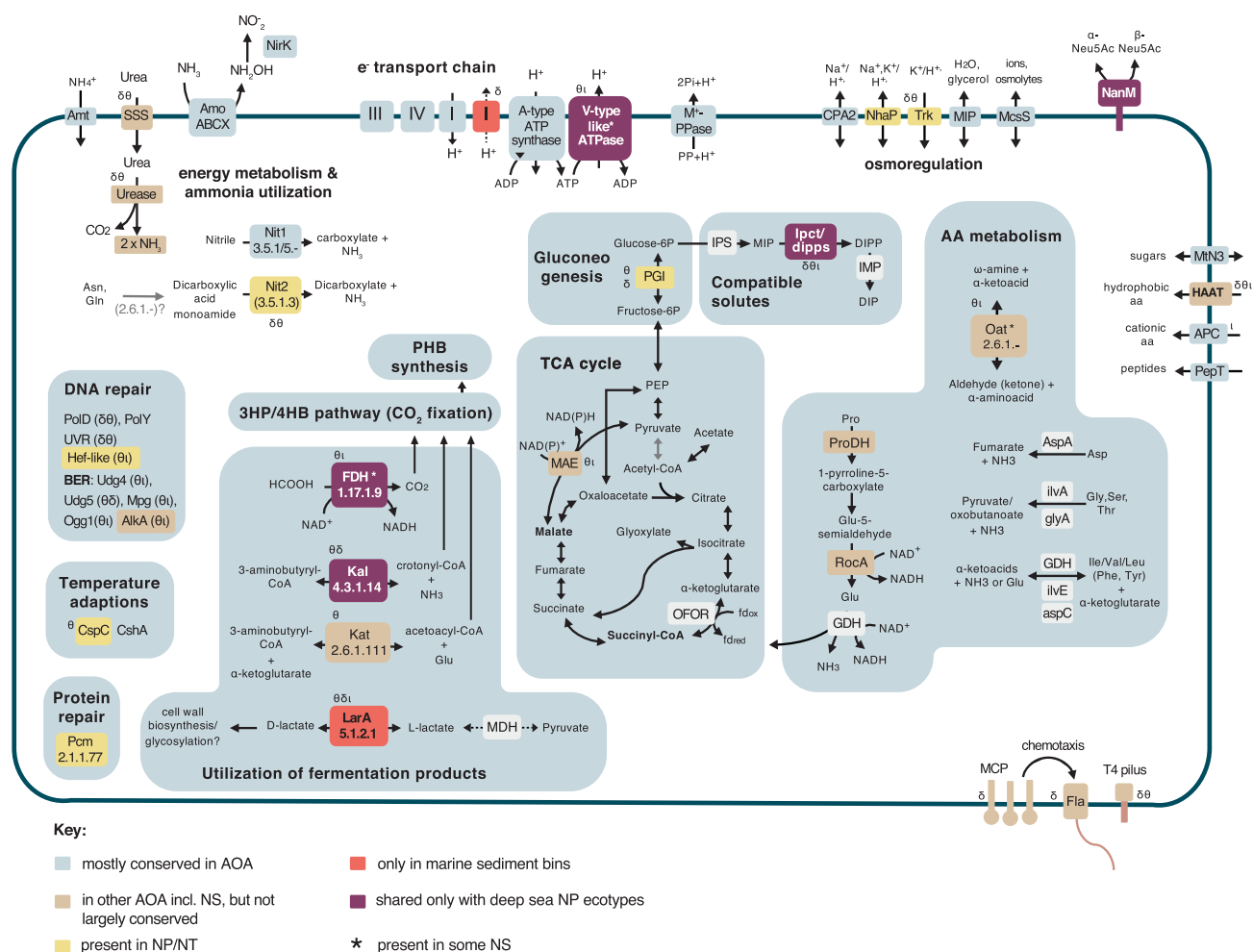


Fig. 5 Metabolic reconstruction of *amoA*-NP-theta, *amoA*-NP-delta, and *amoA*-NP-iota AOA. Schematic reconstruction of the predicted metabolic modules in the sediment MAGs, as discussed in the text. Color code of enzymes/complexes indicates conservation status in AOA. Unless specified by greek letters ($\theta\delta\iota$), enzymes/modules are present in all sediment clades. Dashed lines indicate hypothetical reactions. The gray arrow indicates an alternative OFOR reaction. Complexes of the electron transport chain are labeled with roman numerals. Transporters are named according to TCDB classification. Enzymes, gene accession numbers, and transporter classes are also listed in Supplementary Tables 2 and 4. Amo ammonia monooxygenase, NirK nitrite reductase, Nit1 nitrilase, Nit2 nitrilase/omega-amidase, AA amino acid, Fdh formate dehydrogenase, Kal 3-aminobutyryl-CoA ammonia lyase, Kat putative 3-aminobutyryl-CoA aminotransferase, Lar lactate racemase, Pcm protein-L-isoaspartate carboxylmethyltransferase, MCP methyl-accepting chemotaxis protein, Fla archaeallum, PolD polymerase family D, PolY translesion polymerase family Y, UVR excision repair system, Hef-like Hef/FANCM/Mph1-like helicase, BER base-excision

repair, Udg4/5 Uracil DNA glycosylase family 4/5, Mpg methylpurine/alkyladenine-DNA glycosylase, Ogg1 8-oxoguanine DNA glycosylase, AlkA DNA-3-methyladenine glycosylase, CspA cold-shock protein A, CshA cold-shock DEAD-box protein A, Pcm protein-L-isoaspartate carboxylmethyltransferase, PHB polyhydroxybutyrate, MAE malic enzyme, OFOR 2-oxoacid:ferredoxin oxidoreductase, PGI phosphoglucose isomerase, IPS myo-inositol-1-phosphate synthase, ipct/dipps bifunctional CTP:inositol-1-phosphate cytidylyltransferase/di-myo-inositol-1,3'-phosphate-1'-phosphate synthase, IMP DIPP phosphatase, ProDH proline dehydrogenase, RocA 1-pyrroline-5-carboxylate dehydrogenase, glutamate dehydrogenase, oat putative ornithine--oxo-glutarate aminotransferase/class III aminotransferase, aspA aspartate ammonia lyase, ilvA threonine/serine ammonia lyase, glyA serine/glycine hydroxymethyltransferase, ilvE branched-chain-amino acid transaminase, aspC aspartate/tyrosine/aromatic aminotransferase, MCO1 multicopper oxidase family 1, NanM N-acetylneuraminic acid mutarotase. Transporters are named according to TCDB classification (Supplementary Table 4).

