Taxon-Specific Shifts in Bacterial and Archaeal Transcription of Dissolved Organic Matter Cycling Genes in a Stratified Fjord

Benjamin Pontiller,[a,b] Clara Pérez-Martínez,[a] Carina Bunse,[a,c,d] Christofer M. G. Osbeck,[a] José M. González,[e] Daniel Lundin,[a] Jarone Pinhassi[a]

[a]Centre for Ecology and Evolution in Microbial Model Systems—EEMiS, Linnaeus University, Kalmar, Sweden
[b]GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany
[c]Helmholtz-Institute for Functional Marine Biodiversity at the University of Oldenburg (HIFMB), Oldenburg, Germany
[d]Institute for the Chemistry and Biology of the Marine Environment, University of Oldenburg, Oldenburg, Germany
[e]Department of Microbiology, University of La Laguna, La Laguna, Tenerife, Spain

ABSTRACT A considerable fraction of organic matter derived from photosynthesis in the euphotic zone settles into the ocean's interior and, as it progresses, is degraded by diverse microbial consortia that utilize a suite of extracellular enzymes and membrane transporters. Still, the molecular details that regulate carbon cycling across depths remain little explored. As stratification in fjords has made them attractive models to explore patterns in biological oceanography, we here analyzed bacterial and archaeal transcription in samples from five depth layers in the Gullmar Fjord, Sweden. Transcriptional variation over depth correlated with gradients in chlorophyll *a* and nutrient concentrations. Differences in transcription between sampling dates (summer and early autumn) were strongly correlated with ammonium concentrations, which potentially was linked with a stronger influence of (micro-)zooplankton grazing in summer. Transcriptional investment in carbohydrate-active enzymes (CAZymes) decreased with depth and shifted toward peptidases, partly a result of elevated CAZyme transcription by *Flavobacteriales*, *Cellvibrionales*, and *Synechococcales* at 2 to 25 m and a dominance of peptidase transcription by *Alteromonadales* and *Rhodobacterales* from 50 m down. In particular, CAZymes for chitin, laminarin, and glycogen were important. High levels of transcription of ammonium transporter genes by *Thaumarchaeota* at depth (up to 18% of total transcription), along with the genes for ammonia oxidation and CO_2 fixation, indicated that chemolithoautotrophy contributed to the carbon flux in the fjord. The taxon-specific expression of functional genes for processing of the marine pool of dissolved organic matter and inorganic nutrients across depths emphasizes the importance of different microbial foraging mechanisms over spatiotemporal scales for shaping biogeochemical cycles.

IMPORTANCE It is generally recognized that stratification in the ocean strongly influences both the community composition and the distribution of ecological functions of microbial communities, which in turn are expected to shape the biogeochemical cycling of essential elements over depth. Here, we used metatranscriptomics analysis to infer molecular detail on the distribution of gene systems central to the utilization of organic matter in a stratified marine system. We thereby uncovered that pronounced shifts in the transcription of genes encoding CAZymes, peptidases, and membrane transporters occurred over depth among key prokaryotic orders. This implies that sequential utilization and transformation of organic matter through the water column is a key feature that ultimately influences the efficiency of the biological carbon pump.

KEYWORDS marine bacteria, metatranscriptomics, dissolved organic carbon, carbohydrate-active enzymes, peptidases, transporters, vertical depth gradients, stratification, *Nitrosopumilus*, fjord

Major portions of the primary production in the photic zone—up to ~40% of the photosynthetically fixed carbon—is transported vertically in the form of particulate organic matter into the ocean's interior in a process referred to as the biological carbon pump (1). This sinking organic matter is degraded by heterotrophic bacteria via extracellular enzymes that remineralize large proportions to carbon dioxide (2). Sinking particles cross steep gradients in light, temperature, nutrients, and hydrostatic pressure (3). Commonly, density gradients limit mixing between water masses, which disrupts the connectivity of microbial communities and nutrient fluxes. Stratification thereby strongly influences both the microbial community composition and the ecological function of these communities (4). This is, for example, visible in the replacement of phototrophy genes dominating in surface waters by chemolithoautotrophy genes at depth (5–9). This suggests that a pronounced variability in the genetic repertoire of bacteria and archaea is involved in the processing and uptake of nutrients and organic matter throughout the water column in the open ocean.

In the ocean, the bulk, i.e., community level, extracellular enzymatic hydrolysis rates rapidly decrease from epipelagic to bathypelagic zones, whereas the per-cell rates increase, indicating an increased microbial reliance on high-molecular-weight dissolved organic matter (HMW-DOM) with depth (10, 11). Actually, the hydrolysis of HMW-DOM is considered the rate-limiting step in the marine carbon cycle (12), and bacteria secrete hydrolytic enzymes to utilize particulate organic matter and biopolymers (13), e.g., carbohydrate-active enzymes (CAZymes) and peptidases (PEPs), which cleave HMW-DOM into molecules smaller than ~600 Da that can be transported through the cell membrane (14). A few recent studies applied metagenomics to study the spatial and vertical distribution of CAZyme and PEP genes (15–18). For instance, analysis of 94 metagenome-assembled genomes (MAGs) from the Mediterranean Sea uncovered a pronounced depth-related taxonomic and functional specialization in degradation of polysaccharides dominated by *Bacteroidetes*, *Verrucomicrobia*, and *Cyanobacteria* (16). Zhao and colleagues (18) applied a multi-omics approach encompassing a broad spatial and vertical coverage, showing that both the abundance and diversity of dissolved enzymes excreted by particle-attached prokaryotes consistently increased from epi- to bathypelagic waters. Still, knowledge of the concrete expression of these enzymes by different taxa through the water column is limited.

Fjord ecosystems exhibit pronounced vertical gradients in physicochemical and biological conditions in qualitative aspects reflecting gradients observed in other stratified coastal and offshore marine waters. Due to the reduced scaling of conditions in space and time, fjords are at times referred to as model oceans (19–21). Kristineberg Marine Research Station, created in 1877 and located in the Gullmar Fjord, a sill fjord on the west coast of Sweden, is one of the oldest marine research stations in the world. The extensive knowledge of the physical, chemical, and biological oceanography in the fjord provides a solid background against which to determine aspects of the microbial oceanography (22). Therefore, to obtain novel mechanistic knowledge of the functional degradation of biopolymers with depth and the microbial taxa that produce the required enzymes, we applied environmental metatranscriptomics to investigate the expression of polymer-degrading enzyme systems (i.e., CAZymes and PEPs). As we aimed at studying the combined responses of both degradation and uptake of ecologically important DOM compounds or nutrients in this system, we included membrane transporters in the analysis. We hypothesized that potential divergence in CAZyme, PEP, and transporter expression would be associated with shifts in dominance of transcription levels among taxa across depths.

RESULTS AND DISCUSSION

Variability of biotic and abiotic parameters. Seawater samples were collected in duplicate from five depth layers (2 to 5, 15 to 25, 50 to 55, 75, and 100 m) in July and September 2017 at station Alsbäck (bottom depth, ~120 m) in the Gullmar Fjord, Sweden (Fig. 1A and B). Analysis of conductivity-temperature-depth (CTD) profiles
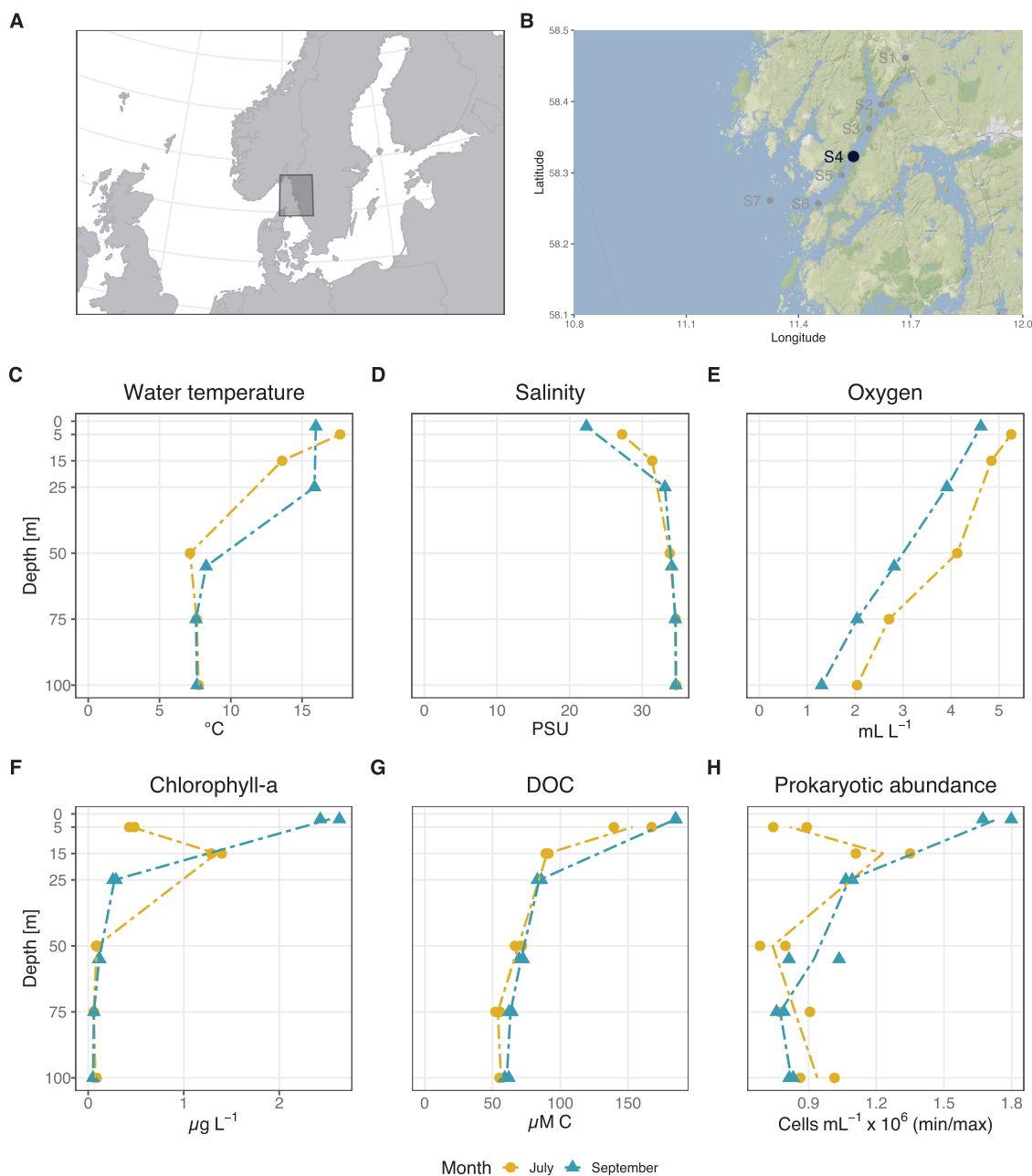
FIG 1 Overview of study site and water column characteristics. (A and B) Locations of the Gullmar Fjord on the Swedish west coast (A) and the sampling site at station 4 (Alsbäck) with a bottom depth of 120 m (B). (C to H) Depth profiles of temperature (C), salinity (D), oxygen levels (E), chlorophyll *a* (B), dissolved organic carbon (DOC) (G), and prokaryotic abundance (H) in July and September.

