environmental microbiology

Environmental Microbiology (2021) **23**(11), 7152–7167 doi:10.1111/1462-2920.15759

Microbes mediating the sulfur cycle in the Atlantic Ocean and their link to chemolithoautotrophy

Daniele De Corte,[1*] **Simone Muck**,[2] **Johanna Tiroch**,[2] **Catalina Mena**,[3] **Gerhard J. Herndl**[2,4] **and Eva Sintes**[2,3]

[1]*Institute for Chemistry and Biology of the Marine Environment, Carl Von Ossietzky University, Oldenburg, Germany.*

[2]*Department of Functional and Evolutionary Ecology, University of Vienna, Vienna, Austria.*

[3]*Instituto Español de Oceanografía, Centro Oceanográfico de Baleares, Palma, Spain.*

[4]*NIOZ, Department of Marine Microbiology and Biogeochemistry, Royal Netherlands Institute for Sea Research, Utrecht University, Den Burg, The Netherlands.*

Summary

Only about 10%–30% of the organic matter produced in the epipelagic layers reaches the dark ocean. Under these limiting conditions, reduced inorganic substrates might be used as an energy source to fuel prokaryotic chemoautotrophic and/or mixotrophic activity. The *apr*A gene encodes the alpha subunit of the adenosine-5′-phosphosulfate (APS) reductase, present in sulfate-reducing (SRP) and sulfur-oxidizing prokaryotes (SOP). The sulfur-oxidizing pathway can be coupled to inorganic carbon fixation via the Calvin–Benson–Bassham cycle. The abundances of *apr*A and *cbb*M, encoding RuBisCO form II (the key CO_2 fixing enzyme), were determined over the entire water column along a latitudinal transect in the Atlantic from 64°N to 50°S covering six oceanic provinces. The abundance of *apr*A and *cbb*M genes significantly increased with depth reaching the highest abundances in meso- and upper bathypelagic layers. The contribution of cells containing these genes also increased from mesotrophic towards oligotrophic provinces, suggesting that under nutrient limiting conditions alternative energy sources are advantageous. However, the *apr*A/*cbb*M ratios indicated that only a fraction of the SOP is associated with inorganic carbon fixation. The *apr*A harbouring prokaryotic community was dominated by Pelagibacterales in surface and mesopelagic waters, while *Candidatus* Thioglobus, Chromatiales and the Deltaproteobacterium_SCGC dominated the bathypelagic realm. Noticeably, the contribution of the SRP to the prokaryotic community harbouring *apr*A gene was low, suggesting a major utilization of inorganic sulfur compounds either as an energy source (occasionally coupled with inorganic carbon fixation) or in biosynthesis pathways.

Received 2 December, 2020; revised 18 August, 2021; accepted 3 September, 2021. *For correspondence. E-mail daniele.de.corte@uni-oldenburg.de. Telephone: (+49) 441-798-3704 Fax: (+49) 441-798-3358

Introduction

Prokaryotes play a key role in the biochemical cycles of the global ocean (Azam *et al*., 1983). 10%–30% of the epipelagic primary production is exported as organic carbon into the meso- and bathypelagic layers (Buesseler *et al*., 2007; Buesseler and Boyd, 2009), limiting heterotrophic prokaryotes growth in the dark ocean (Herndl and Reinthaler, 2013). Recent studies indicate that dissolved inorganic carbon (DIC) fixation driven by chemical energy represents an alternative source of organic carbon (Walsh *et al*., 2009; Reinthaler *et al*., 2010; Swan *et al*., 2011; Sintes *et al*., 2013), together with non-sinking particulate organic matter for deep ocean prokaryotes (Baltar *et al*., 2010; Herndl and Reinthaler, 2013).

Several reduced inorganic compounds, such as ammonia, hydrogen and diverse sulfur compounds, can potentially be used as an energy source to fuel mixotrophic and/or chemoautotrophic processes in the dark ocean (Herndl *et al*., 2005; Swan *et al*., 2011; Anantharaman *et al*., 2013). The oxidation of reduced sulfur compounds is a widespread process in hydrothermal vent plumes and oxygen minimum zones (OMZs) (Sievert *et al*., 2007; Walsh *et al*., 2009; Ulloa *et al*., 2012). However, the extent to which reduced sulfur is used as an electron donor or energy source in the oxygenated oceanic water column remains poorly investigated. Reduced sulfur is typically low in concentration and under the detection limit of current methods outside of upwelling regions and the vicinity of hydrothermal vents (Radford-Knoery *et al*., 2001; Sievert *et al*., 2007). A potential source for reduced

sulfur in the otherwise oxygenated water column are marine snow particles, where respiration can be fuelled by sulfate reduction and sulfide can be generated in oxygen-depleted or anoxic microenvironments (Shanks and Reeder, 1993; Bianchi *et al*., 2018). Organosulfur compounds primarily released by eukaryotes such as dimethylsulfoniopropionate (DMSP) and taurine are utilized by marine microbes (Clifford *et al*., 2019; Durham *et al*., 2019). Sulfate-reducing and sulfur-oxidizing prokaryotes (SOP) harbour the *apr* gene (Blazejak *et al*., 2006; Meyer and Kuever, 2007a). The *apr* gene encodes APS reductase, which forms a heterodimer with one alpha- and one beta-subunit (Lampreia *et al*., 1994; Hipp *et al*., 1997; Fritz *et al*., 2000) and is a key enzyme in the dissimilatory sulfate reduction pathway (Canfield *et al*., 2005; Meyer and Kuever, 2007a) of sulfate-reducing prokaryotes (SRP). However, this enzyme also operates in reverse direction in SOP, catalysing the transformation of sulfite to sulfate (Hipp *et al*., 1997; Kappler and Dahl, 2001; Meyer and Kuever, 2007a, 2007b). APS reductase has also been suggested to function as an intracellular detoxification mechanism of sulfite, a metabolite of organic sulfur degradation, in prokaryotes with an incomplete dissimilatory sulfur oxidation pathway (Meyer and Kuever, 2007b).

The aim of this study was to determine the distribution patterns and the phylogenetic affiliation of the *apr*A-harbouring prokaryotic community throughout the water column along a latitudinal transect in the Atlantic Ocean covering six different oceanographic provinces (Fig. 1) (Longhurst, 2007). The potential link between chemoautotrophy and sulfur oxidation was also explored by relating the *apr*A to *cbb*M gene abundances and characterizing their biogeographical patterns. Quantitative-PCR (q-PCR) of two functional genes *cbb*M encoding the 1,5-bisphosphate carboxylase (RuBisCO), the key CO_2 fixing enzyme of the Calvin–Benson–Bassham cycle, and *apr*A (encoding the APS reductase alpha subunit) was used to assess the biogeographic and depth-related distribution of prokaryotes harbouring these genes in the Atlantic Ocean. This quantification of *apr*A and *cbb*M was related to physicochemical and biological environmental factors. Similar patterns of increasing abundance towards the meso- and bathypelagic layers and towards low latitude regions were observed for both genes.

Additionally, tag sequencing of the *apr*A gene was used to characterize the SRP and SOP community composition in the different biogeographic provinces and depth zones, as well as to identify prokaryotes with a potentially incomplete sulfur oxidation pathway. Our results suggest that sulfur-oxidizing bacteria are important players in the sulfur cycling of the meso- and bathypelagic Atlantic Ocean and that sulfur oxidation is only partially coupled with DIC fixation as revealed by low *apr*A/*cbb*M ratios.

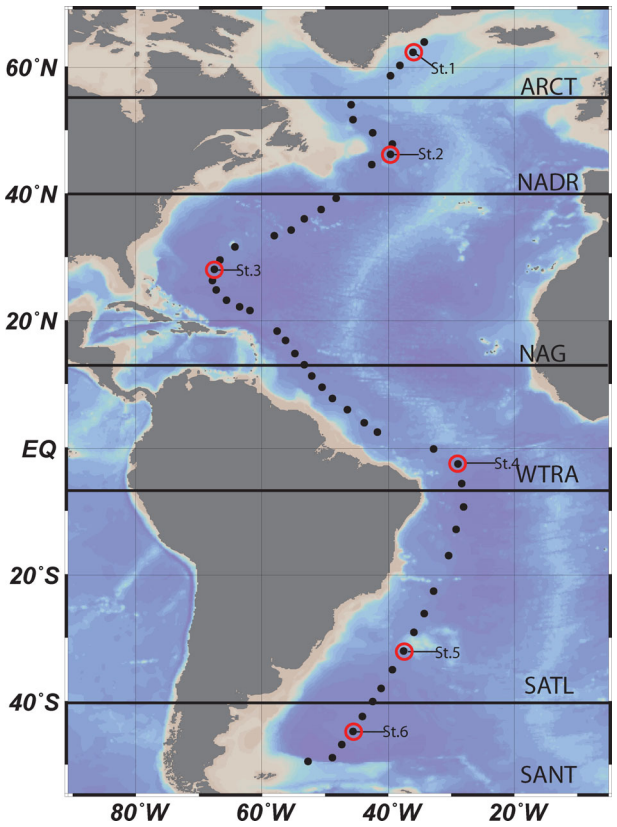

Fig. 1. Location of stations sampled during the GEOTRACES expeditions. Oceanographic regions are classified according to Longhurst (2007): ARCT, North Atlantic Arctic Province; NADR, North Atlantic Drift Province; NAG, North Atlantic Gyral Province; WTRA, Western Tropical Atlantic; SATL, South Atlantic Gyral Province; SANT, Subantarctic Province. Tag sequencing of the *apr*A gene was conducted in the red encircled stations.

Results

Distribution of prokaryotes harbouring the apr*A and* cbb*M genes in the Atlantic*

The abundance of the *apr*A harbouring prokaryotic community throughout the Atlantic was highest in the meso- and upper bathypelagic layers (with a median of 4.2×10^2 and 3.4×10^2 genes ml^{-1} respectively). Generally, a higher abundance of *apr*A genes was found at low latitude regions (NAG, WTRA and SATL) with a median of 3.1×10^2 genes ml^{-1} (Table S1; Fig. 2A and B) than at high latitude regions (with a median of 1.5×10^2 genes ml^{-1}). The *apr*A containing cells to total prokaryotic abundance (PA) ratio (i.e. abundance of *apr*A gene divided by the total PA) significantly increased with depth in all the oceanographic provinces peaking in the lower bathypelagic layer of the low latitude regions (with a median ratio ranging between 1.89×10^{-2} and 3.46×10^{-2}) (Fig. S1a, b; Table S1).