potential. This finding expands the range of organic carbon and nitrogen substrates suggested earlier for deep ocean AOA (previously comprising amino acids, peptides, and compatible solutes) and reinforces their role as key players in nutrient cycling in these biomes [6, 10, 23, 24, 27, 34].

A putative soluble NAD^+ -dependent formate dehydrogenase (Fdh), distinct from the iron-sulfur/molybdenum containing Fdh enzymes traditionally found as part of formate-hydrogen lyase systems [75], is found in the NP-theta and NP-iota clades (as well as in certain NS representatives and NP-alpha Figs. 4 and 5). However, no

additional hydrogenases or known formate transport systems were identified in the marine sediment bins. Our phylogenetic analysis (Fig. S4) indicates that the enzyme is a bona fide NAD⁺-dependent Fdh within the superfamily of D-2-hydroxyacid dehydrogenases [76]. This indicates the capacity to use formate for supplementing CO₂ needs while concomitantly supplying reducing equivalents (as in methylotrophs [77, 78]).

A putative 3-aminobutyryl-CoA aminotransferase (Kat, EC 2.6.1.111) and a 3-aminobutyryl-CoA ammonia lyase (Kal, EC 4.3.1.14) were identified in the NP-theta and NP-delta bins, and are also found in some NP-alpha, NP-gamma and NT lineages (Figs. 4 and 5). These enzymes participate in lysine fermentation pathway variants in fermentative bacteria [79]. Although the key pathway enzymes are not present in the sediment bins or any other AOA, this intermediate compound (3-aminobutyryl-CoA) could be scavenged from fermenting microorganisms in the sediment community. Both enzymes can remove ammonia from 3-aminobutyryl-CoA either by transferring it to α -ketoglutarate resulting in the formation of acetoacetyl-CoA and glutamate (Kat) or by an elimination reaction that produces crotonyl-CoA and free ammonia (Kal). Both products are intermediates of the 3HB/4HP (CO₂ fixation-) pathway and could be processed accordingly, generating reducing potential in the subsequent steps.

The presence of a putative lactate racemase family protein (LarA), specific to the NP-theta, -delta, and -iota clades (Figs. 4, 5, and S3), suggests that lactate is another fermentation product that could be utilized by these lineages. This is one of the very few protein families with a putative function prediction shared specifically between the sediment AOA clades to the exclusion of all other AOA, suggesting an essential role. LarA in lactobacilli catalyzes the interconversion of D- and L-lactate, ensuring an adequate supply of D-lactate which is an important cell wall component conferring resistance to vancomycin [80] (see Supplementary Information). Given the importance of cell envelope maintenance in the adverse conditions of the sediments, it is possible that D-lactate has a similar use in sediment AOA, conferring resistance to exogenous toxic compounds. Alternatively, the lactate dehydrogenase-like malate dehydrogenase homologs found in AOA possess features indicating that they could have a broad substrate specificity, being able to utilize pyruvate in addition to oxaloacetate, and producing the L-stereoisomers of the products (see Supplementary Information and Figs. S5 and S6), with the concomitant reduction of NAD⁺ [81].

As mentioned above, the only protein family specifically shared among the sediment MAGs and the terrestrial NS lineages is an F₄₂₀-dependent luciferase-like oxidoreductase [82]. While the metabolic role of these proteins in AOA in general is still unclear, the ability to degrade recalcitrant carbon via oxygenases in a manner similar to terrestrial

organisms [83] has been observed in sediment and crust communities [63, 72], and is proposed to provide an opportunistic advantage for expanded substrate utilization in limiting conditions.

Adaptations to low energy and high-pressure environments

Deep sea sedimentary environments found under the oligotrophic ocean present manifold challenges to microbial life, namely energy limitation, high hydrostatic pressure (HHP), low temperatures (<4 °C), and potential microoxic or anoxic conditions detrimental to aerobic metabolisms [16, 18, 19, 74, 84–86]. Microorganisms respond with global metabolic changes rather than stress responses [87], some of which are found in the deep sediment AOA clades.

Many organisms possess distinct electron transport, ion gradient generating and ATP synthase complexes that are differentially regulated under HHP [74, 88–90]. Interestingly, both NP-iota MAGs and two out of four high-quality NP-theta MAGs encode complete gene clusters for both the A-type ATPase found in neutrophilic AOA and V-type ATPase variant found in acidophilic/acidotolerant/piezotolerant archaea (and AOA) which is homologous to the proton/ion pumping ATPases from eukaryotes and enterococci (Figs. 4 and 5) [91, 92]. The remaining high-quality NP-theta MAGs encode either the A-type (NPMR_NP_theta_3, which encodes additionally a V-type *atpI* on the edge of a contig) or the V-type ATPase (NPMR_NP_theta_2), suggesting the general presence of the V-type ATPase type in this clade. None of the ATPase gene clusters were detected in the remaining two lower-quality NP-theta MAGs (NPMR_NP_theta_4 and theta_5), most probably due to their incompleteness and fragmentation levels.