showed a strong stratification of the water column in both July and September (Fig. 1C to E). The water temperature decreased rapidly from 13 to 17°C at the surface (2 to 25 m) to ~8°C from 50 m and below. In September, temperatures were similar, at 16°C from 2 to 25 m and ~8°C from 50 m (Fig. 1C). The surface salinity in July was 28 practical salinity units (PSU) and decreased to 22 PSU in September and at both times increased with depth to ~36 PSU at 100 m (Fig. 1D), which corresponds to North Sea values (23). Dissolved oxygen levels steadily decreased with depth from oxic (~4.5 to 5.2 ml liter^{-1}) at the surface to dysoxic (1 to 2 ml liter^{-1}) at the bottom (Fig. 1E). The vertical distribution of chlorophyll *a* (Chl *a*) concentrations differed substantially between months. In July, Chl *a* at the surface reached ~0.5 μg liter^{-1}, increased to

$\sim$1.2 μg liter^{-1} in the Chl a maximum at 15 m, and decreased to $\sim$0.1 μg liter^{-1} below 50 m depth (Fig. 1F). In September, concentrations were higher at the surface, reaching $\sim$2.8 μg liter^{-1}, but rapidly decreased to below $\sim$0.2 μg liter^{-1} from 25 m downward (Fig. 1F). Interestingly, the variability between months seen for Chl a was not reflected in the distribution of dissolved organic carbon (DOC) concentrations, which remained at maximum levels ($\sim$170 μM C) at the surface and decreased $\sim$3-fold to 100 m (Fig. 1G). Inorganic nutrients like $NO_3^+ + NO_2^+$ and PO_4^{3-} increased from <1 μM to elevated levels at depth (up to 40 and 4.6 μM, respectively), although, for example, total nitrogen ($\sim$35 μM) remained relatively constant (see Fig. S1E in the supplemental material). The distribution of bacterial abundance resembled the vertical distribution of Chl a. Accordingly, bacterial abundance in July was generally low, at $\sim$0.9 $\times$ 10^6 cells ml^{-1}, except for a peak at $\sim$1.2 $\times$ 10^6 cells ml^{-1} in the Chl a maximum at 15 m depth. In September, bacterial abundance was highest in the surface, at $\sim$1.8 $\times$ 10^6 cells ml^{-1}, and steadily decreased with depth to $\sim$0.9 $\times$ 10^6 cells ml^{-1} at 100 m (Fig. 1H).

The Chl a concentrations measured here emphasize the mesotrophic nature of the Gullmar Fjord (22). The $\sim$2-fold difference between Chl a concentrations in July and September was likely due to elevated grazing pressure in July (personal observation), in line with previous observations (22). In terms of physicochemical water column characteristics and nutrient dynamics during the time of sampling, the Gullmar Fjord compares to other fjord systems (19, 20). Still, we recognize that our current study is limited to a single location in the Gullmar Fjord, which, although sampled at two ecologically relevant time points, does not cover the full range of (a)biotic gradients along the extension of the fjord or how they change during a full year in this or other fjord systems.

Overall patterns in prokaryotic community transcription by depth and month. Bacterial and archaeal transcription changed with depth and between samplings (Fig. 2). A principal-component analysis (PCA) separated the widely spread surface samples (down to 25 m) from a tight cluster of deeper samples (50 to 100 m) (Fig. 2A). Detailed hierarchical cluster analysis showed four distinct clusters consisting of a July surface cluster (5 to 15 m), a September surface cluster (12 to 25 m depth), and a separate cluster consisting of a mix of deep-water samples (50 to 100 m depth) (Fig. S3A). Further, redundancy analysis (RDA) showed that the influence of sampling date (July or September) on bacterial and archaeal transcription was strongest in the surface layer and decreased with depth (Fig. 2A and B). These differences were mainly explained by NH_4^+ (13% of variation) and partly by Chl a (9%), whereas depth variation was predominantly driven by DOC (12%) and $NO_3^- + NO_2^-$ (8%) (Fig. 2B; Fig. S2). This decrease in variability of transcription with depth was recently demonstrated for genes and taxa that showed diel oscillation in the surface but essentially diminished with depth through strong attenuation of light (24). Here, we show that this variability in transcription also applies to other gene systems, including genes coding for CAZymes, PEPs, and transporters (TPs).

Alphaproteobacteria and *Gammaproteobacteria* accounted for high portions (around 25 to 30%) of the prokaryotic community transcription in the surface waters and maintained elevated transcription ($\sim$11%) throughout the water column (Fig. 2C). *Cyanobacteria* (primarily *Synechococcus*) were most active in the upper water column (reaching $\sim$15% of community transcription). These patterns are largely in agreement with marine metagenomic and metatranscriptomic surveys from the open ocean, which typically report a dominance of *Picocyanobacteria* (*Prochlorococcus*), *Alphaproteobacteria* (primarily the SAR11 clade), and *Bacteroidetes* in the upper water column (6, 16, 25). The relatively high activity of *Verrucomicrobia* in July (reaching $\sim$45% of community transcription) is in line with a report from the Baltic Sea, where 16S rRNA gene analyses show that this taxon reaches higher relative abundances during summer coinciding with a dominance of *Cyanobacteria* (26). Metagenomic analysis of *Verrucomicrobia* shows that this taxon carries an extensive repertoire of CAZymes (e.g., alpha- and beta-galactosidases, xylanases, fucosidases, agarases, and endoglucanases), suggesting that they play a vital role in the
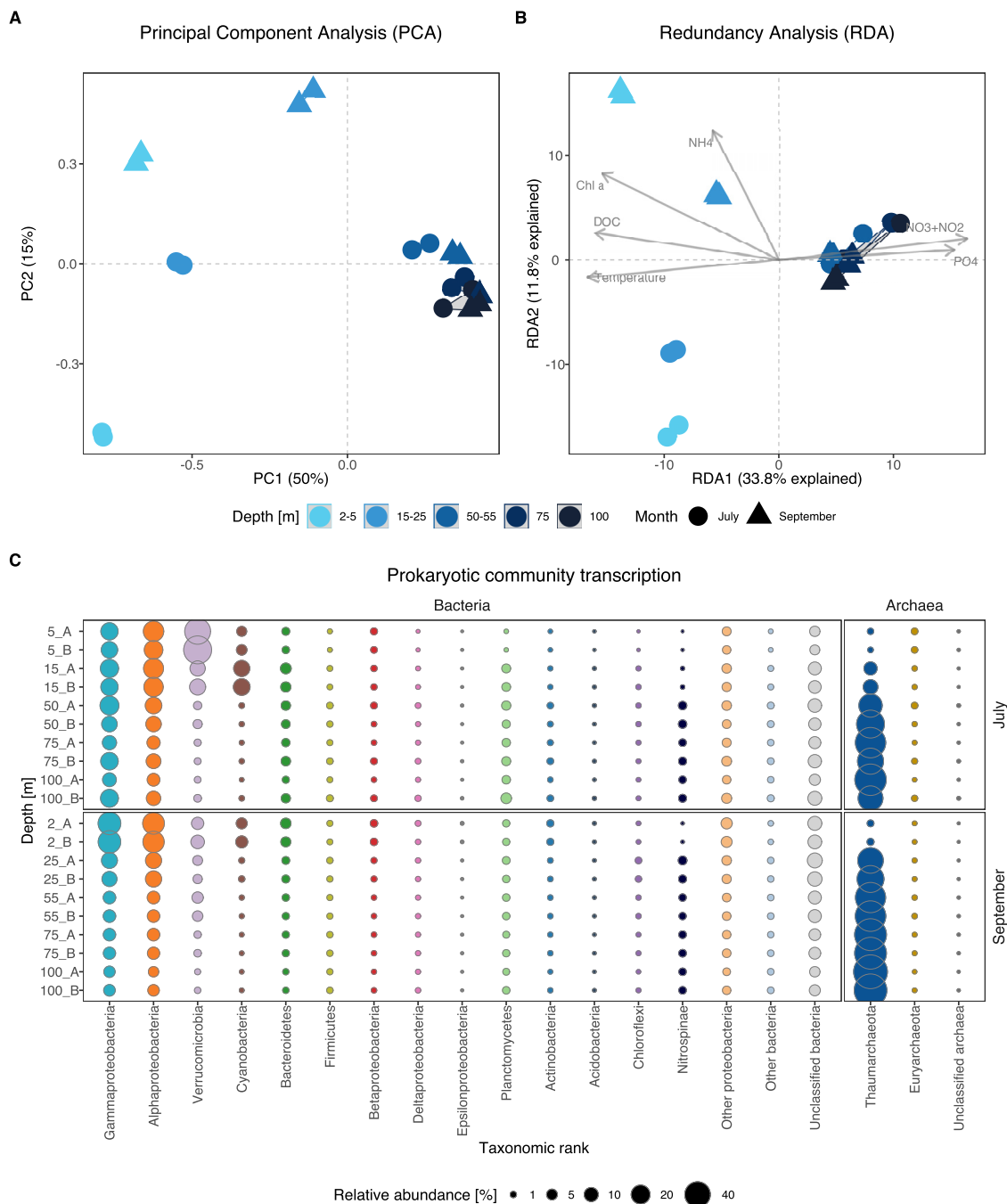
FIG 2 Variability in prokaryotic community transcription across depth and between sampling dates. (A) Principal-component analysis (PCA) based on 82,842 quality filtered open reading frames (ORFs). (B) Redundancy analysis (RDA) constrained by environmental variables. (C) Taxonomic affiliation of the transcriptionally most abundant prokaryotic phyla. Note that *Proteobacteria* are shown at the class level.

degradation of phytoplankton-derived DOM and complex polysaccharides such as fucoidans (16, 27, 28).