The *cbb*M gene abundance increased from the epipelagic waters (with a median of 3.5×10^0 genes ml^{-1}) to

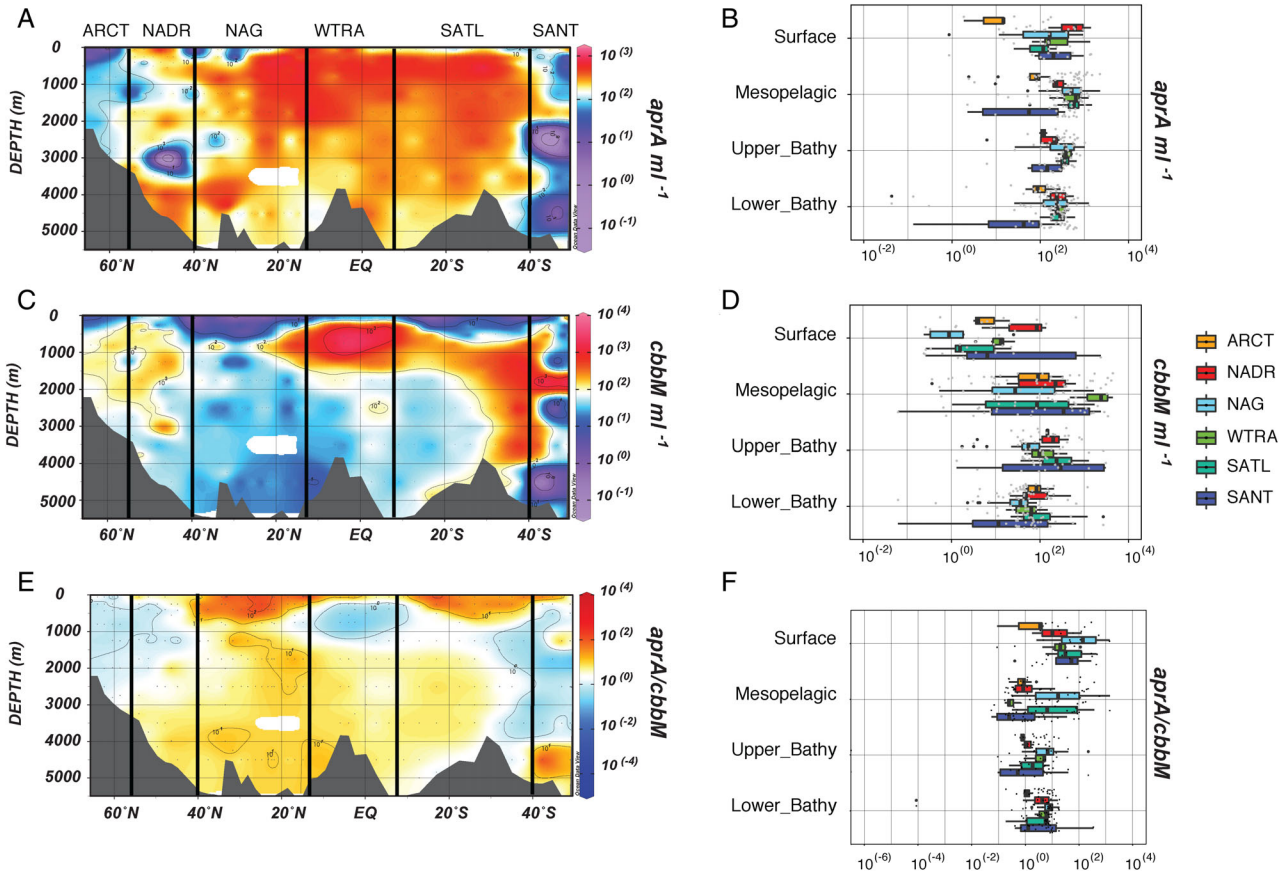

Fig. 2. *Apr*A ml^{-1} (A), *cbb*M ml^{-1} (C) and *apr*A/*cbb*M (E) throughout the Atlantic Ocean determined by q-PCR. Box plot showing *apr*A ml^{-1} (B), *cbb*M ml^{-1} (D) and *apr*A/*cbb*M (F) ratios in different regions (colour coded) and depth layers. The left and right of the box represent the first and the third quartiles, while the thick vertical line represents the median. Outliers (larger black points) and jitter values (smaller black points) are shown.

the mesopelagic layer (with a median of 1.30×10^2 genes ml^{-1}) and decreasing again in the bathypelagic waters (Table S1; Fig. 2C and D). The ratio of *cbb*M containing cells to total prokaryotic cell abundance (i.e. abundance of *cbbM* genes divided the total PA) showed similar depth and latitudinal patterns as the *apr*A to PA ratio, with the highest ratio found in the mesopelagic layer of the WTRA region (with a median of 3.94×10^{-2}) (Fig. S1c, d; Table S1).

The *apr*A/*cbb*M ratio decreased from the epipelagic layers (with a median of 31 for the whole dataset) to the mesopelagic layers (median 1.73) (Table S1; Fig. 2E and F). The *apr*A/*cbb*M ratios were highest in the epipelagic communities of the NAG province (with a median of 179), whereas the lowest ratio was found in the mesopelagic prokaryotic community of the WTRA province (with a median of 0.25) (Table S1; Fig. 2E and F).

Distance-based redundancy analysis (dbRDA) was used to explain the variation of *apr*A and *cbb*M gene abundances and *apr*A/*cbb*M ratios along the transect in different depth layers. The percent variation explained by the selected environmental parameters ranged between 0% and 42% for the *apr*A gene abundance, between 48% and 84% for the *cbb*M gene abundance and between 0% and 44% for the *apr*A/*cbb*M ratio, depending on the pelagic zone (Table 1). The environmental parameters analysed did not explain most of the variation of *apr*A gene in the upper and lower bathypelagic layers, nor the variation of the *apr*A/*cbb*M ratio in the epipelagic and upper bathypelagic layers. Latitude and temperature were the main explanatory factors influencing the distribution of the sulfur-oxidizing/-reducing and chemolithoautotrophic genes in the different depth layers (Table 1).

*Diversity and richness of the prokaryotic community harbouring the adenylyl-sulfate reductase (APS) subunit alpha (*apr*A) gene*

Twenty-four samples were taken for determining the composition of the *apr*A harbouring prokaryotic community from four depth layers at six oceanic provinces (Fig. 1). In total, 787 447 paired reads were obtained, with $3.3 \pm 1.3 \times 10^4$ reads per sample. After trimming and low-quality reads filtering, the total number of reads

Table 1. Distance-based redundancy analysis results of the environmental parameters best explaining the variation of the *apr*A ml^{-1}, *cbb*M ml^{-1} and *apr*A/*cbb*M ratios in different depth layers (Cont = Contribution, *F* = *F*-test, *p* = *p*-value).

	***apr*A ml^{-1}**											
	Epipelagic			**Mesopelagic**			**Upper_Bathypelagic**			**Lower_Bathypelagic**		
Variables	**Cont**	***F***	***p***	**Cont**	***F***	***p***	**Cont**	***F***	***p***	**Cont**	***F***	***p***
Abs_Latitude				0.294	29.72	0.001	0.183	11.45	0.001			
Depth (m)												
Temperature (°C)	0.167	6.18	0.023	0.118	11.95	0.003						
Salinity	0.126	4.64	0.044									
Oxygen (μM)												
Silicate (μM)												
Phosphate (μM)												
Nitrate (μM)												
Nitrite (μM)												
DIC (μM)	0.126	4.67	0.059									
Fe (nM)												
Cd (nM)							0.07	4.38	0.037			
Zn (nM)												
Mn (nM)												
Cumulative	**0.42**			**0.41**			**0.25**					
	***cbb*M ml^{-1}**											
	Surface			**Mesopelagic**			**Upper_Bathypelagic**			**Lower_Bathypelagic**		
Variables	**Cont**	***F***	***p***	**Cont**	***F***	***p***	**Cont**	***F***	***p***	**Cont**	***F***	***p***
Abs_Latitude	0.064	32.54	0.009	0.312	48.28	0.001	0.123	13.78	0.002	0.052	7.64	0.006
Depth (m)												
Temperature (°C)	0.170	62.00	0.004	0.260	40.20	0.001	0.175	19.61	0.001	0.180	26.27	0.001
Salinity	0.084	30.61	0.024							0.041	6.03	0.03
Oxygen (μM)							0.204	22.83	0.002			
Silicate (μM)	0.422	153.8	0.017									
Phosphate (μM)										0.090	13.10	0.001
Nitrate (μM)	0.059	21.41	0.006							0.051	7.43	0.017
Nitrite (μM)												
DIC (μM)							0.067	7.46	0.016			
Fe (nM)							0.046	5.09	0.04	0.032	4.73	0.048
Cd (nM)										0.038	5.61	0.036
Zn (nM)												
Mn (nM)	0.04	14.19	0.019	0.023	3.67	0.041						
Cumulative	**0.84**			**0.60**			**0.62**			**0.48**		
	***apr*A/*cbb*M**											
	Surface			**Mesopelagic**			**Upper_Bathypelagic**			**Lower_Bathypelagic**		
Variables	**Cont**	***F***	***p***	**Cont**	***F***	***p***	**Cont**	***F***	***p***	**Cont**	***F***	***p***
Abs_Latitude										0.102	13.73	0.001
Depth (m)				0.048	5.83	0.018						
Temperature (°C)				0.248	30.25	0.001				0.179	24.10	0.001
Salinity												
Oxygen (μM)				0.114	13.92	0.001						
Silicate (μM)										0.067	9.08	0.006
Phosphate (μM)										0.037	4.98	0.035
Nitrate (μM)												
Nitrite (μM)												
DIC (μM)												
Fe (nM)							0.082	5.56	0.024	0.051	6.97	0.01
Cd (nM)												
Zn (nM)												
Mn (nM)				0.035	4.37	0.055	0.099	6.71	0.012			
Cumulative				**0.44**			**0.18**			**0.44**		

decreased to 403 613 paired reads with $1.7 \pm 0.7 \times 10^4$ reads per sample. These total reads resulted in 2472 amplicon sequence variants (ASVs) after applying the DADA2 algorithm. The 2472 ASVs clustered in 775 different amino acid sequences at 100% identity. Therefore, the *apr*A ASVs might not always represent different

protein sequences, and the nucleotide-based diversity might not always correspond to a similar functional diversity. However, in this study, we sequenced only a fraction of the *apr*A gene (~410 bp vs. ~1700 bp) and thus, we cannot explore the diversity of the entire protein. Taking this into consideration, we decided to use the ASVs to obtain more detailed information on the diversity of the *apr*A harbouring community.

The epi- and mesopelagic layers harboured 718 (29%) and 612 (25%) unique ASVs respectively (Fig. 3). In the lower and upper bathypelagic layers 339 (14%) and 476 (19%) unique ASVs respectively, were identified. Epi- and mesopelagic layers shared ~8% ASVs, whereas the upper and lower bathypelagic layers shared ~3% ASVs. Only eight (0.3%) ASVs were ubiquitous throughout the four pelagic zones (Fig. 3).

The cumulative read abundance showed clear differences between epi- and mesopelagic communities and bathypelagic *apr*A harbouring communities (Fig. 4). The cumulative abundance of the bathypelagic communities increased more steeply than the epipelagic communities with few very abundant ASVs (Fig. 4A). The lower bathypelagic layer exhibited the lower Chao richness and Pielou evenness indices of all the depth layers. Shannon diversity was significantly lower in the lower bathypelagic compared to the surface and mesopelagic communities (ANOVA on rank, Bonferroni corrected, $p < 0.05$). The phylogenetic diversity of the *apr*A harbouring community significantly differed between the mesopelagic and lower bathypelagic realms (ANOVA on rank, Bonferroni corrected, $p < 0.05$) (Fig. 4C). The epipelagic, mesopelagic and bathypelagic *apr*A harbouring communities were clearly separated, however, the upper and lower bathypelagic communities overlapped (Fig. 4B). The first and second coordinates of the principal coordinates analyses explained 59% and 13% of the community variation respectively.