The V-type ATPase has been suggested to confer physiological advantages in high-pressure environments by virtue of its proton-pumping function [91]. This would enable the maintenance of intracellular pH, which is disrupted by the accelerated release of protons from weak acids (such as carbonic acid) under HHP [93]. The presence of both ATPase variants is also observed in abyssal/hadopelagic NP-gamma AOA lineages, while the deep marine NP-alpha encode only the V-type ATPase (Figs. 4 and 5 and Supplementary Information for further discussion) [91]. In contrast, all three NP-delta MAGs encode only the canonical A-type ATPase (Fig. 5), but intriguingly at least two of them seem to contain a partially duplicated NADH dehydrogenase (complex I) operon which could similarly be responsible for alleviating cytoplasm acidification (see Supplementary Information).

The cytoplasmic membrane is severely affected by HHP, which induces a tighter packing of the lipids and a transition to a gel state, resulting in a decrease in fluidity and permeability [94, 95]. The presence of an N-acetylneuraminic

acid mutarotase in NP-theta, NP-iota, and NP-delta MAGs (Fig. 4, shared with a few abyssopelagic/hadal NP-gamma species) indicates the ability to acquire sialic acid [96] putatively as a component of the S-layer associated glycan. This important component of glycoconjugates found on cell walls has multiple functions including concentrating water on cell surfaces [97] and regulating membrane permeability [98]. It can also enable the regulation of the thickness of the hydration layer surrounding the cell membrane [99], which could prevent system volume change and stabilize membrane protein complexes and membrane structure under pressure [99, 100], while also regulating membrane permeability [98]. Modification of the hydration layer properties has also been identified as a specific adaptation mechanism of the piezophilic archaeon *Thermococcus barophilus* [101]. Genes associated with sialic acid synthesis (NeuA and NeuB) have been identified in certain NP [102], in our dataset though only NPMR_NP_theta_1 encoded the respective genes. This however implies the potential production and usage of sialic acid from AOA marine populations.

An ABC-type branched-chain amino acid transport system of the HAAT family (3.A.1.4) is present in all three sediment clades as well as in one NP-alpha MAG, sponge-associated and few other lineages of the NP-gamma clade, NS and non-AOA Thaumarchaea (Figs. 4 and 5, Tables S2 and S4). The uptake of amino acids has been interpreted earlier as indicative of the possibility of organic carbon utilization via enzymes participating in canonical amino acid biosynthesis pathways and present in all or most AOA (e.g., aspA, ilvA, ilvE, aspC glyA, GDH, ProDH) [10, 23, 38, 64]. Such mixotrophic strategies are also responsible for the enormous ecological success in oligotrophic environments of oceanic cyanobacterial lineages [103]. However, canonical amino acid degradation key enzymes, such as amino acid hydroxylases, the branched-chain α -keto acid dehydrogenase complex or 2-ketoacid:ferredoxin oxidoreductases, have not been detected in deep sea or sediment AOA clades, nor are their genomes particularly enriched in proteases (Fig. S3). On the other hand, a metabolic shift from expensive de novo biosynthesis of cellular materials (with proteins accounting for 56% of total energy investment in oxic environments) to recycling of exogenous or endogenous resources is observed in HHP-adapted microorganisms [16, 18, 72, 87, 88, 104, 105]. Therefore, it seems more likely that amino acids are used for recycling, as suggested earlier for AOA by isotope tracer and NanoSIMS experiments with sediment and oceanic crust communities [16, 33, 34, 72], and as observed in the piezophile *T. barophilus* and other facultative piezophiles [105–107]. Moreover, amino acids (mostly glutamate, proline, and glutamine) can be accumulated as compatible solutes to ensure the stabilization of macromolecular

structures upon pressure or temperature related stress [107, 108]. It cannot be ruled out though that amino acids are also used for replenishing the intracellular ammonia pool, with minimal production (if at all) of reducing equivalents (Fig. 5) (see Supplementary Information for detailed discussion).

It is worth noting that most of the AOA genomes were recovered from shallow sediments, whose estimated temperatures should not pronouncedly deviate from those of the bottom seawater (Table 1). Evidence for adaptation to low temperatures in the deep sediments is also given through the presence of homologs of the cold-shock protein CspC [109] in the NP-theta, NP-gamma, NP-alpha, and NP-epsilon clades, and the cold-shock DEAD-box protein A (CshA) in all NP (Figs. 4 and 5 and Tables S2 and S4), neither present in thermophilic archaea or bacteria. Both have been implicated in the cold-shock response in bacteria and archaea, enabling growth in temperatures within the psychrophilic range [110, 111]. Therefore, while we do not have data to suggest optimal growth conditions for the deep marine sediment clades, their high abundance, environmental distribution, previously inferred in situ growth and activity [10, 12, 28] enable us to infer that are mesophilic/psychrophilic or at least psychrotolerant (see additional discussion in Supplementary Information).

The sediment clades, especially NP-theta, encode an extended repertoire of enzymes for DNA and protein repair compared to other NP (details in Supplementary Information and Figs. 4, 5, and S3). This is an indication of energy investment toward maintenance of cellular components, rectifying damage due to low turnover rates and cellular aging rather than active and fast growth [74, 85, 112]. This strategy together with dormancy is presumed to be responsible for persistence in subseafloor energy limited environments [113].

Osmoregulation

All sediment clades encode a putative bifunctional CTP: inositol-1-phosphate cytidylyltransferase/di-myoinositol-1,3'-phosphate-1'-phosphate synthase (ipct/dipps), responsible for the synthesis of the compatible solute di-myoinositol-1,3'-phosphate [114]. The enzyme is also found in deep marine AOA clades [23] (Figs. 4 and 5). Biosynthetic genes for this compatible solute have so far only been observed in organisms growing above 55 °C, and have been extensively transferred between archaea and bacteria [115], making these AOA clades the first nonthermophilic organisms with the ability to synthesize this inositol derivative. Compatible solutes can confer resistance to various types of stress, so it is possible that this anionic solute has multiple roles in these polyextremophilic organisms [116], especially since no pathways for synthesis/uptake of known osmolytes such as

mannosylglycerate, ectoine/hydroxyectoine or glycine/betaine were identified in the NP-theta, NP-iota, and NP-delta MAGs (Figs. 4, 5, and S3).