Strikingly, *Thaumarchaeota* dominated transcription from 50 m downwards in July and from ~25 m in September, accounting for up to ~75% of total transcripts. Also, the functional contribution of *Nitrospinae* (~2.5%, <0.5% at 5 m) was noticeable in deeper water layers that were associated with lower oxygen concentrations (Fig. 2C and 1E). High proportions of *Thaumarchaeota* and *Nitrospinae* at depth have also been found in the northern Pacific and the Mediterranean Sea (6, 16, 25). *Thaumarchaeota*

are key players in the oceanic N cycle by oxidizing ammonia to nitrite (first step in nitrification) (7, 9) and generally increase in abundance with depth, accounting for up to 40% of total cells in the mesopelagic zone (29). *Nitrospinae*, in turn, are the most abundant and ubiquitously distributed nitrite-oxidizing bacteria (NOB) in the ocean (performing the second step in nitrification, oxidation of nitrite to nitrate) (8). In natural systems, *Thaumarchaeota* are typically ~10-fold more abundant than *Nitrospinae* (30). Thus, the transcriptional activity of *Thaumarchaeota* together with *Nitrospinae* below 50 m depth suggests an important contribution of these two groups to the N cycle in this stratified fjord and potentially to the C cycle, since both taxa are lithoautotrophs.

Variation in bulk prokaryotic transcription of DOM transformation genes (CAZymes, PEPs, and TPs). Averaged over the entire water column, *Bacteria* accounted for 97% of CAZyme (~0.3% of total transcripts) and ~90% of PEP (~3.4% of total transcripts) gene transcription. In contrast, *Archaea* devoted little transcriptional effort to CAZymes and PEPs but instead accounted for ~55% of TP transcription (~10.4% of total transcripts) (see Table S2 at https://doi.org/10.6084/m9.figshare.17029547). These patterns prompted us to determine the relationship between the three gene systems over depth (see Fig. S4 in the supplemental material). The prokaryotic CAZyme transcription relative to that of PEPs was up to 5- to 10-fold higher in the surface waters than at depths from 50 m and below; note in particular the high proportion of CAZymes at 15 m in July (Fig. S4A). In addition, we found that the proportion of CAZyme and PEP transcription relative to TPs decreased with depth (Fig. S4B and S4C), mainly because of the increasing dominance of *Archaea*. The higher proportion of CAZymes in the Chl *a* maximum layer agrees with enzymatic activity dynamics during phytoplankton blooms, particularly during bloom senescence, and in Chl *a* maximum layers in marine and limnic systems (11, 31). If the concentration of cleavage end products determines prokaryotic ectoenzyme activities, the relative increase in PEP transcription with depth suggests that prokaryotic communities increase their efforts to acquire organic N and P because exported DOM typically has high C-to-N ratios (32). This has been shown for cell-specific enzymes such as leucine aminopeptidases and alkaline phosphatases, which typically increase in activity with depth (10).

These pronounced patterns in depth distributions, in turn, inspired further analysis of the allocation of transcriptional efforts to CAZymes, PEPs, and TPs among the dominant prokaryotic orders (Fig. 3), in an attempt to mitigate the limitation of relative metatranscriptomic data (i.e., the abundance of gene transcripts calculated as percentage of the sequence library) to disentangle the extent or direction of change in gene transcription of complex natural communities with depth. This highlighted the disproportionate contribution of *Thaumarchaeota* in nutrient uptake over polymer degradation at depth. *Alphaproteobacteria* such as *Pelagibacterales* showed a relatively high and stable proportion of TP transcripts compared to degrading enzymes with depth and across sampling dates. A somewhat different trend was noticed for *Rhodobacterales*, which potentially engaged more in carbohydrate degradation in the upper water column but changed toward proteins with depth; this was more pronounced in September than in July (Fig. 3). Interestingly, a preference for carbohydrates in comparison to proteins was more accentuated in *Flavobacteriales* in the upper water column but shifted toward PEPs with depth, whereas *Cellvibrionales* invested consistently more in CAZymes than PEPs throughout the water column in July but showed a trend similar to that of *Flavobacteriales* in September, emphasizing their significant role in the turnover of carbohydrates in surface waters. *Alteromonadales*, in turn, contributed more to the expression of TPs than CAZymes and seemingly favored peptides over carbohydrates in subsurface layers (Fig. 3). *Synechococcales* showed a disproportionate investment in TPs compared to degrading enzymes from 2 to 15 m depth. However, PEPs and TPs increased over CAZymes from 50 m downwards with a disproportionate transcription of PEPs over TPs.

These results suggest that the prokaryotic community attempted to acquire N and P with depth, as seen by the disproportional investment in PEPs and TPs over CAZymes. However, the relative proportions of these gene systems varied substantially between orders throughout the water column and were surprisingly consistent
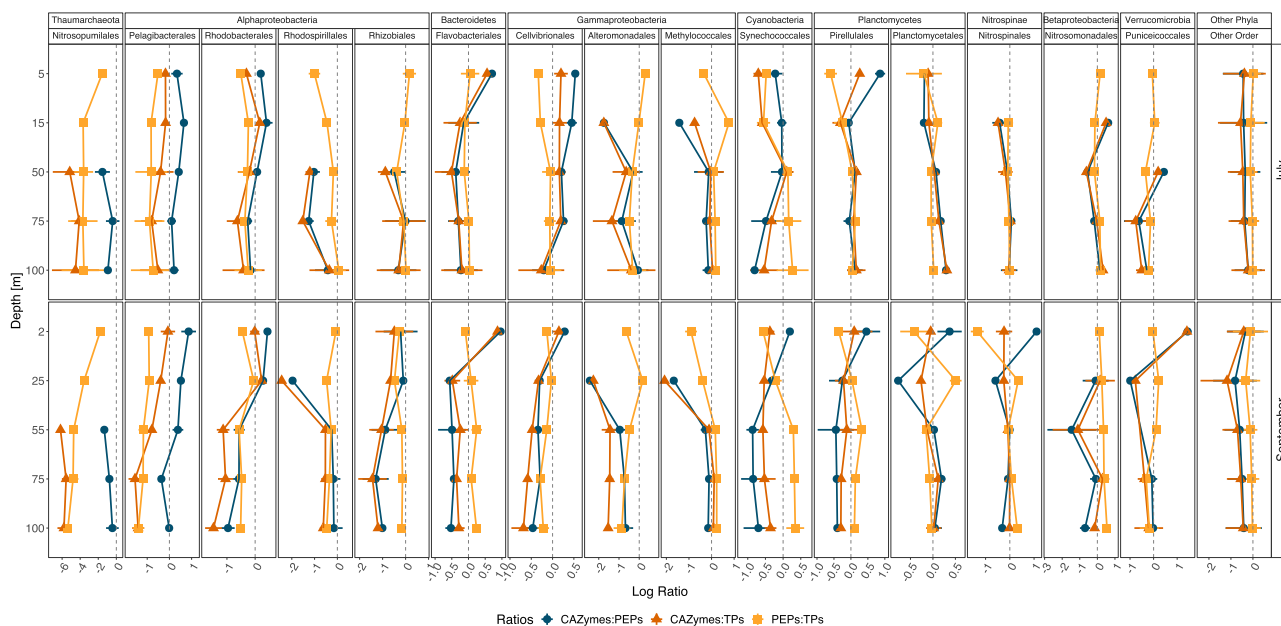
FIG 3 DOM cycling gene transcription ratios of the most active prokaryotic orders across spatiotemporal gradients. Depicted are the mean log ratios ± SD ($n = 2$) of total CAZyme to PEP transcription within an order for a given depth and date (blue points). In addition, the mean log ratios ± SD ($n = 2$) of CAZyme to TP transcription (orange triangles) and PEPs to TPs (yellow squares) are shown for each order. For instance, *Alteromonadales* depicted a negative CAZyme/PEP log ratio, indicating that the relative proportion of PEP transcription was higher than the CAZyme transcription, in particular at 15 m depth. Moreover, negative CAZyme/TP log ratios suggested that the transcription of TPs was more important than that of CAZymes. The slightly decreasing trend of PEP/TP, from initially positive at 5 m depth to negative at 100 m depth, indicated that peptide degradation was more important in the upper water layer than uptake that increased with depth. Overall, *Alteromonadales* transcribed more PEPs and TPs than CAZymes. Note the different signs and values on the y and x axes.

between sampling dates for some orders (e.g., *Thaumarchaeota*, *Pelagibacterales*, and *Alteromonadales*). In addition, our findings show that a few bacterial groups, such as *Bacteroidetes*, *Cellvibrionales*, *Planctomycetes*, and *Verrucomicrobia*, showed relatively higher transcription of CAZymes than of transporters, particularly in the surface, suggesting that marine microbes regulate their transcriptional efforts in degrading enzymes and transporters in a depth-dependent manner. The variation in the ratios implies a crucial role of different bacterial and archaeal taxa in the processing of HMW-DOM, emphasizing their metabolic/transcriptional plasticity in transcribing CAZymes, PEPs, and TPs. Possible triggers of such variation could be changing elemental ratios of DOM and the desire to maintain a balanced stoichiometry. If confirmed in future analysis, this would have implications for the constraints of DOM cycling and microbial population dynamics.

Divergence in CAZyme transcription. The richness of expressed CAZymes differed significantly between different depth layers (analysis of variance [ANOVA]; $F_{4,15} = 14.87$, $P < 0.00004$) (Fig. 4A), with ~4-fold lower values in the surface (2 to 5 m depth) than in the deeper water layers (15 to 100 m depth) (Tukey; $P < 0.03$ to 0.0001). While both richness (Fig. 4A) and evenness (Fig. 4B) were fairly similar in July, the evenness in September at 25 m was lower than that in the deep.

The microbial community expression of CAZyme genes differed significantly with both depth (permutational multivariate ANOVA [PERMANOVA], $R^2 = 0.61$, adjusted P value [P_{adj}] < 0.0003) and month (PERMANOVA, $R^2 = 0.1$, $P_{adj} < 0.001$) (Fig. 4C). Both PEP and CAZyme transcription in the surface layer (2 to 5 m) clustered with the samples from 15 to 25 m and distantly from the deeper samples (50 to 100 m) (Fig. S3B and C). Differences in clustering of samples between July and September were observed for the surface samples (2 to 5 and 15 to 25 m) but only minimally for the samples below 50 m depth (Fig. 4C; Fig. S3B and S3C).