Taxonomic assignment and community composition of apr*A harbouring prokaryotes*

Taxonomic assignment yielded 12 unclassified ASVs and two assigned only at the kingdom level, together accounting for ~0.2% of the total reads. Within the classified ASVs (at or above the phylum level), Alphaproteobacteria dominated the epipelagic *apr*A containing community (77 ± 24%), whereas Gammaproteobacteria dominated the mesopelagic (54 ± 18%) and bathypelagic (48 ± 9%) *apr*A harbouring community. Deltaproteobacteria also significantly contributed to the bathypelagic community accounting for 19 ± 9% of the *apr*A containing prokaryotic community (Table 2; Fig. 5). At the order level, Pelagibacterales dominated the epi- and the mesopelagic *apr*A containing communities, contributing 77 ± 23% and 33 ± 17% to the communities respectively. Pelagibacterales also significantly contributed (16 ± 13%) to bathypelagic *apr*A harbouring communities. Three taxa [Delta_proteobacterium_SCGC (19 ± 9%), *Candidatus*_Thioglobus (18 ± 9%) and Chromatiales (13 ± 3%)] dominated the upper and lower bathypelagic *apr*A containing prokaryotic communities.

Significant differences were observed between *apr*A containing communities of high and low latitude regions. Specifically, Alphaproteobacteria contributed more to epipelagic communities at low latitudes than at high latitudes (96 ± 2% and 58 ± 18% respectively; *T*-test $p < 0.05$). Conversely, Alphaproteobacteria accounted for a substantially lower fraction of the community in meso- and bathypelagic communities of low latitude regions compared to the high latitude communities (Table 2; Figs 5B and S2). Gammaproteobacteria showed opposite trends compared to the Alphaproteobacteria in epipelagic zones, and no significant differences between high and low latitudes for bathypelagic communities (*T*-test $p > 0.05$). Two groups (Pelagibacterales and Chromatiales) were

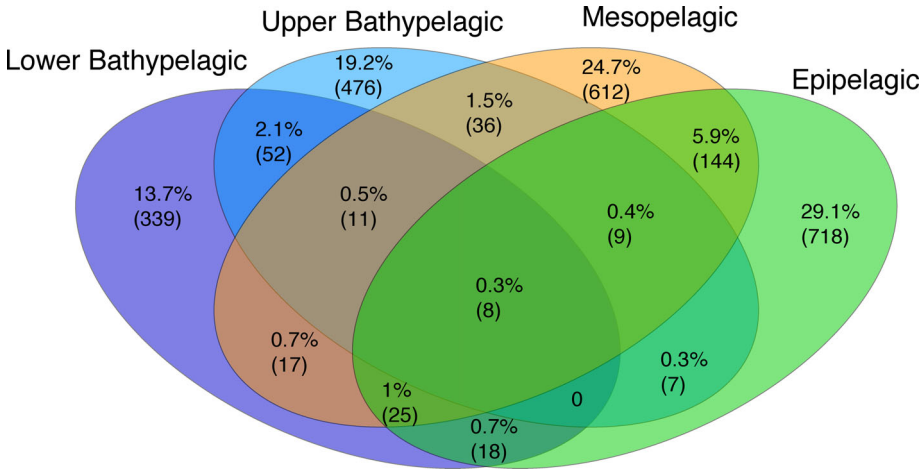

Fig. 3. Venn diagram showing the shared and unique *apr*A amplicon sequence variants (ASVs) and their relative contribution (in %) in four different depth layers throughout the Atlantic Ocean.

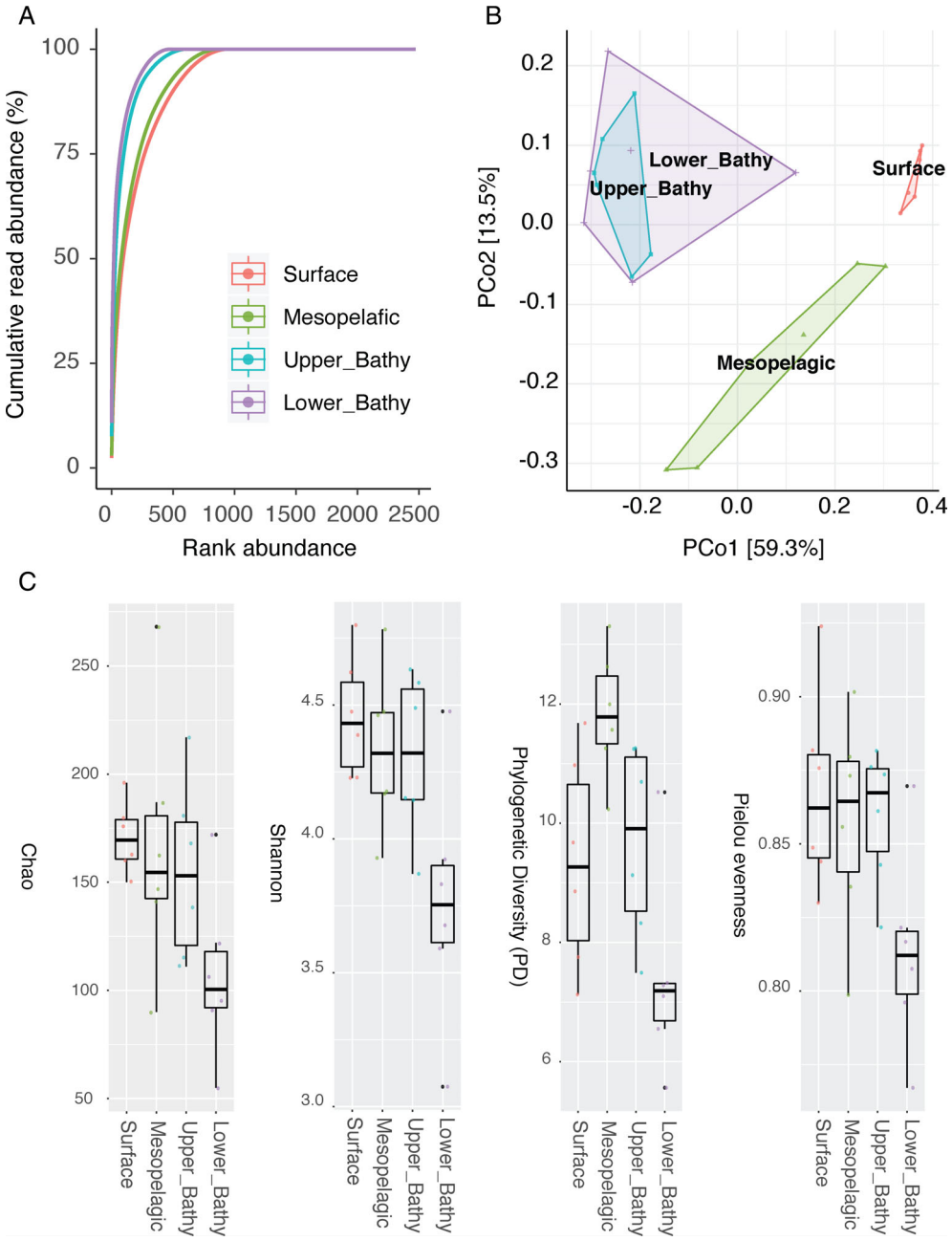

Fig. 4. Cumulative abundance (A) and principal coordinate analysis (B) of the *apr*A harbouring prokaryotic community at the four depth layers studied. Box plot showing Chao richness, Shannon and Phylogenetic Diversity (PD) indexes and Pielou evenness (C) of the *apr*A containing prokaryotic community at different depths. The bottom and the top of the box represent the first and the third quartiles, while the thick horizontal line represents the median. Outliers (black dots) and jitter values (coloured dots) are shown.

ubiquitously present and therefore significantly contributed to the epi-, meso- and bathypelagic communities. However, their contribution differed between different regions and depth layers.

Four ASVs, two closely related to the Delta_proteobacterium_SCGC and the other two putatively identified as closely related to Chromatiales, and unassigned Gammaproteobacteria composed the core bathypelagic bacterial community harbouring *apr*A genes. Only one ASV related to Pelagibacterales composed the core community of the mesopelagic layer. Epipelagic prokaryotic communities did not exhibit a core *apr*A gene harbouring community. The bathypelagic resident *apr*A containing community was dominated by ASVs with similar taxonomic affiliation as those found in the core community with the addition of *Candidatus*

Table 2. Phylogenetic affiliation and relative contribution (%, ±SD) of the *apr*A containing prokaryotic community in four different depth layers at high and low latitude oceanic provinces.

	Surface				Mesopelagic				Upper bathypelagic				Lower bathypelagic			
	High_latitude		Low_latitude		High_latitude		Low_latitude		High_latitude		Low_latitude		High_latitude		Low_latitude	
	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD
Acidiferrobacterales	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.9	1.0	1.1
Alpha_environmental_samples	0.0	0.0	0.0	0.0	0.1	0.2	1.2	0.5	5.9	4.5	4.6	1.5	14.1	9.9	7.2	5.9
Candidatus_Thioglobus	4.6	7.0	0.1	0.1	5.8	2.2	10.8	4.5	10.7	6.9	17.7	4.0	19.4	13.2	24.7	7.1
Chromatiales	8.2	6.9	2.2	2.4	16.5	8.1	9.6	1.2	11.3	4.0	14.4	2.8	10.8	2.8	14.3	1.1
Delta_proteobacterium_SCGC	0.0	0.0	0.0	0.0	0.3	0.3	0.0	0.0	18.5	8.5	23.4	3.0	9.0	3.2	25.4	10.7
Desulfobacterales	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0
Environmental_marine_bacterium_NA	24.5	18.8	1.2	2.0	6.7	11.2	0.0	0.1	0.0	0.0	0.0	0.0	0.6	0.6	0.0	0.0
Gamma_environmental_samples	0.0	0.0	0.1	0.2	3.0	3.4	11.7	8.9	4.9	0.4	9.3	1.5	1.4	1.0	4.5	1.7
Gamma_proteobacterium_SCGC	0.1	0.1	0.0	0.0	7.7	10.3	1.9	2.5	6.2	2.8	1.8	1.0	6.3	4.7	0.8	0.8
Nitrosomonadales	0.0	0.0	0.1	0.2	0.0	0.0	0.7	1.2	1.5	0.5	1.1	1.4	0.0	0.0	0.0	0.0
Pelagibacterales	57.8	18.0	95.1	3.1	44.0	13.2	22.0	13.2	16.9	3.0	13.3	0.3	22.1	28.0	13.4	7.1
Proteobacteria_Alphaproteobacteria_NA	0.0	0.1	0.2	0.2	0.3	0.5	1.1	1.1	0.0	0.0	0.7	0.5	0.0	0.0	0.0	0.0
Proteobacteria_Deltaproteobacteria_NA	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Proteobacteria_Gammaproteobacteria_NA	3.1	4.6	0.2	0.3	6.2	1.5	33.3	1.5	11.0	2.2	6.7	1.0	6.1	4.8	5.3	1.8
Proteobacteria_NA_NA	0.9	1.4	0.0	0.0	8.1	9.3	7.5	3.7	9.9	4.6	7.1	1.4	9.2	8.0	3.2	2.1
SAR116_cluster	0.2	0.2	0.5	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Sulfur_oxidizing_symbionts	0.8	0.9	0.2	0.2	1.4	2.1	0.1	0.1	3.2	5.6	0.0	0.0	0.0	0.1	0.0	0.0

Thioglobus and Pelagibacterales. Eight ASVs (two closely related to Pelagibacterales, two related to Gamma environmental samples, and the other four putatively identified as closely related to Chromatiales, *Candidatus* Thioglobus, unassigned Gammaproteobacteria and unassigned Proteobacteria) composed the resident mesopelagic bacterial community harbouring *apr*A genes. The epipelagic *apr*A harbouring resident community was dominated by ASVs putatively assigned to Pelagibacterales.