Conclusions

Our comparative and phylogenomic analyses using 11 sediment-derived MAGs reported in this study, together with a large collection of AOA genomes with a broad phylogenetic and ecological distribution, allowed us to study the evolution, diversification, and adaptation mechanisms of AOA into deep marine environments. Based on phylogenomic analyses and different from earlier scenarios [29], we conclude that AOA from deep marine sediments evolved independently within (at least) three lineages. Although it seems that the ancestor of the *amoA*-NP-theta clade was pelagic and descendants of it occupied the deep marine sediments and the oxic seafloor crust, it is likely that in the case of *amoA*-NP-iota and *amoA*-NP-delta, there was a transition from terrestrial ecosystems/freshwater sediment to marine sediments without having colonized the ocean water column first. Interestingly, all extended capacities and adaptations discussed in this manuscript are found to be combined in lineage *amoA*-NP-theta, which represents the most widely distributed and abundant clade ranging over different marine sediment layers, whereas the other two clades that share some of these features exhibit a more distinct distribution pattern.

All AOA adapted to marine sediments and investigated in this study are able to perform ammonia oxidation in combination with CO₂ fixation like all other described AOA. In addition, all three lineages seem to be capable of utilizing exogenous organic fermentation products that they convert into intermediates of their central carbon metabolism, a feature they share with pelagic AOA from the deep ocean and a few other AOA. This, together with the capability of taking up amino acids, putatively for recycling into proteins or utilization of amine groups, would support growth in this extremely oligotrophic environment and contribute to organic nitrogen and carbon turnover in the sediments. In the absence of any components indicating increased capacity of amino acid degradation in these AOA, we argue that recycling of amino acids rather than catabolism as otherwise suggested in [10, 23, 38] represents an advantageous and more plausible strategy for the sedimentary AOA clades. It is also a trait frequently observed in other sedimentary and crustal population groups to overcome the prohibitive energetic costs of de novo monomer biosynthesis [72]. In addition, a broad repertoire of DNA and protein repair enzymes seems to enable the deep sediment-adapted AOA to counteract the most severe consequences of cellular aging. An important feature shared with HHP-adapted deep marine clades is the presence of two ATPase complexes

in *amoA*-NP-theta and *amoA*-NP-iota, with putatively opposing functions that would alleviate the effects of pH imbalance due to HHP, as well as the PMF-destabilizing effects of age-induced membrane leakage. These features shed light onto the mechanisms underlying AOA persistence in the benthic environments beneath the open ocean, from the surface sediments down to the underlying oceanic crust, and further consolidate the central role of these archaea in the global biogeochemical cycles.

Acknowledgements The authors thank Philipp Weber for help with visualizations. The LABGeM (CEA/Genoscope & CNRS UMR8030), the France Génomique and French Bioinformatics Institute national infrastructures (funded as part of Investissement d'Avenir program managed by Agence Nationale pour la Recherche, contracts ANR-10-INBS-09 and ANR-11-INBS-0013) are acknowledged for support within the MicroScope annotation platform. The authors thank onboard scientists, officers, and crews of RV Yokosuka and the manned submersible Shinkai 6500 for their help and operation during the YK1309 and YK1312 cruises. The authors wish to acknowledge the entire scientific party and all crew members onboard Joides Resolution during IODP expedition 336, for their help and expertise, especially co-chief scientists Katrina Edwards and Wolfgang Bach. The authors also wish to thank the chief scientist of the CGB summer cruise 2014, Rolf Berger Pedersen, for the sediment coring opportunity and thank Ingeborg Ørklund, Desiree Roerdink, Tamara Bumburger, and Ingunn H Thorseth for their help with porewater extraction and nutrient analysis. This project was funded by FWF grant 27017 and ERC AdvGr TACKLE (695192). SSA was funded by a Marie Curie Action fellowship (grant THAUMECO-PHYL 701981). SLJ was funded by the Trond Mohn starting grant BFS2017REK03 and the Census of Deep Life phase V grant. TN was partially supported by JSPS KAKENHI Grant Number JP19H05684 within JP19H05679 (Post-Koch Ecology).