The relative proportions of CAZyme classes were relatively stable throughout the water column (Fig. 4D), with the most abundant classes being glycoside hydrolases
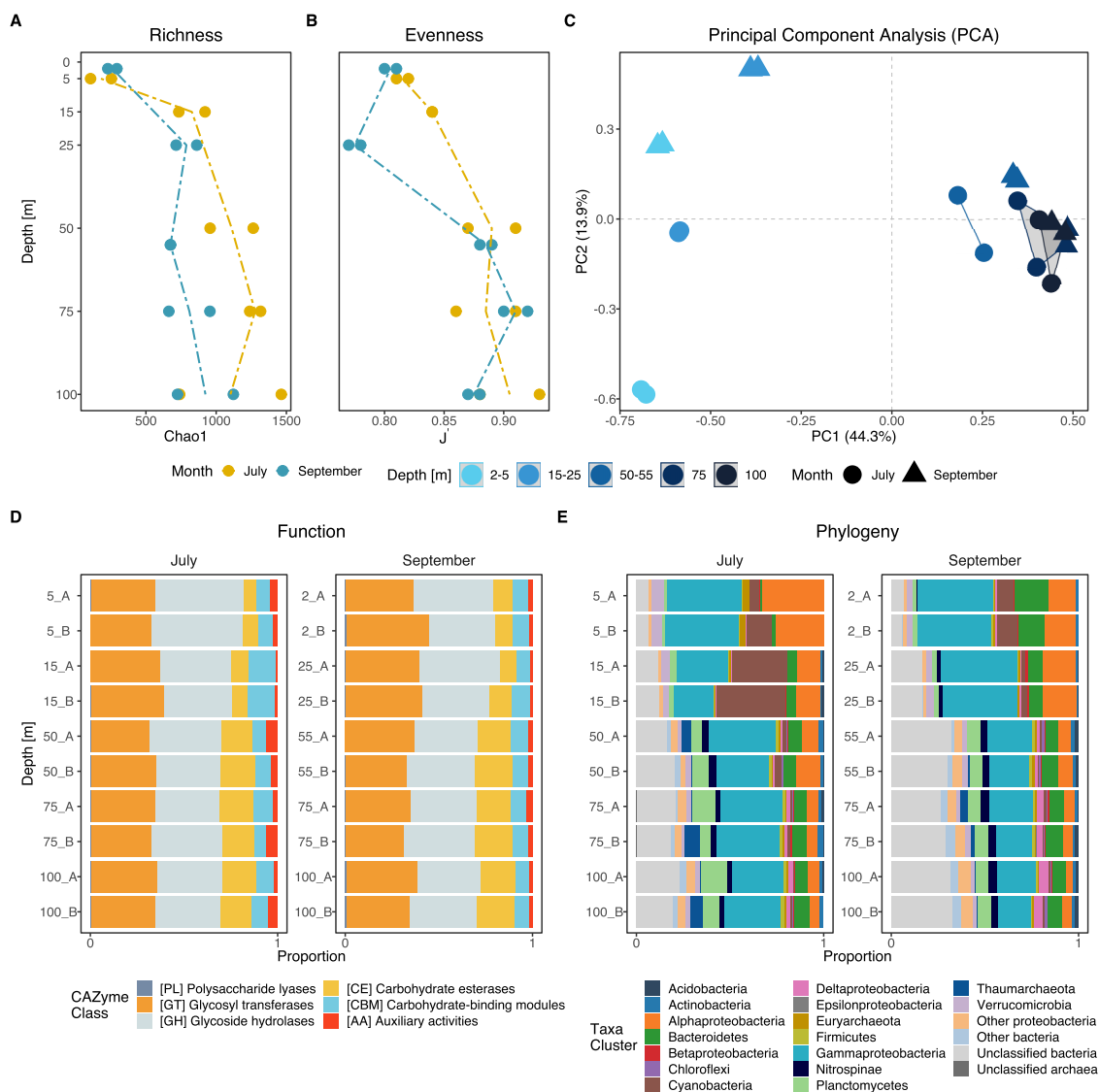
FIG 4 Taxonomic and functional profiles of carbohydrate-active enzyme (CAZyme) transcripts. (A and B) The estimated richness (Chao1) (A) and evenness (J′) (B) in depth profiles of CAZymes at the ORF level. (C) Principal-component analysis (PCA) of 1,350 CAZyme-affiliated open reading frames (ORFs). (D and E) Functional transcripts at the CAZyme class level (D) and taxonomic affiliation of the most active phyla transcribing CAZymes (E). *Proteobacteria* are grouped at the class level.

(GHs; 37.9% ± 4.4%, *n* = 20) and glycosyltransferases (GTs; 35.6% ± 3.4%). This was surprising given that the taxonomic affiliation of these genes changed with depth (Fig. 4E). Nevertheless, Zhao et al. found a similar pattern among epi- to bathypelagic samples from the Pacific, Atlantic, and Southern Oceans (18). This suggests that the CAZyme class level is too coarse to identify functional changes across depths, which is in stark contrast to the distribution of CAZyme families (see below) (Fig. 4 and 5).

Gammaproteobacteria dominated CAZyme expression in both July and September, accounting for up to ~40% of CAZymes in the upper water layers and 10 to 30% from 50 m and below (Fig. 4E). *Alphaproteobacteria* were also abundant in transcription in the surface layers (up to ~30%), with expression levels down to 5% of CAZymes at depth. Interestingly, *Cyanobacteria* dominated in July in the subsurface Chl *a* maximum layer, where they accounted for ~34% of total CAZyme transcription. The significantly lower evenness observed in September (Fig. 4B) was likely the result of the *Gammaproteobacteria* being composed of a diverse set of taxa with a more uneven distribution of transcripts (Fig. 4B). Notably, from 50 m downward, *Planctomycetes*,
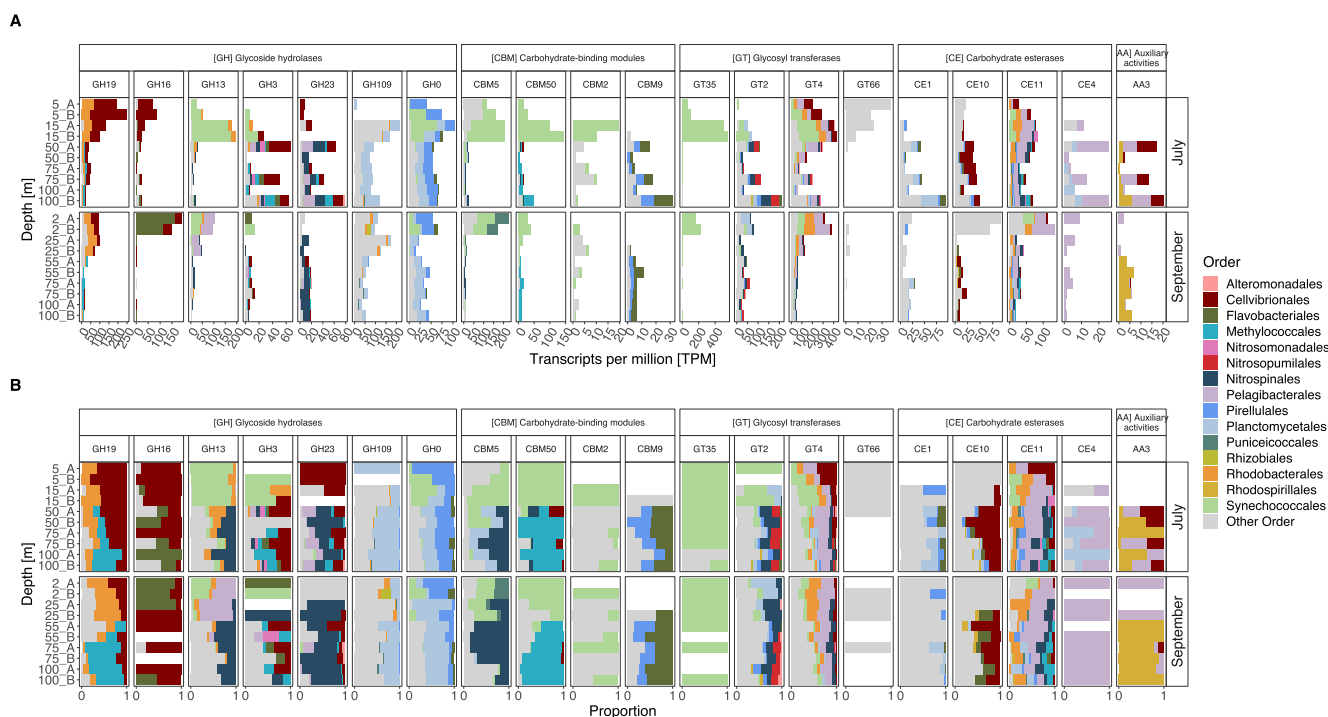
FIG 5 Transcription of the 20 most abundant carbohydrate-active enzymes (CAZymes) grouped into CAZyme families and taxonomic orders. (A) Relative transcript abundance in transcripts per million (TPM); note different scales on y axes. (B) Proportions of CAZyme transcripts. Individual CAZyme families are further grouped into CAZyme classes (top facets).

Nitrospinae, and *Deltaproteobacteria* accounted for ~8%, 3.6%, and ~2.6% (both months) of total CAZyme transcription, respectively (Fig. 4E). *Verrucomicrobia* showed a relatively stable activity profile with depth, contributing 1 to 6% of CAZyme expression.

Already at the order level, there were pronounced differences in the transcription of the most abundant CAZyme families between months and depth (Fig. 5; Fig. S5). For instance, *Cellvibrionales* (e.g., *Halieaceae* family) and *Rhodobacterales* (e.g., *Rhodobacteraceae* family) dominated transcription of glycoside hydrolase family 19 (GH19) in the surface layer, whereas *Methylococcales* (e.g., *Methylococcaceae* family) dominated below 50 m depth (Fig. 5; Fig. S6). Actually, GH19 explained 18.5% of the variation in community transcription with depth and ~27% by month (see Table S3 at https://doi.org/10.6084/m9.figshare.17029559). GH19 consists primarily of chitinases that hydrolyze chitin, the primary component of cell walls in fungi and exoskeletons of crustaceans (33). Incidentally, we noticed a high abundance of copepods in microscopy samples from July that coincided with the low Chl *a*, in line with previous observations from this time of year (22). These findings suggest a crucial role of chitin and chitodextrins as resources for prokaryotic plankton. *Cellvibrionales* (e.g., *Cellvibrionaceae*, *Halieaceae*, and *Porticoccaceae* families) and *Flavobacteriales* (e.g., *Flavobacteriaceae* family) both transcribed GH16, with a larger proportion of *Flavobacteriales* in September (Fig. 5; Fig. S6). The GH16 family comprises laminarinases that allow bacteria to decompose the algal storage glucan laminarin (a beta-1,3-glucan) (31). *Cellvibrionales* and *Flavobacteriales* are typically abundant and active during phytoplankton blooms and likely fulfill different roles in the degradation of phytoplankton-derived organic matter (31). The CAZyme transcription patterns support pronounced intra- and interspecific niche partitioning in the turnover of organic matter over spatiotemporal scales.