The epi- and mesopelagic were clearly separated from bathypelagic samples as indicated by a constrained principal coordinates analysis (CAP) based on weighted UniFrac distance with the first and the second coordinates accounting for 58.3% and 9.4% of the community variation respectively (Fig. 6). Furthermore, the upper and lower bathypelagic communities partially overlapped, whereas epi- and mesopelagic communities were clearly separated. Pelagibacterales ASVs comprised most of the variation of epi- and mesopelagic communities, while Chromatiales and Deltaproteobacteria accounted for most of the variation in the bathypelagic communities (Fig. 6A). Spatial and physicochemical factors explained a significant portion of the beta-diversity of the a*pr*A harbouring communities. All the selected parameters accounted together for 79% ($p < 0.001$) of the community variation. Specifically, depth (32%, $p = 0.001$), temperature (13%, $p = 0.001$) and latitude (11%, $p = 0.001$) accounted for most of the variation (summing up 56%), followed by oxygen, nitrite, DIC and silicate (Fig. 6).

Additionally, variation partition analysis revealed that 39% of the total variability in the *apr*A containing community was explained by the selected environmental parameters. Specifically, physicochemical parameters explained 27% of the variation followed by the spatial variables (depth and latitude) with 6% and temporal factors with 0.6%. The combination of the physicochemical and spatial variables explained an additional 5% of the *apr*A harbouring community variation (Fig. 7).

Distribution of Candidatus *Pelagibacter oligotypes*

Oligotyping analysis at the amino acid level of the *Candidatus* Pelagibacter *apr*A containing community identified 35 oligotypes (Fig. 8A). The epi- and mesopelagic oligotypes clearly differed from the composition of bathypelagic oligotypes. The *z*-score distribution of *apr*A containing *Candidatus* Pelagibacter oligotypes exhibited four main clusters. The first cluster grouped communities from epipelagic, another cluster grouped high latitude communities of different depths, two other clusters grouped communities obtained from the upper and lower bathypelagic layers respectively.

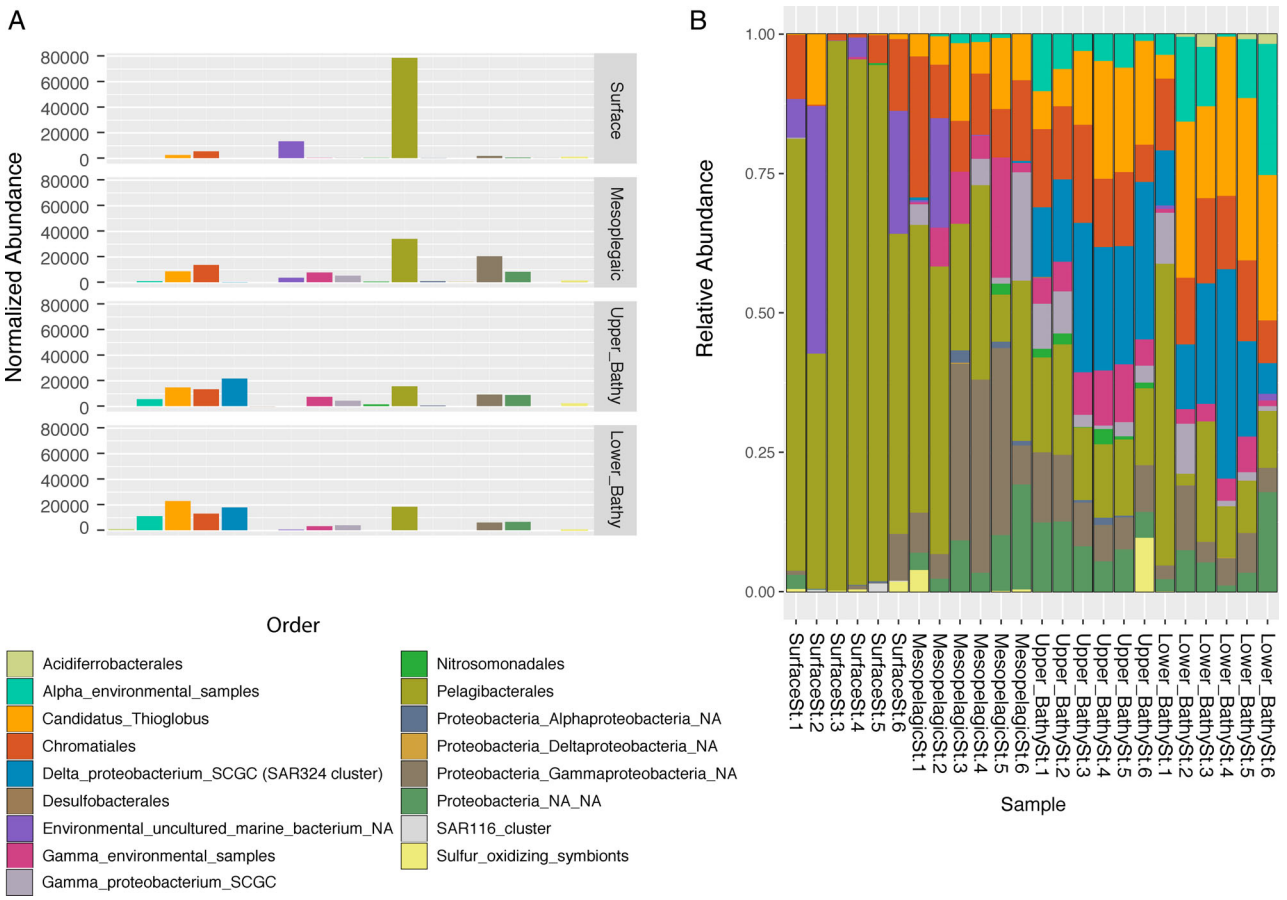

Fig. 5. Normalized abundance (by the median sequencing depth) of abundant *apr*A harbouring taxa at the order level in different depth layers (A) and their relative contribution at different stations (B) along the Atlantic transect.

Discussion

Factors driving the depth-related and biogeographic distribution of the prokaryotes with a role in sulfur cycle and chemolithoautotrophy

The distribution of prokaryotes involved in the sulfur cycle, such as those harbouring the *apr*A gene involved in sulfur oxidation or reduction, has been studied in marine sediments (Blazejak and Schippers, 2011; Watanabe *et al*., 2013; Aoki *et al*., 2015) and hydrothermal vents (Hügler *et al*., 2010; Frank *et al*., 2013). However, limited information is available on the distribution of *apr*A harbouring prokaryotes across different oceanic provinces and throughout the water column. Single amplified genomes (SAGs) and metagenomics studies revealed the importance of the potential for sulfur oxidation/reduction in the open ocean (Swan *et al*., 2011; Ulloa *et al*., 2012; Li *et al*., 2018), yet to which extent prokaryotes harbouring these genes contribute to the global sulfur cycle remains unknown.

The increase in the *apr*A abundance from the surface to the bathypelagic layers indicates an increasing contribution of genes involved in sulfur metabolism with depth coinciding with a decrease in bioavailable organic carbon (Hansell, 2013) potentially caused by dilution of dissolved organic matter (DOM) compounds (Arrieta *et al*., 2015; Wagner *et al*., 2020). This finding suggests that the limitation of bioavailable DOM in the deep ocean promotes the proliferation of alternative metabolic pathways, including chemoautotrophy or litho-heterotrophy, utilizing inorganic compounds as a supplementary source of energy. The increasing contribution of *apr*A containing prokaryotes to the communities in the deep ocean and in low latitude oligotrophic regions (reaching up to a median of 3.5% of the total prokaryotic community in the lower bathypelagic of the NAG, Table S1) supports the notion of an important role of inorganic sulfur compounds as a supplementary energy source in organic matter-limited environments. The contribution of *apr*A containing prokaryotes to the community determined in this study falls within the range (~3% to 20%) of values reported in a recent metagenomic study of the global deep ocean (Acinas *et al*., 2021). However, it is worth mentioning that our