Author contributions CS, SLJ, and TN conceived the study. RZ and SLJ sampled and processed the Atlantic sediments, while TN, HN, MH, and YT sampled and processed the Pacific sediments. RZ, SSA, and RIP-T assembled the MAGs. RIP-T performed the phylogenetic and comparative genomic analyses. MK annotated and analyzed the genomes. MK, RIP-T, RZ, SSA, SLJ, CS, and TN interpreted data. MK, RIP-T, and CS wrote the manuscript with contributions from TN, RZ, SLJ, and SSA.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Schleper C, Nicol GW. Ammonia-oxidising archaea—physiology, ecology and evolution. *Adv Microb Physiol.* 2010;57:1–41.
- Stahl DA, de la Torre JR. Physiology and diversity of ammonia-oxidizing archaea. *Annu Rev Microbiol.* 2012;66:83–101.
- Hatzenpichler R. Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. *Appl Environ Microbiol.* 2012;78:7501–10.
- Pester M, Schleper C, Wagner M. The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. *Curr Opin Microbiol.* 2011;14:300–6.
- Lehtovirta-Morley LE. Ammonia oxidation: ecology, physiology, biochemistry and why they must all come together. *FEMS Microbiol Lett.* 2018;365:fny058.
- Santoro AE, Richter RA, Dupont CL. Planktonic marine archaea. *Ann Rev Mar Sci.* 2019;11:131–58.
- 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. *Proc Natl Acad Sci USA.* 2005;102:14683–8.
- Durbin AM, Teske A. Archaea in organic-lean and organic-rich marine subsurface sediments: an environmental gradient reflected in distinct phylogenetic lineages. *Front Microbiol.* 2012;3:168.
- Park SJ, Park BJ, Rhee SK. Comparative analysis of archaeal 16S rRNA and amoA genes to estimate the abundance and diversity of ammonia-oxidizing archaea in marine sediments. *Extremophiles.* 2008;12:605–15.
- Vuillemin A, Wankel SD, Coskun ÖK, Magritsch T, Vargas S, Estes ER, et al. Archaea dominate oxic subseafloor communities over multimillion-year time scales. *Sci Adv.* 2019;5:eaaw4108.
- Nunoura T, Nishizawa M, Hirai M, Shimamura S, Harnvoravongchai P, Koide O, et al. Microbial diversity in sediments from the bottom of the Challenger Deep, the Mariana Trench. *Microbes Environ.* 2018;33:186–94.
- Zhao R, Hannisdal B, Mogollon JM, Jørgensen SL. Nitrifier abundance and diversity peak at deep redox transition zones. *Sci Rep.* 2019;9:1–12.
- Jørgensen SL, Zhao R. Microbial inventory of deeply buried oceanic crust from a young ridge flank. *Front Microbiol.* 2016;7:820.
- Zhao R, Dahle H, Ramírez GA, Jørgensen SL. Indigenous ammonia-oxidizing archaea in oxic subseafloor oceanic crust. *mSystems.* 2020;5:e00758–19.
- Nunoura T, Nishizawa M, Kikuchi T, Tsubouchi T, Hirai M, Koide O, et al. Molecular biological and isotopic biogeochemical prognoses of the nitrification-driven dynamic microbial nitrogen cycle in hadopelagic sediments. *Environ Microbiol.* 2013;15:3087–107.
- Morono Y, Terada T, Nishizawa M, Ito M, Hillion F, Takahata N, et al. Carbon and nitrogen assimilation in deep subseafloor microbial cells. *Proc Natl Acad Sci USA.* 2011;108:18295–300.
- Wheat CG, Fisher AT, McManus J, Hulme SM, Orcutt BN. Cool seafloor hydrothermal springs reveal global geochemical fluxes. *Earth Planet Sci Lett.* 2017;476:179–88.
- Orsi WD. Ecology and evolution of seafloor and subseafloor microbial communities. *Nat Rev Microbiol.* 2018;16:671–83.
- D'Hondt S, Pockalny R, Fulfer VM, Spivack AJ. Subseafloor life and its biogeochemical impacts. *Nat Commun.* 2019;10:3519.
- Qin W, Martens-Habbena W, Kobelt JN, Stahl DA. *Candidatus Nitrosopumilales*. In: *Bergey's manual of systematics of archaea and bacteria*. Hoboken, NJ: John Wiley & Sons; 2016.
- Santoro AE, Dupont CL, Richter RA, Craig MT, Carini P, McIlvin MR, et al. Genomic and proteomic characterization of '*Candidatus Nitrosopelagicus brevis*': an ammonia-oxidizing archaeon from the open ocean. *Proc Natl Acad Sci USA.* 2015;112:1173–8.
- Nunoura T, Takaki Y, Hirai M, Shimamura S, Makabe A, Koide O, et al. Hadal biosphere: insight into the microbial ecosystem in the deepest ocean on Earth. *Proc Natl Acad Sci USA.* 2015;112: E1230–6.
- Wang Y, Huang J-M, Cui G-J, Nunoura T, Takaki Y, Li W-L, et al. Genomics insights into ecotype formation of ammonia-oxidizing archaea in the deep ocean. *Environ Microbiol.* 2019;21:716–29.
- León-Zayas R, Novotny M, Podell S, Shepard CM, Berkenpas E, Nikolenko S, et al. Single cells within the Puerto Rico Trench suggest hadal adaptation of microbial lineages. *Appl Environ Microbiol.* 2015;81:8265–76.
- Sintes E, Bergauer K, De Corte D, Yokokawa T, Herndl GJ. Archaeal amoA gene diversity points to distinct biogeography of ammonia-oxidizing Crenarchaeota in the ocean. *Environ Microbiol.* 2013;15:1647–58.
- Nunoura T, Hirai M, Yoshida-Takashima Y, Nishizawa M, Kawagucci S, Yokokawa T, et al. Distribution and niche separation of planktonic microbial communities in the water columns from the surface to the hadal waters of the Japan Trench under the eutrophic ocean. *Front Microbiol.* 2016;7:1261.
- Santoro AE, Saito MA, Goepfert TJ, Lamborg CH, Dupont CL, DiTullio GR. Thaumarchaeal ecotype distributions across the equatorial Pacific Ocean and their potential roles in nitrification and sinking flux attenuation. *Limnol Oceanogr.* 2017;62:1984–2003.
- Alves RJE, Minh BQ, Urich T, von Haeseler A, Schleper C. Unifying the global phylogeny and environmental distribution of ammonia-oxidising archaea based on amoA genes. *Nat Commun.* 2018;9:1517.
- 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.
- Stolper DA, Buchholz CE. Neoproterozoic to early Phanerozoic rise in island arc redox state due to deep ocean oxygenation and increased marine sulfate levels. *Proc Natl Acad Sci USA.* 2019;116:8746–55.
- Abby SS, Kerou M, Schleper C. Ancestral reconstructions decipher major adaptations of ammonia-oxidizing archaea upon radiation into moderate terrestrial and marine environments. *MBio.* 2020;11:e02371–20.
- Konneke M, Schubert DM, Brown PC, Hugler M, Standfest S, Schwander T, et al. Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO₂ fixation. *Proc Natl Acad Sci USA.* 2014;111:8239–44.
- Molari M, Manini E, Dell'Anno A. Dark inorganic carbon fixation sustains the functioning of benthic deep-sea ecosystems. *Glob Biogeochem Cycles.* 2013;27:212–21.
- Dekas AE, Parada AE, Mayali X, Fuhrman JA, Wollard J, Weber PK, et al. Characterizing chemoautotrophy and heterotrophy in marine archaea and bacteria with single-cell multi-isotope nanoSIP. *Front Microbiol.* 2019;10:2682.
- Kits KD, Sedlacek CJ, Lebedeva EV, Han P, Bulaev A, Pjevac P, et al. Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle. *Nature.* 2017;549:269–72.
- Eloe EA, Fadrosch DW, Novotny M, Zeigler Allen L, Kim M, Lombardo M-J, et al. Going deeper: metagenome of a hadopelagic microbial community. *PLoS ONE.* 2011;6:e20388.
- Lloyd KG, Schreiber L, Petersen DG, Kjeldsen KU, Lever MA, Steen AD, et al. Predominant archaea in marine sediments degrade detrital proteins. *Nature.* 2013;496:215–8.