Synechococcales (e.g., the genus *Synechococcus*) dominated transcription of GH13 at 15 m in July, with some contribution also from *Alphaproteobacteria* (e.g., *Rhodobacterales* and *Pelagibacteraceae* family) (Fig. 5; Fig. S6). GH13 contains, for example, alpha-amylases

that allow hydrolysis of the storage polymers glycogen and starch (34). In addition, *Synechococcales* showed a high transcription of glycosyltransferases (GT35 and GT4) and carbohydrate-binding modules (CBM50 and CBM5). The glycosyltransferase GT35 is associated with the synthesis of starch and glycogen (35), whereas GT4 is involved in the synthesis of cellular structures and energy storage (i.e., sucrose, mannose, and trehalose synthase) upon photosynthesis (15, 35). The carbohydrate-binding modules CBM50 and CBM5 are involved in chitin binding to facilitate the degradation of chitin or peptidoglycan (36). Interestingly, macroalgae and cyanobacteria express an extensive suite of CAZymes (e.g., cellulases, amylases, galactosidases), PEPs, and lipases (37). Given that *Cyanobacteria*, including *Synechococcales*, synthesize the polysaccharide glycogen as an internal carbon and energy storage compound (38), the relatively high expression of GH13 is likely associated with the utilization of internal glycogen sources rather than being a sign of extracellular degradation. Indeed, cyanobacterial genomes encode a relatively low number of genes for secretory enzymes (18). The strong transcriptional response in CAZyme transcription by *Synechococcales* (especially *Synechococcus*) highlights the importance of these genes in regulating glycogen metabolism across pronounced light and nutrient gradients by balancing organic carbon synthesis and degradation.

Two taxa stood out as having higher CAZyme transcription from 50 m and down: *Planctomycetales* and *Nitrospinales* (Fig. 4). *Planctomycetales* (e.g., *Planctomyces*, *Gimesia*, and *Rubinisphaera* genera) dominated transcription of α-N-acetylgalactosaminidase (GH109) and glycoside hydrolases (G0, not yet classified in the CAZy database) (Fig. 5; Fig. S6). GH109 is involved in the degradation of bacterial cell walls (39) and has been shown to be abundant in metagenome-assembled genomes (MAGs) throughout the water column in the Mediterranean Sea (16) (Fig. 5). Since the cell walls of *Planctomycetales* lack the polymer peptidoglycan (40), the transcription of GH109 suggests a role in the degradation of cell walls of other bacteria. *Nitrospinales* increased transcription of GH23, CBM5, and GT2 with depth (Fig. 5). These CAZymes are involved in the degradation of peptidoglycan or chitin and the synthesis of chitin or cellulose (35, 39).

The transcriptional CAZyme responses by a diverse microbial community confirm the crucial role of the storage polysaccharide laminarin in fueling the heterotrophic carbon demand in this stratified fjord. Moreover, we noted strong transcriptional responses associated with the structural polysaccharides chitin and peptidoglycan. Remarkably, our results highlight a depth-layer-dependent functional partitioning in CAZyme transcription by well-known polymer degraders like *Cellvibrionales* and *Flavobacteriales*. Importantly, our analysis identified taxa like *Nitrospinales* (e.g., *Nitrospina* genus), *Nitrosomonadales* (e.g., *Nitrosopira* genus), and *Methylococcales* (e.g., *Methylococcaceae* family) that previously seem to have been overlooked in the context of polysaccharide degradation.

Divergence in membrane transporter transcription. The estimated richness of transcribed transporters increased with depth in July but was relatively constant in September (Fig. 6A). The evenness of transcribed transporters, however, was lower below 20 m depth, especially in September (Fig. 6B). Moreover, transcription of prokaryotic membrane transporter genes differed significantly with depth (PERMANOVA; $R^2 = 0.62$, $P_{adj} < 0.0003$) and sampling date (PERMANOVA; $R^2 = 0.08$, $P_{adj} < 0.002$) (Fig. 6C). As for overall transcription, cluster analysis of transporter transcription grouped samples into four distinct depth clusters (Fig. S3D). The largest variability in transcription was associated with the upper water layers, which clustered away from the samples from 50 m and below.
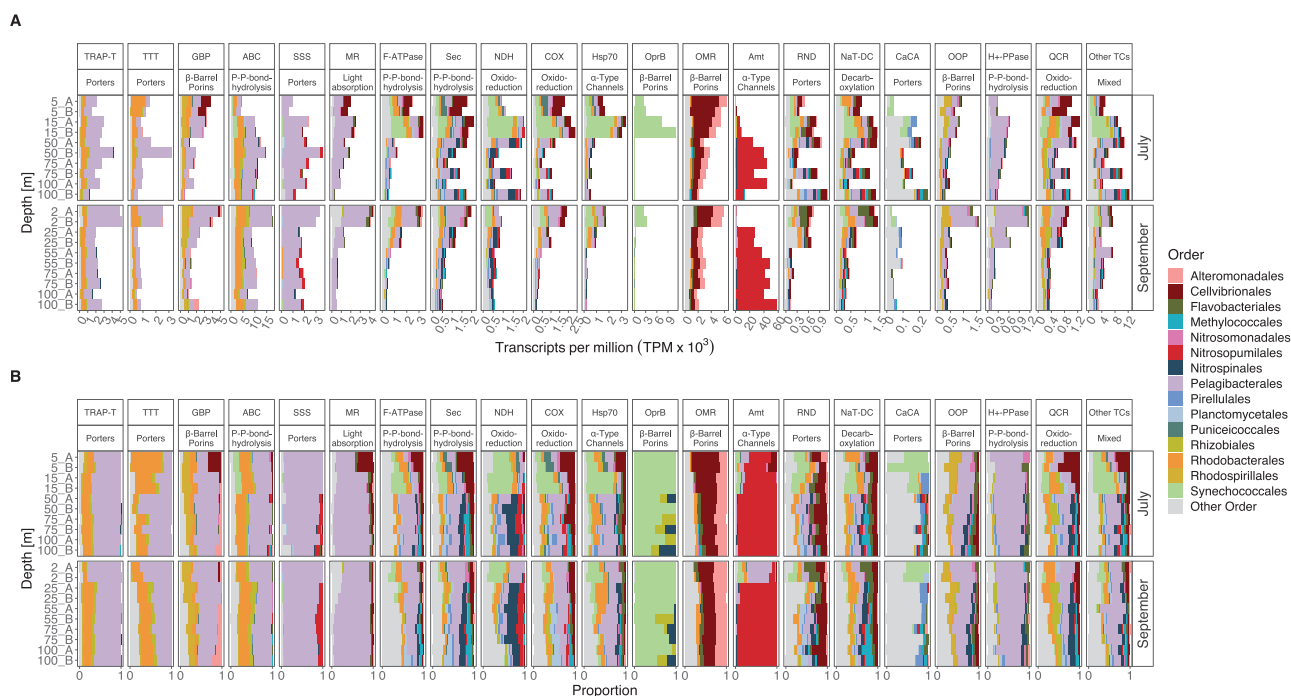
The most abundant transporter families were ammonium transporters (Amt; $47\% \pm 28\%$, number of samples [n] = 20), ATP-binding cassette transporters (ABC; $14\% \pm 6\%$, $n = 20$), and outer membrane receptors (OMR; $4\% \pm 3\%$, $n = 20$) (Fig. 6D). PCA of transporter transcription showed that these transporters contributed between 38% and 12% of variation in community gene expression with depth (PC1) and between 27% and 11% of variation by sampling date (PC2) (see Table S3 at https://doi .org/10.6084/m9.figshare.17029559). Generally, transporter transcription in the surface
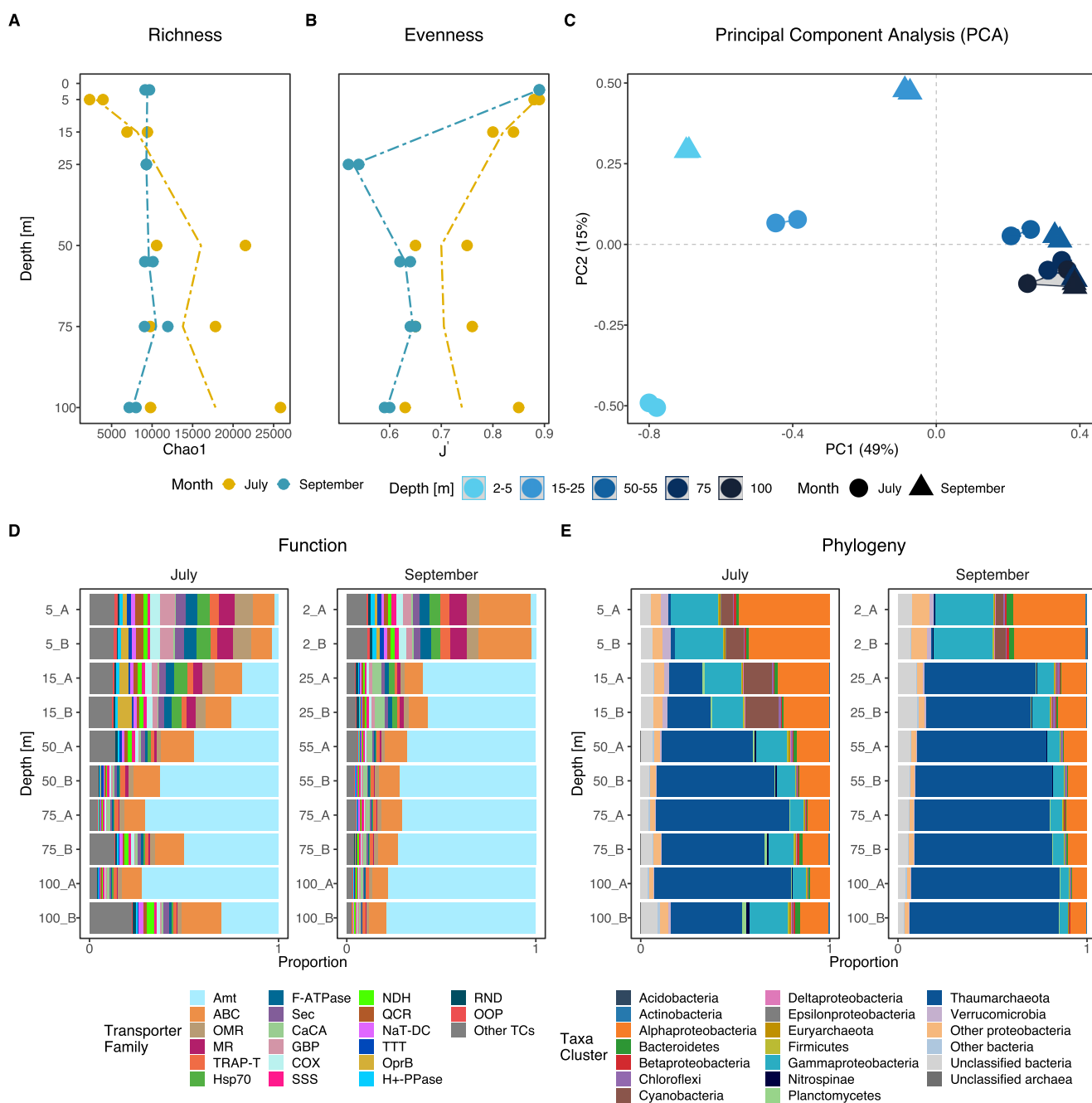
FIG 6 Taxonomic and functional transcription of membrane transporters (TPs). (A and B) The estimated richness (Chao1) (A) and evenness (J′) (B) in depth profiles of TPs at the ORF level. (C) PCA of 25,395 TP affiliated open reading frames (ORFs). (D and E) Functional transcripts of the most abundant TP families (D), and taxonomic affiliation of the most active phyla transcribing TPs (E). Proteobacteria are shown at the class level. Abbreviations of transporter families: Amt, ammonium transporter channel; ABC, ATP-binding cassette; OMR, outer membrane receptor; MR, ion-translocating microbial rhodopsin; TRAP-T, tripartite ATP-independent periplasmic transporter; Hsp70, cation channel-forming heat shock protein 70; F-ATPase, H^+- or Na^+-translocating F-type, V-type, and A-type ATPase; Sec, general secretory pathway; CaCA, Ca^{2+}:cation antiporter; GBP, general bacterial porin; COX, H^+-translocating cytochrome oxidase; SSS, solute-sodium symporter; NDH, H^+- or Na^+-translocating NADH dehydrogenase; QCR, H^+-translocating quinol: cytochrome c reductase; NaT-DC, Na^+-transporting carboxylic acid decarboxylase; TTT, tricarboxylate transporter; OprB, glucose-selective porin; H^+-PPase, H^+, Na^+-translocating pyrophosphatase; RND, resistance-nodulation-cell division; OOP, OmpA-OmpF porin.