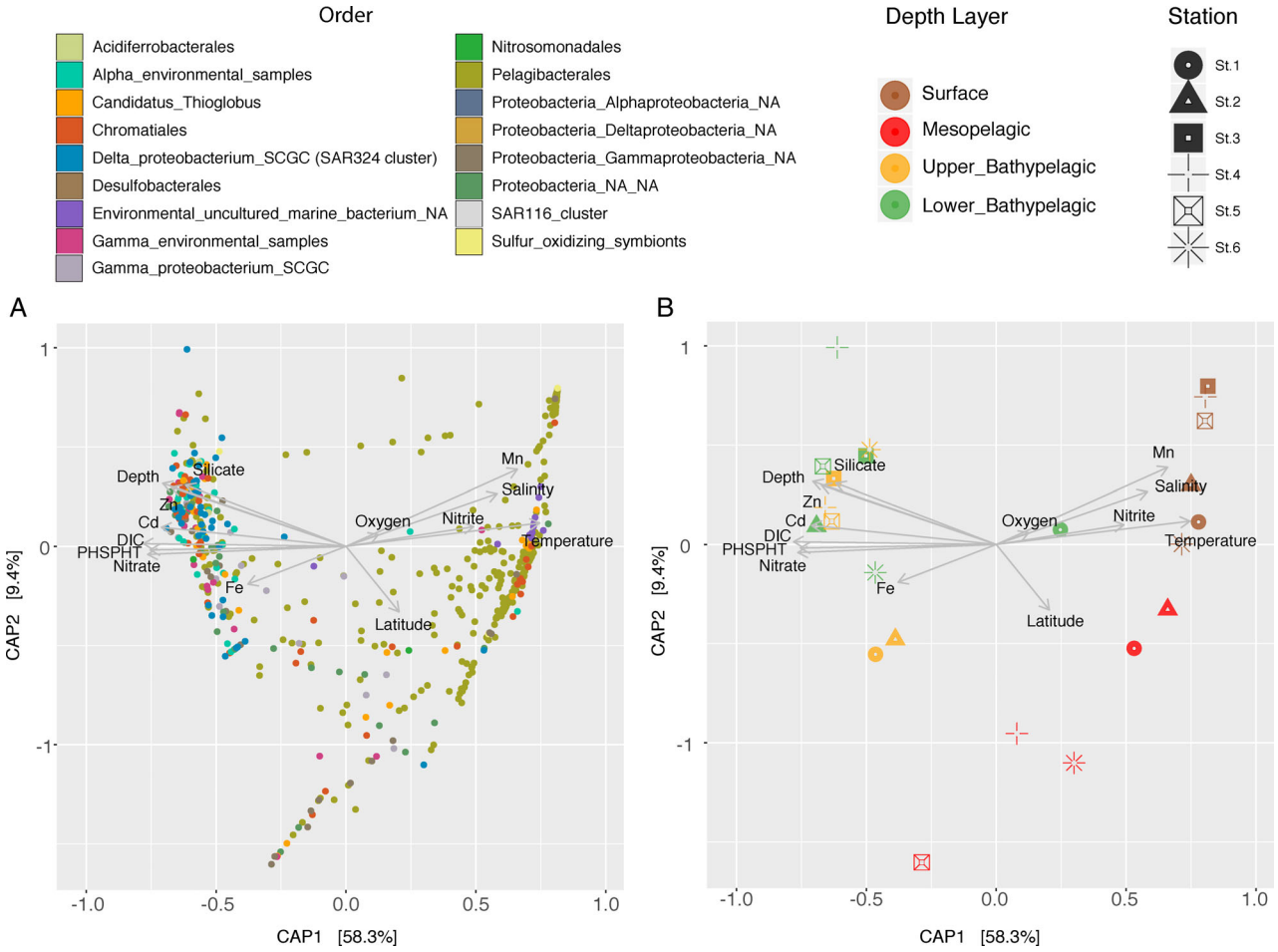

Fig. 6. Constrained principal coordinate analysis (CAP) based on weighted UniFrac distance of the *apr*A harbouring bacterial community showing the environmental parameters related to the bacterial variation. Variation of the most abundant orders (A) and samples (B).

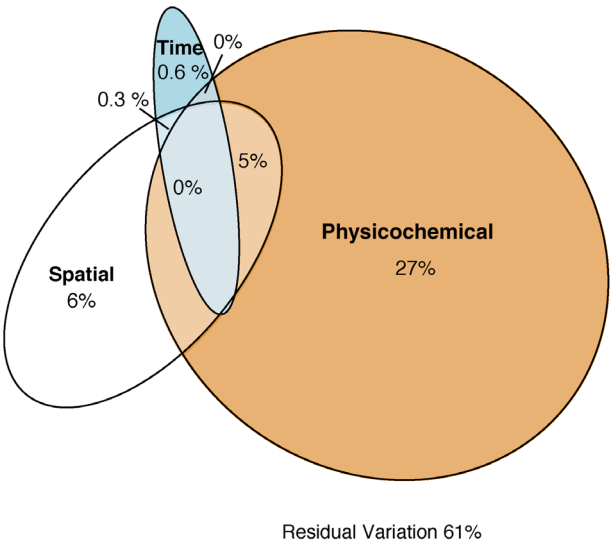

Fig. 7. Variation partitioning of physicochemical, spatial and time factors explaining the *apr*A-harbouring bacterial community composition.

values might be underestimated compared to those of Acinas *et al*. (2021), since we calculated the *apr*A contribution to the community normalizing to total PA (bacteria and archaea), whereas Acinas *et al*. (2021) normalized to *rec*A gene abundance (a housekeeping gene present only in bacteria). Moreover, some underestimation could also be present due to the DNA extraction efficiency when comparing gene abundance with PA determined by flow cytometry.

Chemolithoautotrophy has been demonstrated to occur in the meso- and bathypelagic realms at higher rates than previously assumed (Reinthaler *et al*., 2010). Chemoautotrophs such as ammonia-oxidizing archaea (AOA) show similar biogeographic and depth-related distribution patterns in the Atlantic (Sintes *et al*., 2016) as *apr*A containing prokaryotes (this study). Also, AOA increase in their contribution to the prokaryotic community towards oligotrophic regions and with depth (Sintes *et al*., 2016, Santoro *et al.,* 2017). In contrast, relatively low contributions of AOA to total PA are found in highly

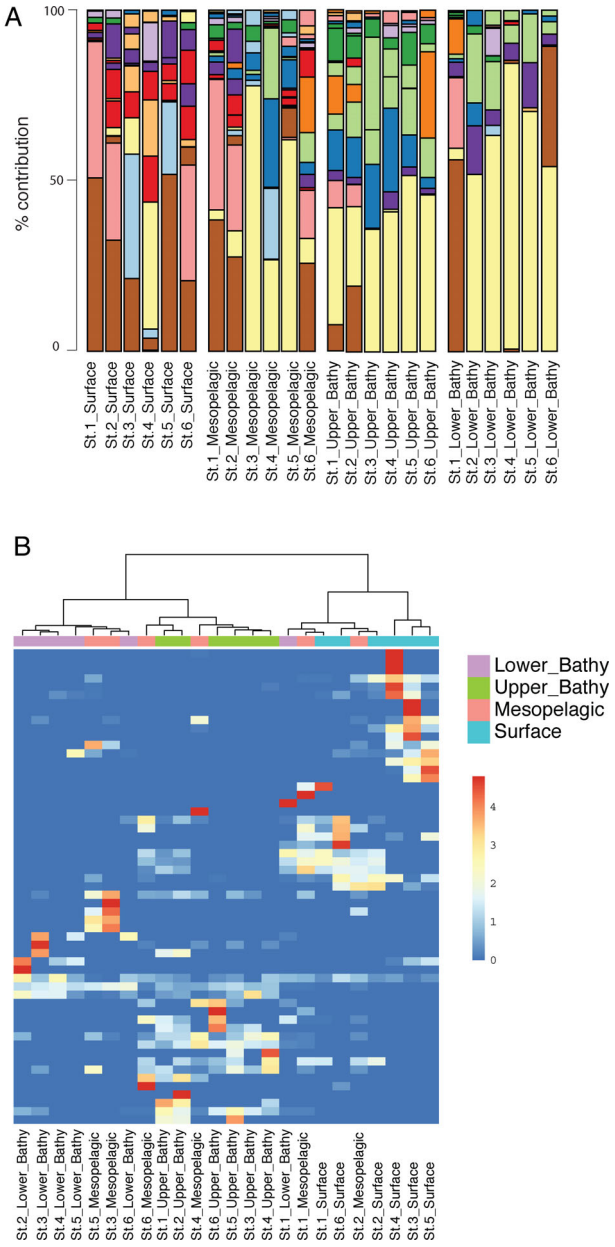

Fig. 8. Distribution of *Pelagibacter* sp. oligotypes at six stations and four depths (A), different colours represent specific oligotypes. Heatmap of the *z*-score distribution of the *Pelagibacter* sp. oligotypes (B). The dendrogram clusters the samples according to Bray–Curtis similarity.

productive regions such as the Arctic and Antarctic (Sintes *et al*., 2016), in agreement with the higher depth-integrated DIC fixation rates measured in the oligotrophic compared to mesotrophic North Atlantic Ocean regions (Reinthaler *et al*., 2010). Taken together, these findings suggest that the availability of organic compounds under mesotrophic conditions favours heterotrophic life strategies over slow-growing chemolithoautotrophic and mixotrophic prokaryotes.

Adenylyl sulfate reductase is present in both sulfate reducers and sulfur oxidizers playing different roles (Meyer and Kuever, 2007b). Therefore, to assess the link between potential chemoautotrophy and the *apr*A containing community, the distribution of the *cbb*M encoding form II ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) enzyme (Swan *et al*., 2011) responsible for DIC fixation was determined. The *cbb*M containing prokaryotes showed a similar distribution pattern as the *apr*A harbouring prokaryotes, exhibiting a higher contribution in oligotrophic regions and the meso- and bathypelagic prokaryotic communities compared to high latitudes and epipelagic communities. The ratio between *apr*A and *cbb*M gene abundance indicates that at least a fraction of the prokaryotic community carrying *apr*A genes oxidizes inorganic sulfur compounds as a source of energy coupled to CO_2 fixation, in agreement with previous findings in the mesopelagic Atlantic (Swan *et al*., 2011).

Multiple forms of RuBisCO occur in prokaryotes (Tabita *et al*., 2008), and some sulfur oxidizers might harbour other types of *cbb*M than *cbb*M form II. Consequently, the potential contribution of DIC fixation associated with sulfur oxidation in the present study is likely a conservative estimate. The ratios of *apr*A/*cbb*M in the mesopelagic and bathypelagic realm (with a median of 1.7 and 4.8 respectively; Fig. 2E and F) support the notion of a larger role of the prokaryotic community carrying *apr*A as chemoautotrophs than in epipelagic waters where the *apr*A/*cbb*M ratio had a median of 31 (Fig. 2E and F). This finding suggests that in the epipelagic communities most adenylyl sulfate reductase activity is not coupled to DIC fixation. The dominance of Pelagibacterales in the *apr*A harbouring community (Fig. 5) and the proposed role of APS reductase in detoxification of sulfite in this taxon (see below) corroborates this notion. High productivity regions such as the high latitude provinces and the equatorial area showed lower *apr*A/*cbb*M ratios than oligotrophic provinces (Fig. 2). The relatively high abundance of the *cbb*M gene in these provinces (ARCT, WTRA and SANT) suggests that the inorganic carbon fixation was coupled to alternative sources of energy (such as ammonia, iron, hydrogen, methane) rather than sulfur. Our findings are in agreement with a previous study conducted in the Atlantic Ocean (Swan *et al*., 2011) showing that not all the SAGs containing RuBisCO (*cbb*M and *cbb*L) also contain sulfur-oxidizing genes.

Adenylyl sulfate reductase can alternatively perform dissimilatory sulfate reduction as a form of anaerobic respiration using sulfate as an electron acceptor. This process is strictly anaerobic and therefore limited to OMZs and anaerobic microenvironments (Ulloa *et al*., 2012; Bianchi *et al*., 2018). Suspended and sinking marine snow can provide anoxic or microaerobic

microenvironments, thus promoting the proliferation of sulfate-reducing bacteria (Ploug *et al*., 1997; Bianchi *et al*., 2018). Therefore, a fraction of the *apr*A harbouring community in the meso- and bathypelagic layers could inhabit these microniches. However, our results indicate that sulfate reducers constitute a minor fraction of the *apr*A containing prokaryotes in the open oxygenated ocean, in agreement with the patchiness of suitable marine snow in these waters (Silver *et al*., 1978).

Major apr*A harbouring taxa in the oxygenated water column of the open Atlantic*

The *apr*A harbouring prokaryotic community shares only a few ASVs and amino acid sequence variants (AASVs) between different depth layers throughout the Atlantic Ocean (Figs 3 and S2), suggesting a clear niche differentiation according to environmental gradients associated to depth layers. The phylogenetic characterization reveals that the *apr*A-containing community in the oxygenated water column is dominated by sulfur-oxidizing bacteria in the open Atlantic (Fig. 5). Neither Archaea nor bacterial sulfate reducers significantly contributed to the *apr*A harbouring community, suggesting a minor role of these groups in the sulfur cycle of the pelagic open ocean. However, the occurrence of anaerobic microenvironments in oxygenated waters, e.g. within marine snow, may sporadically promote the proliferation of sulfate reducers mediating sulfate reduction in the otherwise oxygenated waters (Bianchi *et al*., 2018).