38. Li M, Baker BJ, Anantharaman K, Jain S, Breier JA, Dick GJ. Genomic and transcriptomic evidence for scavenging of diverse organic compounds by widespread deep-sea archaea. *Nat Commun.* 2015;6:8933.
39. Zhao R, Mogollón JM, Abby SS, Schleper C, Biddle JF, Roerdink D, et al. Geochemical transition zone powering microbial growth in subsurface sediments. *Proc Natl Acad Sci USA.* 2020;117:32617–26.
40. Villinger HW, Müller P, Bach W, Becker K, Orcutt BN, Kaul N, et al. Evidence for low temperature diffuse venting at North Pond, western flank of the Mid-Atlantic Ridge. *Geochem Geophys Geosyst.* 2019;20:2572–84.
41. Malinverno A, Martinez EA. The effect of temperature on organic carbon degradation in marine sediments. *Sci Rep.* 2015;5:17861.
42. Schmidt-Schierhorn F, Kaul N, Stephan S, Villinger H. Geophysical site survey results from North Pond (Mid-Atlantic Ridge). *Proc Integr Ocean Drill Progr.* 2012;336:1–62.
43. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114–20.
44. 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.
45. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, et al. Binning metagenomic contigs by coverage and composition. *Nat Methods.* 2014;11:1144–6.
46. Kang DD, Froula J, Egan R, Wang Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ.* 2015;3:e1165
47. Wu Y-W, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics.* 2016;32:605–7.
48. 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.
49. Vallenet D, Calteau A, Cruveiller S, Gachet M, Lajus A, Josso A, et al. MicroScope in 2017: an expanding and evolving integrated resource for community expertise of microbial genomes. *Nucl Acids Res.* 2017;45:D517–28.
50. Makarova KS, Wolf YI, Koonin EV. Archaeal Clusters of Orthologous Genes (arCOGs): An Update and Application for Analysis of Shared Features between Thermococcales, Methanococcales, and Methanobacteriales. *Life* 2015;5:818–840.
51. Kristensen DM, Kannan L, Coleman MK, Wolf YI, Sorokin A, Koonin EV, et al. A low-polynomial algorithm for assembling clusters of orthologous groups from intergenomic symmetric best matches. *Bioinformatics.* 2010;26:1481–7.
52. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics.* 2012;28:3150–2.
53. Graham ED, Heidelberg JF, Tully BJ. Potential for primary productivity in a globally-distributed bacterial phototroph. *ISME J.* 2018;12:1861–6.
54. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF, et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature.* 2013;499:431–7.
55. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30:772–80.
56. Criscuolo A, Gribaldo S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol.* 2010;10:210.
57. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol.* 2020;37:1530–4.
58. Durbin AM, Teske A. Sediment-associated microdiversity within the Marine Group I Crenarchaeota. *Environ Microbiol Rep.* 2010;2:693–703.
59. Zhang S, Song W, Wemheuer B, Reveillaud J, Webster N, Thomas T. Comparative genomics reveals ecological and evolutionary insights into sponge-associated Thaumarchaeota. *mSystems.* 2019;4:e00288–19.
60. Hug LA, Thomas BC, Sharon I, Brown CT, Sharma R, Hettich RL, et al. Critical biogeochemical functions in the subsurface are associated with bacteria from new phyla and little studied lineages. *Environ Microbiol.* 2016;18:159–73.
61. Ngugi DK, Blom J, Alam I, Rashid M, Ba-Alawi W, Zhang G, et al. Comparative genomics reveals adaptations of a halotolerant thaumarchaeon in the interfaces of brine pools in the Red Sea. *ISME J.* 2015;9:396–411.
62. Dombrowski N, Teske AP, Baker BJ. Expansive microbial metabolic versatility and biodiversity in dynamic Guaymas Basin hydrothermal sediments. *Nat Commun.* 2018;9:4999.
63. Tully BJ, Heidelberg JF. Potential mechanisms for microbial energy acquisition in oxic deep sea sediments. *Appl Environ Microbiol.* 2016;82:4232–43.
64. Moeller FU, Webster NS, Herbold CW, Behnam F, Domman D, Albertsen M, et al. Characterization of a thaumarchaeal symbiont that drives incomplete nitrification in the tropical sponge *Lanthella basta*. *Environ Microbiol.* 2019;21:3831–54.
65. Jung M-Y, Islam MA, Gwak J-H, Kim J-G, Rhee S-K. *Nitrosarchaeum koreense* gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon member of the phylum Thaumarchaeota isolated from agricultural soil. *Int J Syst Evol Microbiol.* 2018;68:3084–95.
66. 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.
67. 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. *Proc Natl Acad Sci USA.* 2016;113: E7937–46.
68. Walker CB, de la Torre JR, Klotz MG, Urakawa H, Pinel N, Arp DJ, et al. *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc Natl Acad Sci USA.* 2010;107:8818–23.
69. Bayer B, Vojvoda J, Offre P, Alves RJ, Elisabeth NH, Garcia JA, et al. Physiological and genomic characterization of two novel marine thaumarchaeal strains indicates niche differentiation. *ISME J.* 2016;10:1051–63.
70. Offre P, Kerou M, Spang A, Schleper C. Variability of the transporter gene complement in ammonia-oxidizing archaea. *Trends Microbiol.* 2014;22:665–75.
71. Spang A, Poehlein A, Offre P, Zumbragel S, Haider S, Rychlik N, et al. The genome of the ammonia-oxidizing *Candidatus Nitrososphaera gargensis*: insights into metabolic versatility and environmental adaptations. *Environ Microbiol.* 2012;14:3122–45.
72. Li J, Mara P, Schubotz F, Sylvan JB, Burgaud G, Klein F, et al. Recycling and metabolic flexibility dictate life in the lower oceanic crust. *Nature.* 2020;579:250–5.
73. Pejchal R, Ludwig ML. Cobalamin-independent methionine synthase (MetE): a face-to-face double barrel that evolved by gene duplication. *PLoS Biol.* 2004;3:e31.