layer was taxonomically similar to that of CAZymes. However, from 15 m depth and below, we noticed an exceptionally high contribution of *Thaumarchaeota* on both samplings, accounting for up to 79% of transporter transcription (Fig. 6E).

Our findings on membrane transporter transcription indicated that *Pelagibacterales* (e.g., "*Candidatus* genus Pelagibacter") and *Rhodobacterales* (e.g., *Rhodobacteraceae*

FIG 7 Transcription of the 20 most abundant transporter families (TPs) grouped into transporter subclasses and taxonomic orders. (A) Relative transcript abundance in transcripts per million × 10³; note different scales on y axes. (B) Proportions of TP transcription of prokaryotic orders. Abbreviations of transporter families are the same as in the legend to Fig. 6.

family) were actively engaged in the turnover of low-molecular-weight compounds but depicted different nutritional preferences with noticeable changes between depths and sampling dates (Fig. 7; Fig. S7 and S8). Both orders showed important contributions to, e.g., tripartite ATP-independent periplasmic transporter (TRAP-T), tricarboxylate transporter (TTT), and transporters for branched-chain amino acids (PF02653) and sugars (PF00532) (Fig. 7; Fig. S7). However, pronounced differences in transcription in a variety of transporter families, including solute-sodium symporter (SSS) and proteorhodopsin (microbial rhodopsin [MR], ~3% of transcription) were noted. As anticipated, *Rhodobacterales* did not transcribe proteorhodopsins but showed a higher activity in TTT transcription in July than did *Pelagibacterales*, which contributed more in the surface in September (Fig. 7). Also, *Pelagibacterales* (e.g., "*Candidatus* genus Pelagibacter") expressed high levels of glycine betaine transporters (PF04069), whereas *Rhodobacterales* (e.g., *Rhodobacteraceae* family) transcribed more extracellular solute-binding proteins for amino acids, oligopeptides, and oligosaccharides (PF00497, PF00496, PF01547). The common osmolyte glycine betaine is readily degraded by marine bacteria, including *Pelagibacterales*, who require reduced sulfur compounds (e.g., dimethylsulfoniopropionate [DMSP] or serine) for optimal growth (41, 42). Thus, they encode high-affinity glycine betaine transporters which have a broad specificity for substrates with similar structures to glycine betaine, such as DMSP, choline, and proline (43, 44). Similarly, for *Alphaproteobacteria* in the Atlantic Ocean, the lowest relative proportions of transporter-related membrane proteins derive from *Rhodobacterales* in the mesopelagic zone (300 to 850 m), whereas *Pelagibacterales* show the highest proportions in the euphotic zone down to the mesopelagic zone (45). Taken together, these differences in transcribed transporters together with differences in the depth distribution suggest pronounced functional partitioning of nutrients between *Pelagibacterales* and *Rhodobacterales* with depth.

Pelagibacterales (an unknown genus along with "*Candidatus* genus Pelagibacter") largely dominated the expression of proteorhodopsin (MR transporter family), along with some expression by *Cellvibrionales* (e.g., *Porticoccaceae* and *Halieaceae* families) and *Flavobacteriales* (e.g., *Flavobacteriaceae* family) (Fig. S8). As expected, proteorhodopsin

expression was highest in the surface (especially in September), yet substantial expression by *Pelagibacterales* (by an unclassified genus within the *Pelagibacteraceae*) was also detected down to 100 m (i.e., well below the deep chlorophyll maximum at 25 m) (Fig. 7). Proteorhodopsins (PR) are light-activated proton pumps that generate a proton motive force that in turn can be utilized by ATPases to generate energy (5). These retinal-binding rhodopsin proteins are widespread in bacteria (see references 46 and 47 and references therein), including *Pelagibacterales* (SAR11 clade) (see reference 42 and references therein). The precise metabolic processes that are stimulated by PR photoheterotrophy appear to differ between prokaryotes with different life strategies (46, 47). The physiological roles of PR below the euphotic zone, if any, are still unknown. Sunlight can penetrate several hundred meters deep into the water column, where even a few remaining photons may suffice for operating PR at maintenance levels (48). Yet, it is possible that the PR expression we observed in our deep samples is constitutive, given that experimental analyses show that PR expression in some strains of taxa as divergent as vibrios (*Gammaproteobacteria*) and pelagibacters (SAR11 clade; *Alphaproteobacteria*) is independent of light (49, 50).

Synechococcales showed the highest transcript abundance of transporters in the subsurface Chl *a* maximum layer in July and to a lesser extent in September. Photoautotrophs which are distributed between the surface and the deep Chl *a* maximum experience sharp gradients in light and nutrients and therefore face the challenge to optimize growth by adapting their metabolic machinery (51). The transcriptional increase of ion and electron transporters (e.g., F-ATPase, Sec, NDH, COX, and Hsp70) in the Chl *a* maximum, together with the active usage of storage polysaccharides, indicates an active struggle to maintain a beneficial metabolic balance for growth, given their involvement in respiration and stress responses (52). Interestingly, *Synechococcales* (e.g., *Synechococcus*) dominated transcription of the carbohydrate-selective porin (OprB) family in July (in the Chl *a* maximum) (Fig. 7; Fig. S8). To our knowledge, such high transcription of *oprB* by this taxon across a vertical depth profile has not been reported before. In *Pseudomonas putida*, these porins are part of a high-affinity uptake system for a variety of carbohydrates, including glucose (53). Unicellular picocyanobacteria (*Prochlorococcus* and *Synechococcus*) also encode proteins with an OprB domain (54–56), and uptake of glucosyl-glycerol, sucrose, and trehalose under temperature and salinity stress has been demonstrated (see reference 57 and references therein). Thus, picocyanobacteria may acquire a variety of low-molecular-weight carbohydrates through these porins either for osmotic adjustments (57) or in response to P starvation through uptake of sugar-phosphates (55). These results provide additional support for the importance of a certain capacity for mixotrophy in picocyanobacteria (58, 59).

An important transporter family dominated by *Cellvibrionales* (e.g., *Halieaceae* and *Porticoccaceae* families) and *Alteromonadales* (e.g., *Pseudoalteromonadaceae* and *Alteromonadaceae* families) (*Gammaproteobacteria*) was the OMR family, which reached the highest relative transcription in the surface layer and decreased with depth (Fig. 7; Fig. S8). Although the relative transcription of OMR by *Cellvibrionales* (e.g., *Halieaceae* family) was more than 3-fold higher than that of *Alteromonadales* (e.g., *Alteromonadaceae*) in the July surface layer, the two taxa contributed equally in September. Transporters in the OMR family, which include TonB receptors, have broad substrate specificities and are involved in the uptake of iron-siderophore complexes, vitamin B_{12}, nickel complexes, colicins, and carbohydrates (60). In the North Pacific Subtropical Gyre, *Alteromonadales* increased their transcriptional activity with depth (6) and showed the highest transporter protein abundance between 300 and 800 m depth in the Atlantic Ocean (6, 25, 45). Knowledge of the ecophysiology and ecology of *Cellvibrionales* is still scarce, but they appear to reach highest relative abundances in surface waters and coastal areas (61, 62). Here, we showed that both bacterial groups dominated OMR transcription throughout the water column, suggesting a vital role of the two gammaproteobacterial orders in nutrient cycling (potentially related to carbohydrates).

Strikingly, below 15 to 25 m depth, there was a dramatic increase in ammonium transporter (Amt) family transcripts by *Thaumarchaeota* (*Nitrosopumilus* genus), accounting for up to ~79% of transporter transcripts (~22% of total transcripts) at 100 m depth (Fig. 2 and 7; Fig. S8). Other transporter families transcribed by *Thaumarchaeota* were sodium-solute (SSS) and H^+ or Na^+-translocating carboxylic acid decarboxylase (NDH). Elevated transcription of the *amt* gene by *Thaumarchaeota* has been reported from the North Pacific Subtropical Gyre from 25 to 500 m, and a high abundance of Amt transporter proteins has been reported from the Atlantic Ocean (25, 45). *Thaumarchaeota* are abundant and widespread in marine waters and sediments (29, 63, 64) and account for major proportions of ammonia oxidation in mesopelagic waters of the open ocean through the key enzyme ammonia monooxygenase (encoded by *amoABC*) (9). *Thaumarchaeota* genomes encode high-affinity ammonium transporters that enable them to outcompete ammonia-oxidizing bacteria (AOB) under low ambient ammonium concentrations (nM range) (7, 65). The variety of ABC type transporters (e.g., for amino acids, oligopeptides, phosphonates), urea transporters, and solute-sodium symporter (SSS) family transporters (45, 66, 67) enables utilization of nitrogen sources other than ammonia (68) and supports previous findings that some *Thaumarchaeota* are capable of using inorganic and organic matter to supplement their metabolism (45, 66, 69).