SAR11, one of the most abundant and ubiquitous bacterial groups in marine ecosystems (Morris *et al*., 2002) and of paramount significance in the global ocean biogeochemical cycles, dominates the *apr*A-harbouring community in the epi- and mesopelagic layers (Fig. 3). *Candidatus* Pelagibacter *apr* belongs to the sulfur-oxidizing apr lineage-I (Meyer and Kuever, 2008), however, the lack of other enzymes from the pathway prevents the dissimilatory function in these organisms (Meyer and Kuever, 2007b). Consequently, it has been suggested that the primary role of APS reductase in SAR11 is intracellular detoxification of sulfite, a metabolite of organic sulfur degradation (Meyer and Kuever, 2007b). Indeed, SAR11 clade members are deficient in assimilatory sulfate reduction genes and thus, require organic sulfur compounds for growth (Tripp *et al*., 2008). DMSP, released by phytoplankton in the surface ocean, and taurine, released by zooplankton and other metazoans down to bathypelagic waters, potentially match the sulfur requirements of SAR11 in the ocean (Clifford *et al*., 2019; Durham *et al*., 2019). The reported decreased transcription of *apr*AB in SAR11 under sulfur limited conditions (Smith *et al*., 2016) supports the notion that APS reductase plays a role in detoxification.

Oligotyping provides evidence that there are different SAR11 ecotypes adapted to different environmental conditions. The oligotype composition of the *apr*A-harbouring SAR11 community at the amino-acid level is depth-stratified (Fig. 8), suggesting different adaptations to metabolize sulfur compounds of different origins, such as DMSP in the epipelagic layers or taurine in deeper pelagic zones. Another possibility would be that deep ocean ecotypes might be using inorganic sulfur oxidation as a source of additional energy (i.e. chemolithotrophy). Chemolithotrophy has not been found in *Candidatus* Pelagibacter ubique (Meyer and Kuever, 2007a, 2007b) as a consequence of lacking sulfate adenylyltransferase (*sat*). However, knowledge on the metabolic potential of deep ocean ecotypes is generally limited, and the capacity (or absence of it) for chemolithoheterotrophy based on sulfur oxidation has not been assessed yet.

In the bathypelagic realm, the *apr*A harbouring community was dominated by Chromatiales, Deltaproteobacteria SCGC and SUP05 clade members (Fig. 5). Members of the Chromatiales order are known as anoxygenic phototrophic bacteria that utilize sulfide as an electron donor, and subsequently oxidize it to sulfate (Imhoff, 2014). Nevertheless, a number of species are aerophilic sulfur oxidizers common in illuminated stratified coastal waters (Imhoff, 2014), near hydrothermal vents (Han *et al*., 2018; Pjevac *et al*., 2018) in deep-sea sponges (Kennedy *et al*., 2014) and in deep sediments (Li *et al*., 2013). Members of the Chromatiales order have been also reported throughout the oxygenated water column in the global ocean based on 16S rRNA gene sequencing, e.g. in the Gulf of Alaska (Muck *et al*., 2019), North Atlantic (Frank *et al*., 2016), or the Pacific (Pham *et al*., 2008) with increased relative abundances in meso- and bathypelagic waters. However, our knowledge on their metabolic strategies in the open oxygenated water column remains limited and further studies with isolates are needed to shed light on their metabolic adaptations.

Candidatus Thioglobus (SUP05 clade) together with the uncultivated Deltaproteobacterium SCGC (SAR324 cluster) increased in its contribution to the prokaryotic *apr*A containing community with depth. Deltaproteobacterium SCGC (SAR324), contributing $19 \pm 9\%$ to the bathypelagic *apr*A containing bacterial community, exhibits DIC fixation in conjunction with sulfur oxidation (Swan *et al*., 2011). It contributes significantly to the chemoautotrophy of the deep ocean based on sulfur oxidation. The isolate *Candidatus* Thioglobus singularis PS1 (member of the SUP05 clade) has the potential to use organic substrates and also to perform DIC fixation and sulfur oxidation (Spietz *et al*., 2019b). However, *Candidatus* Thioglobus NP1 isolated from the open ocean lacks the RuBisCO encoding gene in its genome, suggesting that DIC fixation is not performed by all SUP05 (Spietz *et al*., 2019a), in agreement with the high *apr*A/*cbb*M ratios (Fig. 2).

In conclusion, our study showed a clear niche differentiation of the prokaryotes carrying *apr*A genes in the open ocean. The *apr*A harbouring prokaryotic community was dominated by sulfur oxidizers, partly capable of DIC fixation (chemoautotrophy). Adenylyl sulfate reductase putatively assigned to Archaea and sulfate reducers was neglectable in this study. This suggests that in the ocean's oxygenated water column the adenylyl sulfate reductase is mainly used in the chemolithoautotrophic or chemolithoheterotrophic pathways as an energy source or in the biosynthesis pathways. This takes place especially in areas depleted in bioavailable organic carbon, such as oligotrophic regions and/or the deep ocean. Taken together, our results suggest that sulfur oxidizers may play an important role in the sulfur and carbon cycle of the oxygenated waters of the global ocean.

Material and methods

Study area and sampling

Water samples were collected during the GEOTRACES (legs 1–3) cruises on board of R/V *Pelagia* (in April and July 2010) and R/V *James Cook* (in February–March 2011) along a latitudinal transect in the Atlantic Ocean from 64°N to 50°S (Fig. 1). Sampling was performed using a rosette water sampler equipped with 24 25-L Niskin bottles, a CTD (conductivity-temperature-depth; SBE43 Seabird, Bellevue, WA, USA), with additional sensors for chlorophyll fluorescence, turbidity and oxygen concentration. Samples for inorganic nutrients, trace elements and microbial parameters were collected at 24 depth layers and 51 stations. Samples for quantitative PCR analyses (*apr*A and *cbb*M genes) were collected at 6–8 depth layers from epi- to bathypelagic waters. Three pelagic zones were distinguished throughout the depth profiles: epipelagic (50–70 m), mesopelagic (200–1000 m) and bathypelagic (1000–5000 m). The bathypelagic layer was subsequently divided into upper bathypelagic (1000–2000 m) and lower bathypelagic (2000–5000 m) based on the different physicochemical characteristics of these two depth layers, resulting in a total of four pelagic zones sampled. The upper bathypelagic (collected at ~1250 m depth) included samples from Iceland-Scotland Overflow Water, Labrador Sea Water, upper North Atlantic Deep Water (NADW) and Antarctic Intermediate Water masses. The lower bathypelagic layer included samples collected between 2500 and 4500 m depth and included the Antarctic Bottom Water, intermediate NADW and Denmark Strait Overflow Water, depending on the oceanographic region and depth.

Six oceanic provinces were discriminated based on the description of Longhurst (2007): the North Atlantic Arctic province (ARCT; 70°N–55°N), the North Atlantic Drift province (NADR; 55°N–40°N), the North Atlantic Gyre (NAG; 40°N–12°N), the Western Tropical Atlantic province (WTRA; 12°N–6°S), the South Atlantic Gyre (SATL; 6°S–40°S), and the Subantarctic province (SANT; 40°S–55°S) (Fig. 1).

Inorganic nutrients

The concentrations of dissolved inorganic nutrients (NO_3, NO_2, PO_4^{3-}) were determined on 0.2 μm filtered water samples in a TRAACS 800 autoanalyser system immediately after collecting the samples following established protocols (Reinthaler *et al*., 2008).

Microbial counts

The PA was estimated by flow cytometry after nucleic acid staining with Sybr Green I (Giorgio *et al*., 1996; Brussaard, 2004).

DNA extraction and amplification

Two (epipelagic) to ten litres (mesopelagic to lower bathypelagic) of seawater were filtered through 0.22 μm GTTP polycarbonate filters (Millipore) for DNA analyses. The extraction was performed using the Ultraclean soil DNA isolation kit (Mobio) following the manufacturer's protocol.

Quantitative-PCR

q-PCR was used to evaluate the abundance of two *cbb*M genes encoding the RuBisCO form II, subsequently named *cbb*M65 and *cbb*M68 according to the primer sets used (Swan *et al*., 2011), and the *apr*A gene encoding the alpha subunit of APS. *cbb*M65 and *cbb*M68 cover the two main clusters of cbbM (rubisco form II) (Swan *et al*., 2011), and were specifically developed for q-PCR.

q-PCR was conducted in triplicate for all functional genes on a LightCycler 480 SW 1.5 (Roche). The reaction mixture (10 μl) contained 1× LightCycler 480 DNA SYBRGreen I Master (Roche), 0.5 μM of specific primers (Table S2), 1 μl of DNA extract made up to 10 μl with PCR-grade water (Roche). All reactions were performed in 96-well q-PCR plates (Roche) with optical tape. Accumulation of newly amplified double-stranded DNA products was followed by the increase of fluorescence due to the binding of the fluorescent dye SYBRGreen to the newly produced DNA. Specificity of the q-PCR reaction was tested on agarose gel electrophoresis and with a melting curve analysis (65–95°C). Each gene was amplified using specific primer sets and annealing temperatures as described in Table S2, and quantified using specific standards. The q-PCR thermocycling for *cbb*M was composed of an initial denaturation step at 95°C for 10 min and 50 amplification cycles

consisting of denaturation at 95°C for 5 s, annealing at the specific primer temperature (Table S2) for 5 s, extension at 72°C for 15 s and plate reading at 76°C for 3 s. The average q-PCR efficiencies for the *cbb*M65 and *cbb*M68 genes were 73% and 91% respectively. The total *cbb*M abundance was calculated as the sum of *cbb*M65 and *cbb*M68 gene abundance. The q-PCR cycling for *apr*A consisted of an initial denaturation step at 95°C held for 10 min, followed by 50 amplification cycles consisting of denaturation at 95°C for 5 s, annealing at 60°C for 15 s, extension at 72°C for 30 s and plate read at 80°C for 3 s. The average efficiency of the q-PCR for *apr*A gene was 81%. The specificity of the q-PCR reaction for the different genes was checked with melting curve analyses (65–95°C).

Sequencing and bioinformatics analyses of the bacterial apr*A genes*

Samples were collected at four depth layers covering the epipelagic to deep bathypelagic realm at six stations corresponding to different geographic provinces to determine the composition of the prokaryotic *apr*A harbouring community. The *apr*A gene was amplified using *apr*A-1F and *apr*A-5R primers (Meyer and Kuever, 2007b) (Table S2). The thermocycling PCR conditions were as follows: initial denaturation at 94°C for 4 min; amplification: 30 cycles, at 94°C for 1 min, 55°C for 1 min, and extension at 72°C for 1 min, followed by a final extension step at 72°C for 7 min and holding at 4°C. The PCR products were purified using PCRExtract MiniKit (5-PRIME) and quality checked in a 2% agarose gel. The purified PCR products were subsequently sequenced with Illumina Miseq (IMGM Laboratories GmbH) with V2 chemistry. The data are publicly available in the NCBI database under the accession number PRJNA641589.