74. Jørgensen BB, Marshall IPG. Slow microbial life in the seabed. *Ann Rev Mar Sci.* 2016;8:311–32.
75. Crable BR, Plugge CM, McInerney MJ, Stams AJM. Formate formation and formate conversion in biological fuels production. *Enzym Res.* 2011;2011:1–8.
76. Matelska D, Shabalin IG, Jabłońska J, Domagalski MJ, Kutner J, Ginalski K, et al. Classification, substrate specificity and structural features of D-2-hydroxyacid dehydrogenases: 2HADH knowledgebase. *BMC Evol Biol.* 2018;18:199.
77. Tishkov VI, Popov VO. Catalytic mechanism and application of formate dehydrogenase. *Biochem.* 2004;69:1252–67.
78. Overkamp KM, Kötter P, van der Hoek R, Schoondermark-Stolk S, Luttk MAH, van Dijken JP, et al. Functional analysis of structural genes for NAD⁺-dependent formate dehydrogenase in *Saccharomyces cerevisiae*. *Yeast.* 2002;19:509–20.
79. Perret A, Lechaplais C, Tricot S, Perchat N, Vergne C, Pellé C, et al. A novel acyl-CoA beta-transaminase characterized from a metagenome. *PLoS ONE.* 2011;6:e22918.
80. Desguin B, Soumillion P, Hausinger RP, Hols P. Unexpected complexity in the lactate racemization system of lactic acid bacteria. *FEMS Microbiol Rev.* 2017;41:S71–S83.
81. Roche J, Girard E, Mas C, Madern D. The archaeal LDH-like malate dehydrogenase from *Ignicoccus islandicus* displays dual substrate recognition, hidden allostery and a non-canonical tetrameric oligomeric organization. *J Struct Biol.* 2019;208:7–17.
82. Greening C, Ahmed FH, Mohamed AE, Lee BM, Pandey G, Warden AC, et al. Physiology, biochemistry, and applications of F420 - and Fo-dependent redox reactions. *Microbiol Mol Biol Rev.* 2016;80:451–93.
83. Ney B, Ahmed FH, Carere CR, Biswas A, Warden AC, Morales SE, et al. The methanogenic redox cofactor F 420 is widely synthesized by aerobic soil bacteria. *ISME J.* 2017;11:125–37.
84. Jørgensen BB, Boetius A. Feast and famine—microbial life in the deep-sea bed. *Nat Rev Microbiol.* 2007;5:770–81.
85. Hoehler TM, Jørgensen BB. Microbial life under extreme energy limitation. *Nat Rev Microbiol.* 2013;11:83–94.
86. Orcutt BN, Sylvan JB, Knab NJ, Edwards KJ. Microbial ecology of the dark ocean above, at, and below the seafloor. *Microbiol Mol Biol Rev.* 2011;75:361–422.
87. Lever MA, Rogers KL, Lloyd KG, Overmann J, Schink B, Thauer RK, et al. Life under extreme energy limitation: a synthesis of laboratory- and field-based investigations. *FEMS Microbiol Rev.* 2015;39:688–728.
88. Vannier P, Michoud G, Oger P, Marteinsson VP, Jebbar M. Genome expression of *Thermococcus barophilus* and *Thermococcus kodakarensis* in response to different hydrostatic pressure conditions. *Res Microbiol.* 2015;166:717–25.
89. Ohke Y, Sakoda A, Kato C, Sambongi Y, Kawamoto J, Kurihara T, et al. Regulation of cytochrome c - and quinol oxidases, and piezotolerance of their activities in the deep-sea piezophile *Shewanella violacea* DSS12 in response to growth conditions. *Biosci Biotechnol Biochem.* 2013;77:1522–8.
90. Campanaro S, De Pascale F, Telatin A, Schiavon R, Bartlett DH, Valle G. The transcriptional landscape of the deep-sea bacterium *Photobacterium profundum* in both a toxR mutant and its parental strain. *BMC Genom.* 2012;13:567.
91. Wang B, Qin W, Ren Y, Zhou X, Jung M-Y, Han P, et al. Expansion of Thaumarchaeota habitat range is correlated with horizontal transfer of ATPase operons. *ISME J.* 2019;13:3067–79.
92. Grüber G, Manimekalai MSS, Mayer F, Müller V. ATP synthases from archaea: the beauty of a molecular motor. *Biochim Biophys Acta - Bioenerg.* 2014;1837:940–52.
93. Abe F, Horikoshi K. Analysis of intracellular pH in the yeast *Saccharomyces cerevisiae* under elevated hydrostatic pressure: a study in baro- (piezo-) physiology. *Extremophiles.* 1998;2:223–8.
94. Bartlett DH. Pressure effects on in vivo microbial processes. *Biochim Biophys Acta - Protein Struct Mol Enzymol.* 2002;1595:367–81.
95. Siliakus MF, van der Oost J, Kengen SWM. Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure. *Extremophiles.* 2017;21:651–70.
96. Severi E, Müller A, Potts JR, Leech A, Williamson D, Wilson KS, et al. Sialic acid mutarotation is catalyzed by the *Escherichia coli* β -propeller protein YjhT. *J Biol Chem.* 2008;283:4841–9.