Collectively, these findings suggest that depth strongly shapes the transcription of transporters, similarly to CAZymes. The most pronounced shift in transcription occurred between the surface layer and 50 m depth, largely consisting of a pronounced increase in thaumarchaeal ammonium transporters. These diverse responses in different membrane transporter families, which were associated with different taxa, highlight the important contribution of a few functional groups (e.g., *Pelagibacterales*, *Rhodobacterales*, and *Thaumarchaeota*) to the turnover of nutrients (e.g., amino acids, carbohydrates, glycine betaine, DMSP, and ammonium) across depth gradients.

Thaumarchaeal ammonia oxidation and carbon fixation. To follow up on the exceptionally high transcriptional activity of *Thaumarchaeota* below 25 m, we analyzed marker genes for thaumarchaeal energy and carbon metabolism (Fig. S9A and B). This revealed elevated transcription of both the gene for the large subunit of archaeal ammonia monooxygenase (*amoA*) mediating the first step of nitrification and the gene for archaeal inorganic carbon fixation (*hcd*; 4-hydroxybutyryl-CoA dehydratase) in the hydroxypropionate/hydroxybutyrate (HP/HB) cycle (7). Relating this expression to the expression of the single-copy gene *radA* (70) showed that the relative investment in ammonia oxidation was high (i.e., *amoA*/*radA* ratios, 3- to 4-fold) compared to that of inorganic carbon fixation (4-hydroxybutyryl-CoA dehydratase gene/*radA* ratios, ~3-fold), especially in the upper surface layer (Fig. S9C and D).

Although archaeal ammonia oxidation is coupled with inorganic carbon fixation (71), the low energy yield from ammonia oxidation ($\Delta G = -307.35$ kJ mol^{-1} NH_3) (72) suggests that a large quantity of ammonium needs to be oxidized per amount of CO_2 fixed (73). Accordingly, in soil *Thaumarchaeota*, the expression of *amoA* is severalfold higher than that of *hcd* (74), and in *Nitrosopumilus adriaticus*, nitrification resulted in a C yield per N of ~0.1 (75). Our gene expression data extend these findings on cultivated isolates to the natural marine environment and suggest that different *Thaumarchaeota* may have a distinct influence on the oceanic nitrogen compared to carbon cycles.

Conclusion. The overall depth distribution of prokaryotic transcription in the Gullmar Fjord agreed with findings in the Mediterranean Sea, North Pacific, and other fjords (16, 21, 25), keeping in mind that we sampled a single station in the fjord only twice. Nevertheless, our findings provide novel insights into the stratification of transcriptional activity of genes involved in transformation of DOM and nutrient uptake in natural microbial communities, which heretofore has been surveyed primarily through metagenomics (16–18). The pronounced transcriptional differences in CAZymes among distinct taxa of bacteria and archaea between samplings and over depth were notable in several respects. Given the succession often observed among phytoplankton and in relation to zooplankton grazers, we infer that ecosystem-level approaches

are urgently needed to disentangle the scale of carbon fluxes associated with bacterial utilization of phytoplankton storage polysaccharides compared to zooplankton structural polysaccharides. Moreover, building on previous studies of the genetic adaptations in nutrient scavenging with depth (see reference 45 and references therein), we suggest that expression analyses at the gene and protein levels are necessary to inform the mechanisms regulating the spatial partitioning of resources between bacterial and archaeal taxa. Characterizing and quantifying the constraints on nutrient and carbon cycling in particular depth layers, inhabited by specific sets of prokaryotic taxa, will be important for understanding how future changes in plankton dynamics influence the efficiency of the biological carbon pump.

MATERIALS AND METHODS

Study site and sampling. This study was conducted in the Gullmar Fjord at station Alsbäck (58°19'22.7N, 11°32'49.0E) (Fig. 1A and B). The fjord is located on the Swedish west coast approximately 100 km north of Gothenburg. As in other fjords, the Gullmar Fjord has pronounced differences in residence time between the dynamic surface (16 to 40 days) and deep waters (46 days to a year) (76), which ultimately structures the distribution of nutrients and microbiota. We sampled a vertical depth profile spanning from surface water to 100 m depth in July and September 2017. Although not measured, we estimated that the base of the euphotic zone was between 25 and 50 m deep, as the 1% light level is often below the deep chlorophyll *a* maximum (DCM) in open ocean systems. CTD profiles were obtained prior to the selection of the desired depth layers onboard the R/V *Oscar von Sydow*. In July, we sampled water from 5, 15, 50, 75, and 100 m depths, whereas in September we took water from 2, 25, 55, 75, and 100 m depths. The different depth layers were chosen based on CTD casts and represented major transitions in physicochemical parameters (see Fig. S1A to D in the supplemental material). From all depth layers, we sampled biological duplicates (i.e., from two separate 30-liter Niskin bottles) for inorganic nutrients, chlorophyll *a* (Chl *a*), dissolved organic carbon (DOC), and metatranscriptomics (detailed information is provided in supplemental material at https://doi.org/10.6084/m9.figshare.17032961).

Nutrients, chlorophyll *a* concentration, and prokaryotic abundance. *In situ* nutrient concentrations were measured with a QuAAtro AutoAnalyzer and XY-3 sampler (Seal Analytics) at the Sven Lovén Centre for Marine Infrastructure in Kristineberg, Sweden. Samples of 12 ml were poured into 13-ml polystyrene tubes (Sarstedt; 55.459) prior to the analysis of $NO_3^+ + NO_2^+$, NH_4^+, PO_4^{3-}, and SiO_2. For total nitrogen (TotN) and total phosphorus (TotP), 20-ml aliquots were analyzed with a QuAAtro AutoAnalyzer and XY-3 sampler (Seal Analytics) with the protocol Q126 R0 Nitrate in seawater MT3B and Q125 R0 phosphate in seawater MT3A. SiO_2 was measured with the protocol NIOZ–Kisel Z06605_3 and NH_4^+ with Q033 R7 ammonia MT3B.

Samples for determining DOC concentrations were filtered through 0.2-μm syringe filters (Acrodisc syringe filters; 32 mm, 514-4131, VWR), acidified with 448 μl of 1.2 M HCl to a pH of ~2 and analyzed with a high-temperature carbon analyzer (Shimadzu TOC-5000) at the intercalibrated facility at Umeå Marine Science Centre, Umeå, Sweden. DOC concentrations were calculated as nonpurgeable organic carbon. Chl *a* concentrations were determined fluorometrically upon ethanol extraction as described previously (77).

Samples for prokaryote abundance were fixed with 1% paraformaldehyde and 0.05% glutaraldehyde (final), stained with SYBR green I nucleic acid stain (Invitrogen) (5 μM final concentration), spiked with 2 μl Flow Check high-intensity green alignment beads (Polysciences, Inc.), and analyzed with a CyFlow Cube 8 flow cytometer (Sysmex Partec) according to reference 78 (see supplemental material at https://doi.org/10.6084/m9.figshare.17032961).

Metatranscriptomics analysis. Water samples for metatranscriptomics were retrieved from every depth layer in biological duplicates. Approximately 3.5 liters of water was filtered through Sterivex filter units (GP 0.22 μm; EMD Millipore) onboard within ~30 min after sampling, preserved in 2 ml RNAlater (Qiagen), immediately flash frozen in liquid nitrogen until samples were transferred to a −80°C freezer at the Sven Loven Center (Kristineberg, Sweden), and later stored at −80°C at Linnaeus University (Kalmar, Sweden) until further processing. RNAlater was removed from the Sterivex filter cartridges by using a sterile 20-ml syringe and applying gentle pressure until all of the preservation solution was removed. Total RNA was extracted with an RNeasy minikit (Qiagen) according to reference 79 with some modifications as described in reference 80. In brief, total RNA was DNase treated using a TURBO DNA-free kit (ThermoFisher Scientific) and afterwards controlled for residual DNA by a 30-cycle PCR with 16S rRNA gene primers (27F and 1492R) including Milli-Q as a negative control and DNA from *Escherichia coli* as a positive control. rRNA was depleted using a RiboMinus transcriptome isolation kit and a RiboMinus concentration module (ThermoFisher Scientific). The rRNA-depleted fraction was reverse transcribed into cDNA (first and second strand) and linearly amplified (i.e., *in vitro* transcribed) using the MessageAmp II-bacterial RNA amplification kit (ThermoFisher Scientific) according to the manufacturer's instructions before library construction (TruSeq) and sequencing at the National Genome Infrastructure SciLifeLab Stockholm (Illumina HiSeq 2500 platform in rapid mode and with HiSeq SBS kit v4 chemistry to obtain 2 × 126 bp long paired-end reads) (see supplemental material at https://doi.org/10.6084/m9.figshare.17032961).

Illumina adapter sequences were removed from quality-controlled paired-end reads with Cutadapt (81) (v1.13), and reads were quality trimmed with Sickle (https://github.com/najoshi/sickle) (v1.33) in

paired-end mode and with Sanger quality values. Our analysis strategy was dictated by the need to identify longer mRNA fragments to improve peptide predictions, which enables the use of hmm profiles (CAZymes, peptidases, and transporters). Therefore, we decided to perform a *de novo* assembly of high-quality reads with MEGAHIT (82) (v1.1.2) and default parameters, resulting in 451,080 contigs with a minimum length of 200 bp, a maximum length of 51,124 bp, and an average contig length of 537 bp with an N_{50} of 538 bp. Subsequently, open reading frames (ORFs) were determined with Prodigal (83) (v2.6.3) and default parameters. rRNAs from bacteria, archaea, and eukaryotes were predicted by running BARNAP (84) (v0.9) on the contigs with an E-value cutoff of 1e−6 and rejection of genes of <0.05 of the expected length. Subsequently, ORFs that showed an overlap with the predicted rRNAs on the contigs were removed from the downstream analysis (6,917 ORFs) (Table S1). The remaining peptides were used for blastp searches in the NCBI RefSeq protein database with DIAMOND (85) (v0.9.24) and an E-value threshold of 0.001. Thereafter, taxonomic annotation was done with MEGAN (86) (v6.12.8), using the lowest common ancestor (LCA) algorithm for asemblies with a suggested bit score of 50 and a top hit of 10%. Reads were mapped with bowtie2 (87) (v2.3.5.1) in paired-end mode to the ORFs. Subsequently, Samtools (88) (v1.9) was used to quantify transcript counts per gene. Sequencing and bioinformatics summary statistics are in Table S1.