The raw reads were quality-filtered and processed using dada2 version 2.1 (Callahan *et al*., 2016b). The obtained ASVs were taxonomically assigned using a naïve Bayesian classifier method with a minimum bootstrap threshold of 40 (implemented in dada2) against a reference *apr*A database. The database was built using *apr*A gene sequences obtained from NCBI database closely related to the ASVs according to blast results and of known taxa (Supplementary Table S3). The obtained data were subsequently analysed with phyloseq R package (McMurdie and Holmes, 2013; Callahan *et al*., 2016a). *Apr*A gene sequences were further classified using custom phylogenetic trees. The ASVs were translated to amino acid sequences, and subsequently, CD-HIT (Huang *et al*., 2010) was used to cluster the amino acid sequences at 100% sequence identity. We named the resulting amino acid cluster sequences AASVs. The AASVs were aligned together with amino acid sequences obtained from NCBI database using MEGA-X (Kumar *et al*., 2018). The phylogenetic tree of the *apr*A AASVs was built using the Neighbour-Joining method (Saitou and Nei, 1987) with 100 bootstraps in MEGA-X. The resulting tree was subsequently drawn using iTOL (Letunic and Bork, 2019) (Fig. S2).

Shared and unique *apr*A ASVs in different depth layers were determined with the 'Venn' package in R software using presence/absence data. Richness and diversity indexes were calculated with phyloseq R package. Weighted UniFrac distances and principal coordinate analysis (Lozupone *et al*., 2011) were calculated with ampvis2 R package and used to compare the *apr*A community composition between samples (Andersen *et al*., 2018). The cumulative read abundance was also estimated using ampvis2 R package. ASVs reported in more than 80% of the samples from a specific depth layer were considered core phylotypes of that depth layer, ASVs reported in 50%–79% of the samples were considered resident phylotypes.

The composition of the *apr*A containing prokaryotic community and its relationship with the environmental variables was analysed by CAP implemented in phyloseq R package using weighted UniFrac distance data. To reveal the major environmental factors affecting the community composition, variance partitioning was performed on the Hellinger-transformed dataset using the varpart function of vegan (Oksanen *et al*., 2018). The variance partitioning results were visualized with eulerr R package (Larsson, 2020).

Oligotyping analysis of *apr*A assigned to the *Candidatus* Pelagibacter family at the amino acid level was conducted following the pipeline available at http://oligotyping.org (Eren *et al*., 2013).

Statistical analyses

Relationships between gene abundances (*apr*A, *cbb*M) and the *apr*A/*cbb*M ratio with environmental variables were analysed using a dbRDA with the 'Vegan' package (Oksanen *et al*., 2018) implemented in R software v.3.4. Environmental variables included latitude (absolute), depth, temperature, salinity, fluorescence, dissolved oxygen, DIC, macronutrients (phosphate, silicate, nitrite and nitrate) and trace elements (Cd, Fe, Mn, Zn). Micronutrients, such as trace metals, play an important role for the functioning of proteins and enzymes. Adenylyl-sulfate reductase consists of two subunits (alpha and beta), the beta subunit contains Fe. Other metals serve as cofactors during the sulfur oxidation processes. Moreover, during sulfate reduction and production of sulfide, Cd precipitates as CdS. Therefore, we decided to test whether the *apr*A gene abundance was influenced by the availability of specific micronutrients (such as trace

metals) as they play an important role in the sulfur oxidation and sulfate reduction processes.

The response variables were normalized and Euclidean distance similarities matrices were calculated. The *F*-test and the *p*-value were obtained applying 999 permutations. The dataset of environmental parameters and the corresponding methods used for the measurements are available at the GEOTRACES website (http://www.geotraces.org) (Mawji *et al*., 2015).

Analysis of variance (ANOVA test) was performed to test possible variations within different depth layers and/or oceanic regions.

Acknowledgements

We thank the captain and crew of the R/V *Pelagia* for their support and splendid atmosphere on board. Shiptime was provided by the Geotraces program. Laboratory work was supported by the Austrian Science Fund (FWF) projects: I486-B09, Z194-B17, P28781-B21 and by the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013)/ERC grant agreement No. 268595 (MEDEA project) to G.J.H. D.D.C. was supported as an overseas researcher under the Postdoctoral Fellowship of Japan Society for Promotion of Science (P16085). E.S. was supported by the Austrian Science Fund (FWF) project P27696-B22. The valuable comments of the reviewers helped to improve the quality of the paper.

Data Availability Statement

The data are publicly available in the NCBI database under the accession number PRJNA641589.

References

Acinas, S.G., Sánchez, P., Salazar, G., Cornejo-Castillo, F.M., Sebastián, M., Logares, R., *et al*. (2021) Deep ocean metagenomes provide insight into the metabolic architecture of bathypelagic microbial communities. *Commun Biol* **4**: 604.

Anantharaman, K., Breier, J.A., Sheik, C.S., and Dick, G.J. (2013) Evidence for hydrogen oxidation and metabolic plasticity in widespread deep-sea sulfur-oxidizing bacteria. *Proc Natl Acad Sci U S A* **110**: 330–335.

Andersen, K.S., Kirkegaard, R.H., Karst, S.M., and Albertsen, M. (2018) ampvis2: an R package to analyse and visualise 16S rRNA amplicon data. *bioRxiv*: 299537.

Aoki, M., Kakiuchi, R., Yamaguchi, T., Takai, K., Inagaki, F., and Imachi, H. (2015) Phylogenetic diversity of aprA genes in subseafloor sediments on the northwestern Pacific margin off Japan. *Microbes Environ* **30**: 276–280.

Arrieta, J.M., Mayol, E., Hansman, R.L., Herndl, G.J., Dittmar, T., and Duarte, C.M. (2015) Dilution limits dissolved organic carbon utilization in the deep ocean. *Science* **348**: 331–333.

Azam, F., Fenchel, T., Field, J., Gray, J., Meyer-Reil, L., and Thingstad, F. (1983) The ecological role of water-column microbes in the sea. *Mar Ecol Prog Ser* **10**: 257–263.

Baltar, F., Arístegui, J., Sintes, E., Gasol, J.M., Reinthaler, T., and Herndl, G.J. (2010) Significance of non-sinking particulate organic carbon and dark CO_2 fixation to heterotrophic carbon demand in the mesopelagic northeast Atlantic. *Geophys Res Lett* **37**: L09602.

Bianchi, D., Weber, T.S., Kiko, R., and Deutsch, C. (2018) Global niche of marine anaerobic metabolisms expanded by particle microenvironments. *Nat Geosci* **11**: 263–268.

Blazejak, A., Kuever, J., Erseus, C., Amann, R., and Dubilier, N. (2006) Phylogeny of 16S rRNA, ribulose 1,5-bisphosphate carboxylase/oxygenase, and adenosine 5′-phosphosulfate reductase genes from gamma- and alphaproteobacterial symbionts in gutless marine worms (oligochaeta) from Bermuda and The Bahamas. *Appl Environ Microbiol* **72**: 5527–5536.

Blazejak, A., and Schippers, A. (2011) Real-time PCR quantification and diversity analysis of the functional genes aprA and dsrA of sulfate-reducing prokaryotes in marine sediments of the Peru continental margin and the Black Sea. *Front Microbiol* **2**: 253.

Brussaard, C.P. (2004) Optimization of procedures for counting viruses by flow cytometry. *Appl Environ Microbiol* **70**: 1506–1513.

Buesseler, K.O., and Boyd, P.W. (2009) Shedding light on processes that control particle export and flux attenuation in the twilight zone of the open ocean. *Limnol Oceanogr* **54**: 1210–1232.

Buesseler, K.O., Lamborg, C.H., Boyd, P.W., Lam, P.J., Trull, T.W., Bidigare, R.R., *et al*. (2007) Revisiting carbon flux through the ocean's twilight zone. *Science* **316**: 567–570.

Callahan, B.J., Sankaran, K., Fukuyama, J.A., McMurdie, P.J., and Holmes, S.P. (2016). Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. *F1000Research*, **5**: 1492.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J., and Holmes, S.P. (2016b) DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**: 581–583.

Canfield, D.E., Erik, K., and Bo, T. (2005) The sulfur cycle. In *Advances in Marine Biology*, Vol. **48**, Canfield, D.E., Kristensen, E., and Thamdrup, B. (eds), pp. 313–381. London: Academic Press.

Clifford, E.L., Varela, M.M., De Corte, D., Bode, A., Ortiz, V., Herndl, G.J., and Sintes, E. (2019) Taurine is a major carbon and energy source for marine prokaryotes in the North Atlantic Ocean off the Iberian Peninsula. *Microb Ecol* **78**: 299–312.

Durham, B.P., Boysen, A.K., Carlson, L.T., Groussman, R.D., Heal, K.R., Cain, K.R., *et al*. (2019) Sulfonate-based networks between eukaryotic phytoplankton and heterotrophic bacteria in the surface ocean. *Nat Microbiol* **4**: 1706–1715.

Eren, A.M., Maignien, L., Sul, W.J., Murphy, L.G., Grim, S.L., Morrison, H.G., and Sogin, M.L. (2013) Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol Evol* **4**: 1111–1119.

Frank, A.H., Garcia, J.A.L., Herndl, G.J., and Reinthaler, T. (2016) Connectivity between surface and deep waters determines prokaryotic diversity in the North Atlantic deep water. *Environ Microbiol* **18**: 2052–2063.

Frank, K.L., Rogers, D.R., Olins, H.C., Vidoudez, C., and Girguis, P.R. (2013) Characterizing the distribution and

rates of microbial sulfate reduction at Middle Valley hydrothermal vents. *ISME J* **7**: 1391–1401.

Fritz, G., Büchert, T., Huber, H., Stetter, K.O., and Kroneck, P. M.H. (2000) Adenylylsulfate reductases from archaea and bacteria are 1:1 αβ-heterodimeric iron–sulfur flavoenzymes – high similarity of molecular properties emphasizes their central role in sulfur metabolism. *FEBS Lett* **473**: 63–66.

Giorgio, P.A.d., Bird, D.F., Prairie, Y.T., and Planas, D. (1996) Flow cytometric determination of bacterial abundance in lake plankton with the green nucleic acid stain SYTO 13. *Limnol Oceanogr* **41**: 783–789.

Han, Y., Gonnella, G., Adam, N., Schippers, A., Burkhardt, L., Kurtz, S., *et al*. (2018) Hydrothermal chimneys host habitat-specific microbial communities: analogues for studying the possible impact of mining seafloor massive sulfide deposits. *Sci Rep* **8**: 10386.

Hansell, D.A. (2013) Recalcitrant dissolved organic carbon fractions. *Ann Rev Mar Sci* **5**: 421–445.

Herndl, G.J., and Reinthaler, T. (2013) Microbial control of the dark end of the biological pump. *Nat Geosci* **6**: 718–724.

Herndl, G.J., Reinthaler, T., Teira, E., van Aken, H., Veth, C., Pernthaler, A., and Pernthaler, J. (2005) Contribution of archaea to total prokaryotic production in the deep Atlantic Ocean. *Appl Environ Microbiol* **71**: 2303–2309.

Hipp, W.M., Pott, A.S., Thum-Schmitz, N., Faath, I., Dahl, C., and Trüper, H.G. (1997) Towards the phylogeny of APS reductases and sirohaem sulfite reductases in sulfate-reducing and sulfur-oxidizing prokaryotes. *Microbiology* **143**: 2891–2902.

Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. (2010) CD-HIT suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26**: 680–682.

Hügler, M., Gärtner, A., and Imhoff, J.F. (2010) Functional genes as markers for sulfur cycling and CO_2 fixation in microbial communities of hydrothermal vents of the Logatchev field. *FEMS Microbiol Ecol* **73**: 526–537.

Imhoff, J.F. (2014) The family Chromatiaceae. In *The Prokaryotes: Gammaproteobacteria*, Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., and Thompson, F. (eds). Berlin, Heidelberg: Springer, pp. 151–178.

Kappler, U., and Dahl, C. (2001) Enzymology and molecular biology of prokaryotic sulfite oxidation1. *FEMS Microbiol Lett* **203**: 1–9.

Kennedy, J., Flemer, B., Jackson, S.A., Morrissey, J.P., O'Gara, F., and Dobson, A.D. (2014) Evidence of a putative deep sea specific microbiome in marine sponges. *PLoS One* **9**: e91092.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* **35**: 1547–1549.

Lampreia, J., Pereira, A.S., and Moura, J.G. (1994) [16] Adenylylsulfate reductases from sulfate-reducing bacteria. In *Methods in Enzymology*, Vol. **243**, pp. 241–260. New York: Academic Press.

Larsson, J. (2020) eulerr: Area-proportional Euler and Venn diagrams with ellipses.

Letunic, I., and Bork, P. (2019) Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**: W256–W259.

Li, G., Lai, Q., Liu, X., Sun, F., Du, Y., Li, G., and Shao, Z. (2013) Maricoccus atlantica gen. nov. sp. nov., isolated from deep sea sediment of the Atlantic Ocean. *Antonie Van Leeuwenhoek* **104**: 1073–1081.

Li, Y., Jing, H., Xia, X., Cheung, S., Suzuki, K., and Liu, H. (2018) Metagenomic insights into the microbial community and nutrient cycling in the Western subarctic Pacific Ocean. *Front Microbiol* **9**: 623.

Longhurst, A.R. (2007) . In *Ecological Geography of the Sea*, 2nd ed, Longhurst, A.R. (ed). Burlington: Academic Press.

Lozupone, C., Lladser, M.E., Knights, D., Stombaugh, J., and Knight, R. (2011) UniFrac: an effective distance metric for microbial community comparison. *ISME J* **5**: 169–172.

Mawji, E., Schlitzer, R., Dodas, E.M., Abadie, C., Abouchami, W., Anderson, R.F., *et al*. (2015) The GEOTRACES intermediate data product 2014. *Mar Chem* **177**: 1–8.

McMurdie, P.J., and Holmes, S.P. (2013) Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**: e61217.

Meyer, B., and Kuever, J. (2007a) Molecular analysis of the diversity of sulfate-reducing and sulfur-oxidizing prokaryotes in the environment, using aprA as functional marker gene. *Appl Environ Microbiol* **73**: 7664–7679.

Meyer, B., and Kuever, J. (2007b) Phylogeny of the alpha and beta subunits of the dissimilatory adenosine-5′-phosphosulfate (APS) reductase from sulfate-reducing prokaryotes—origin and evolution of the dissimilatory sulfate-reduction pathway. *Microbiology* **153**: 2026–2044.

Meyer, B., and Kuever, J. (2008) Homology modeling of dissimilatory APS reductases (AprBA) of sulfur-oxidizing and sulfate-reducing prokaryotes. *PLoS One* **3**: e1514.

Morris, R.M., Rappe, M.S., Connon, S.A., Vergin, K.L., Siebold, W.A., Carlson, C.A., and Giovannoni, S.J. (2002) SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**: 806–810.

Muck, S., De Corte, D., Clifford, E.L., Bayer, B., Herndl, G.J., and Sintes, E. (2019) Niche differentiation of aerobic and anaerobic ammonia oxidizers in a high latitude deep oxygen minimum zone. *Front Microbiol* **10**: 2141.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Simpson, G.L., Solymos, P. et al. (2018) Vegan: community ecology package. URL https://cran.r-project.org/web/packages/vegan/.

Pham, V.D., Konstantinidis, K.T., Palden, T., and DeLong, E. F. (2008) Phylogenetic analyses of ribosomal DNA-containing bacterioplankton genome fragments from a 4000 m vertical profile in the North Pacific Subtropical Gyre. *Environ Microbiol* **10**: 2313–2330.

Pjevac, P., Meier, D.V., Markert, S., Hentschker, C., Schweder, T., Becher, D., *et al*. (2018) Metaproteogenomic profiling of microbial communities colonizing actively venting hydrothermal chimneys. *Front Microbiol* **9**: 680.

Ploug, H., Kühl, M., Buchholz-Cleven, B., and Jørgensen, B.B. (1997) Anoxic aggregates – an ephemeral phenomenon in the pelagic environment? *Aquat Microb Ecol* **13**: 285–294.

Radford-Knoery, J., German, C.R., Charlou, J.-L., Donval, J.-P., and Fouquet, Y. (2001) Distribution and behavior of dissolved hydrogen sulfide in hydrothermal plumes. *Limnol Oceanogr* **46**: 461–464.

Reinthaler, T., Sintes, E., and Herndl, G.J. (2008) Dissolved organic matter and bacterial production and respiration in the sea-surface microlayer of the open Atlantic and the western Mediterranean Sea. *Limnol Oceanogr* **53**: 122–136.

Reinthaler, T., van Aken, H.M., and Herndl, G.J. (2010) Major contribution of autotrophy to microbial carbon cycling in the deep North Atlantic's interior. *Deep-Sea Res II Top Stud Oceanogr* **57**: 1572–1580.

Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406–425.

Santoro, A.E., Saito, M.A., Goepfert, T.J., Lamborg, C.H., Dupont, C.L., DiTullio, G.R. (2017). Thaumarchaeal ecotype distributions across the equatorial Pacific Ocean and their potential roles in nitrification and sinking flux attenuation. *Limnol Oceanogr* **62**: 1984–2003.

Shanks, A.L., and Reeder, M.L. (1993) Reducing microzones and sulfide production in marine snow. *Mar Ecol Prog Ser* **96**: 43–47.

Sievert, S.M., Kiene, R.P., and Schulz-Vogt, H.N. (2007) The sulfur cycle. *Oceanography* **20**: 117–123.

Silver, M.W., Shanks, A.L., and Trent, J.D. (1978) Marine snow: microplankton habitat and source of small-scale patchiness in pelagic populations. *Science* **201**: 371–373.

Sintes, E., Bergauer, K., De Corte, D., Yokokawa, T., and Herndl, G.J. (2013) Archaeal amoA gene diversity points to distinct biogeography of ammonia-oxidizing Crenarchaeota in the ocean. *Environ Microbiol* **15**: 1647–1658.

Sintes, E., De Corte, D., Haberleitner, E., and Herndl, G.J. (2016) Geographic distribution of archaeal ammonia oxidizing ecotypes in the Atlantic Ocean. *Front Microbiol* **7**: 77.

Smith, D.P., Nicora, C.D., Carini, P., Lipton, M.S., Norbeck, A.D., Smith, R.D., and Giovannoni, S.J. (2016) Proteome remodeling in response to sulfur limitation in "Candidatus Pelagibacter ubique". *mSystems* **1**: e00068-16.

Spietz, R.L., Lundeen, R.A., Zhao, X., Nicastro, D., Ingalls, A.E., and Morris, R.M. (2019b) Heterotrophic carbon metabolism and energy acquisition in Candidatus Thioglobus singularis strain PS1, a member of the SUP05 clade of marine Gammaproteobacteria. *Environ Microbiol* **21**: 2391–2401.

Spietz, R.L., Marshall, K.T., Zhao, X., and Morris, R.M. (2019a) Complete genome sequence of "Candidatus Thioglobus sp." strain NP1, an open-ocean isolate from the SUP05 clade of marine Gammaproteobacteria. *Microbiol Resour Announc* **8**: e00097-19.

Swan, B.K., Martinez-Garcia, M., Preston, C.M., Sczyrba, A., Woyke, T., Lamy, D., *et al*. (2011) Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean. *Science* **333**: 1296–1300.

Tabita, F.R., Satagopan, S., Hanson, T.E., Kreel, N.E., and Scott, S.S. (2008) Distinct form I, II, III, and IV rubisco proteins from the three kingdoms of life provide clues about rubisco evolution and structure/function relationships. *J Exp Bot* **59**: 1515–1524.

Tripp, H.J., Kitner, J.B., Schwalbach, M.S., Dacey, J.W.H., Wilhelm, L.J., and Giovannoni, S.J. (2008) SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* **452**: 741–744.

Ulloa, O., Canfield, D.E., DeLong, E.F., Letelier, R.M., and Stewart, F.J. (2012) Microbial oceanography of anoxic oxygen minimum zones. *Proc Natl Acad Sci U S A* **109**: 15996–16003.

Wagner, S., Schubotz, F., Kaiser, K., Hallmann, C., Waska, H., Rossel, P.E., *et al*. (2020) Soothsaying DOM: a current perspective on the future of oceanic dissolved organic carbon. *Front Mar Sci* **7**: 341.

Walsh, D.A., Zaikova, E., Howes, C.G., Song, Y.C., Wright, J.J., Tringe, S.G., *et al*. (2009) Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones. *Science* **326**: 578–582.

Watanabe, T., Kojima, H., Takano, Y., and Fukui, M. (2013) Diversity of sulfur-cycle prokaryotes in freshwater lake sediments investigated using aprA as the functional marker gene. *Syst Appl Microbiol* **36**: 436–443.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. Ratios *apr*A/prokaryotic cell abundance (a), *cbb*M/prokaryotic cell abundance (c) throughout the Atlantic. The gene abundance was determined by q-PCR, while the microbial abundance was determined by flow cytometry. Box-plot showing *apr*A/prokaryotic cell abundance (b), *cbb*M/prokaryotic cell abundance (d) ratios in different regions (colour coded) and depth layers. The left and right of the box represent the first and the third quartiles respectively, while the thick vertical line represents the median. Outliers (larger black points) and jitter values (smaller black points) are shown.

Fig. S2. Phylogenetic tree of the *apr*A AASVs (amino acid sequence variants, see Methods). One representative of each amino acid sequence group 100% identical is shown. The inner ring shows the presence/absence of an AASVs at a specific depth, and the outer ring represents their relative abundance. The colours indicate the depth layers (light green: surface; dark green: mesopelagic; light blue: upper bathypelagic; dark blue: lower bathypelagic). Bootstrap values > than 0.5 are indicated by the grey circle at the branches.

Table S1. Average, standard deviation (SD) and median of aprA and cbbM genes, aprA/ prokaryotic cell abundance, cbbM/prokaryotic cell abundance and aprA/cbbM ratios in epipelagic, mesopelagic, upper and lower-bathypelagic layers throughout the Atlantic Ocean.

Table S2. Primers used to amplify the two *cbb*M genes, subsequently termed *cbb*M65 and *cbb*M68, and *apr*A gene fragments for q-PCR and sequencing.

Table S3. Reference sequences obtained from NCBI database (closely related to the ASVs according to blast results) utilized to taxonomically assign the obtained ASVs.