97. Ophir T, Gutnick DL. A role for exopolysaccharides in the protection of microorganisms from desiccation. *Appl Environ Microbiol.* 1994;60:740–5.
98. Higgins MJ, Polcik M, Fukuma T, Sader JE, Nakayama Y, Jarvis SP. Structured water layers adjacent to biological membranes. *Biophys J.* 2006;91:2532–42.
99. Disalvo EA, Lairion F, Martini F, Tymczynsyn E, Frías M, Almaleck H, et al. Structural and functional properties of hydration and confined water in membrane interfaces. *Biochim Biophys Acta - Biomembr.* 2008;1778:2655–70.
100. Disalvo A, de los Angeles Frias M. The role of water in the responsive properties in lipid interphase of biomimetic systems. In: *Liposomes—advances and perspectives.* IntechOpen; 2019.
101. Martinez N, Michoud G, Cario A, Ollivier J, Franzetti B, Jebbar M, et al. High protein flexibility and reduced hydration water dynamics are key pressure adaptive strategies in prokaryotes. *Sci Rep.* 2016;6:32816.
102. Kandiba L, Eichler J. Analysis of putative nonulosonic acid biosynthesis pathways in Archaea reveals a complex evolutionary history. *FEMS Microbiol Lett.* 2013;345:110–20.
103. Muñoz-Marín MC, Gómez-Baena G, López-Lozano A, Moreno-Cabezuelo JA, Díez J, García-Fernández JM. Mixotrophy in marine picocyanobacteria: use of organic compounds by *Prochlorococcus* and *Synechococcus*. *ISME J.* 2020;14:1065–73.
104. Michoud G, Jebbar M. High hydrostatic pressure adaptive strategies in an obligate piezophile *Pyrococcus yayanosii*. *Sci Rep.* 2016;6:27289.
105. Cario A, Lormières F, Xiang X, Oger P. High hydrostatic pressure increases amino acid requirements in the piezo-hyperthermophilic archaeon *Thermococcus barophilus*. *Res Microbiol.* 2015;166:710–6.
106. Amrani A, Bergon A, Holota H, Tamburini C, Garel M, Ollivier B, et al. Transcriptomics reveal several gene expression patterns in the piezophile *Desulfovibrio hydrothermalis* in response to hydrostatic pressure. *PLoS ONE.* 2014;9:e106831.
107. Amrani A, van Helden J, Bergon A, Aouane A, Ben Hania W, Tamburini C, et al. Deciphering the adaptation strategies of *Desulfovibrio piezophilus* to hydrostatic pressure through metabolic and transcriptional analyses. *Environ Microbiol Rep.* 2016;8:520–6.
108. Empadinhas N, Da Costa MS. Osmoadaptation mechanisms in prokaryotes: distribution of compatible solutes. *Int Microbiol.* 2008;11:151–61.
109. Lim J, Thomas T, Cavicchioli R. Low temperature regulated DEAD-box RNA helicase from the antarctic archaeon, *Methanococcoides burtonii*. *J Mol Biol.* 2000;297:553–67.
110. Cavicchioli R, Thomas T, Curmi PMG. Cold stress response in Archaea. *Extremophiles.* 2000;4:321–31.
111. Hunger K, Beckering CL, Wiegshoff F, Graumann PL, Marahiel MA. Cold-induced putative DEAD box RNA helicases CshA and CshB are essential for cold adaptation and interact with cold shock protein B in *Bacillus subtilis*. *J Bacteriol.* 2006;188:240–8.
112. Mhatre SS, Kaufmann S, Marshall IPG, Obrochta S, Andrèn T, Jørgensen BB, et al. Microbial biomass turnover times and clues to cellular protein repair in energy-limited deep Baltic Sea sediments. *FEMS Microbiol Ecol.* 2019;95:fiz0068.

113. Bradley JA, Amend JP, LaRowe DE. Survival of the fewest: microbial dormancy and maintenance in marine sediments through deep time. *Geobiology*. 2019;17:43–59.
114. Rodrigues MV, Borges N, Henriques M, Lamosa P, Ventura R, Fernandes C, et al. Bifunctional CTP:inositol-1-phosphate cytidyltransferase/CDP-inositol:inositol-1-phosphate transferase, the key enzyme for di-myo-inositol-phosphate synthesis in several (hyper)thermophiles. *J Bacteriol*. 2007;189:5405–12.
115. Gonçalves LG, Borges N, Serra F, Fernandes PL, Dopazo H, Santos H. Evolution of the biosynthesis of di-myo-inositol phosphate, a marker of adaptation to hot marine environments. *Environ Microbiol*. 2012;14:691–701.
116. Empadinhas N, da Costa MS. Diversity, biological roles and biosynthetic pathways for sugar-glycerate containing compatible solutes in bacteria and archaea. *Environ Microbiol*. 2011;13:2056–77.