CAZyme open reading frames (ORFs) were detected and classified with the run-dbcan program (89) (v2.0.11) and default parameters. Consensus CAZyme family classifications were filtered as follows: (i) ORFs, predicted with only one tool, were excluded if their percent identity was <50% (DIAMOND), coverage was <0.5 (HMMER), or hits were <10 (Hotpep); (ii) ORFs with positive hits from more than one prediction tool were ranked based on the number of CAZyme family classifications and their relative percent identity (DIAMOND), percent coverage (HMMER), or relative hits plus frequency (Hotpep) across all prediction tools. The highest-ranking CAZyme family classification was kept for downstream analyses. PEPs were identified using HMMER3 (90) and Pfam profiles matching hits to PEP subunit sequences in the MEROPS database (https://www.ebi.ac.uk/merops). Transporter genes were detected and classified with HMMER3 using profiles from the Transporter Classification databases (TCDB). We noticed that two PFAM domains, PF00909 and PF00654, were assigned to incorrect TCDB families, and therefore we manually reassigned the protein domain PF00909 to the "The Ammonium Channel Transporter (Amt) Family; TC 1.A.11" and protein domain PF00654 to the "The Chloride Carrier/Channel (ClC) Family; 2.A.49." This incorrect assignment is noted in the TCDB database (see supplemental material at https://doi.org/10.6084/m9.figshare.17032961).

The Archaea-specific marker genes for RadA and AmoA peptides were detected by hidden Markov models run with HMMER3 and the profiles TIGR02236 (RadA) and PF12942 (AmoA), in the TIGRFAM (v15.0) and Protein Families (Pfam) (v31.0) databases. Hits were considered valid if the score was equal to, or higher than, the recommended "gathering score" for the model. In the case of 4-hydroxybutyryl-coenzyme A (CoA) dehydratase, an alignment was constructed using MUSCLE (91) with the peptide from the characterized gene in *Nitrosopumilus maritimus* SCM1 (locustag Nmar_0207), its predicted orthologs in other *Thaumarchaeota* (7), and all peptides that contained the two conserved domains according to their PFAMs in Nmar_0207, PF11794 (4-hydroxyphenylacetate 3-hydroxylase N terminal) and PF03241 (4-hydroxyphenylacetate 3-hydroxylase C terminal) in the set of peptides derived from the genomes in the MAR database using the "gathering score" as the cutoff. A maximum likelihood phylogeny of the peptides was estimated using FastTree (92). Nmar_0207 clustered together away from predicted paralogs (peptides with a different function). To quantify Nmar_0207 ortholog expression, the peptides from the assembled metatranscriptome sequences were searched using blastp with a relaxed E value (0.0001) against the database of all MAR peptides together with Nmar_0207 predicted orthologs. Those whose closest hits were to Nmar_0207 orthologs were chosen to align them and place the new peptides on the phylogenetic tree. The packages PaPaRa (93), EPA-ng (94), and gappa (95) were used to confirm the position of the new peptides on the phylogenetic tree inside the group of 4-hydroxybutyryl-CoA dehydratase orthologs after visualization of the tree with iToL (96). The amino acid substitution model was predicted with IQ-TREE (97) (see supplemental material at https://doi.org/10.6084/m9.figshare.17032961).

Statistics and visualization. In total, 523,441 ORFs were predicted from the assembly, of which 6,917 were identified as rRNA and subsequently removed from further analysis. Subsetting the data set for bacteria and archaea resulted in 256,673 ORFs, which were used for downstream analyses. Principal-component analysis (PCA) was performed on Hellinger transformed raw counts according to reference 98 on ORFs with at least 5 cpm in 2 or more samples (resulting in 82,842 ORFs). Redundancy analysis (RDA) was performed on the same input data as described above. Environmental variables were standardized with the function "decostand" in vegan and selected based on pairwise Pearson correlation coefficients of <0.9 and variance inflation factors of <10. NH_4^+ concentrations below the detection limit of 0.2 μM were replaced with a small value (0.001) to enable the estimation of this variable in the tb-RDA. The suitability of an RDA was tested prior to analysis (gradient length of ~3.5). The model consisting of the variables temperature, DOC, Chl *a*, NH_4^+, $NO_3^+ + NO_2^+$, and PO_4^+ was significant ($P < 0.001$, $R^2_{adj} = 57\%$). Monte Carlo permutation tests showed that (i) both RDA axes were significant after Holm's correction for multiple testing ($P_{adj} < 0.006$) and (ii) the variables DOC ($P_{adj} < 0.006$), Chl *a* ($P_{adj} < 0.043$), NH_4^+ ($P_{adj} < 0.04$), and $NO_3^+ + NO_2^+$ ($P_{adj} < 0.043$) were significant, explaining ~12% and ~9%, ~13%, and 8% of the variation in prokaryotic community transcripts, respectively, as derived from the variance partitioning (Fig. S2). Grouping of samples was determined through hierarchical clustering based on scaled Hellinger transformed raw counts, Euclidean distances, and Ward D2 cluster criteria. The optimal number of clusters was graphically determined through Elbow and Silhouette methods. Observed richness (Chao1) and evenness (Pielou's J) were calculated with normalized counts that were scaled by ranked subsampling (SRS) according to reference 99. All statistical analyses and graphical visualizations were conducted in RStudio

(100) and predominantly with functions from the packages tidyverse (101) and vegan (102) (see supplemental material at https://doi.org/10.6084/m9.figshare.17032961).

Data availability. All metatranscriptome data are available at the EMBL-EBI European Nucleotide Archive repository under project accession number PRJEB42919.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, EPS file, 2.5 MB.
FIG S2, EPS file, 0.2 MB.
FIG S3, EPS file, 0.4 MB.
FIG S4, EPS file, 2.1 MB.
FIG S5, EPS file, 0.8 MB.
FIG S6, EPS file, 2.1 MB.
FIG S7, EPS file, 2 MB.
FIG S8, EPS file, 1.9 MB.
FIG S9, EPS file, 0.6 MB.
TABLE S1, PDF file, 0.2 MB.

REFERENCES

1. Ducklow H, Steinberg D, Buesseler K. 2001. Upper ocean carbon export and the biological pump. Oceanography 14:50–58. https://doi.org/10.5670/oceanog.2001.06.
2. Azam F. 1998. Microbial control of oceanic carbon flux: the plot thickens. Science 280:694–696. https://doi.org/10.1126/science.280.5364.694.
3. Delong EF, Franks DG, Alldredge AL. 1993. Phylogenetic diversity of aggregate-attached vs free-living marine bacterial assemblages. Limnol Oceanogr 38:924–934. https://doi.org/10.4319/lo.1993.38.5.0924.
4. Agogué H, Lamy D, Neal PR, Sogin ML, Herndl GJ. 2011. Water mass-specificity of bacterial communities in the North Atlantic revealed by massively parallel sequencing. Mol Ecol 20:258–274. https://doi.org/10.1111/j.1365-294X.2010.04932.x.
5. Béjà O, Aravind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP, Jovanovich SB, Gates CM, Feldman RA, Spudich JL, Spudich EN, DeLong EF. 2000. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. Science 289:1902–1906. https://doi.org/10.1126/science.289.5486.1902.
6. DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, Frigaard NU, Martinez A, Sullivan MB, Edwards R, Brito BR, Chisholm SW, Karl DM. 2006. Community genomics among stratified microbial assemblages in the ocean's interior. Science 311:496–503. https://doi.org/10.1126/science.1120250.
7. Könneke M, Schubert DM, Brown PC, Hügler M, Standfest S, Schwander T, Schada von Borzyskowski L, Erb TJ, Stahl DA, Berg IA. 2014. Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO_2

fixation. Proc Natl Acad Sci U S A 111:8239–8244. https://doi.org/10
.1073/pnas.1402028111.

8. Pachiadaki MG, Sintes E, Bergauer K, Brown JM, Record NR, Swan BK, Mathyer ME, Hallam SJ, Lopez-Garcia P, Takaki Y, Nunoura T, Woyke T, Herndl GJ, Stepanauskas R. 2017. Major role of nitrite-oxidizing bacteria in dark ocean carbon fixation. Science 358:1046–1051. https://doi.org/10 .1126/science.aan8260.

9. Wuchter C, Abbas B, Coolen MJ, Herfort L, van Bleijswijk J, Timmers P, Strous M, Teira E, Herndl GJ, Middelburg JJ, Schouten S, Sinninghe Damsté JS. 2006. Archaeal nitrification in the ocean. Proc Natl Acad Sci U S A 103:12317–12322. https://doi.org/10.1073/pnas.0600756103.

10. Baltar F, Arístegui J, Sintes E, van Aken HM, Gasol JM, Herndl GJ. 2009. Prokaryotic extracellular enzymatic activity in relation to biomass production and respiration in the meso- and bathypelagic waters of the (sub)tropical Atlantic. Environ Microbiol 11:1998–2014. https://doi.org/10 .1111/j.1462-2920.2009.01922.x.

11. Hoppe HG, Ullrich S. 1999. Profiles of ectoenzymes in the Indian Ocean: phenomena of phosphatase activity in the mesopelagic zone. Aquat Microb Ecol 19:139–148. https://doi.org/10.3354/ame019139.

12. Arnosti C. 2011. Microbial extracellular enzymes and the marine carbon cycle. Annu Rev Mar Sci 3:401–425. https://doi.org/10.1146/annurev -marine-120709-142731.

13. Stanley KK, Newby AC, Luzio JP. 1982. What do ectoenzymes do? Trends Biochem Sci 7:145–147. https://doi.org/10.1016/0968-0004(82)90207-9.

14. Weiss MS, Abele U, Weckesser J, Welte W, Schiltz E, Schulz GE. 1991. Molecular architecture and electrostatic properties of a bacterial porin. Science 254:1627–1630. https://doi.org/10.1126/science.1721242.

15. Doane M, Haggerty JM, da Silva Lopes CR, Yates P, Edwards R, Dinsdale E, Lopes FAC, Bruce T. 2020. Latitude and chlorophyll a density drive the distribution of carbohydrate-active enzymes in the planktonic microbial fraction of the epipelagic zone. Environ Microbiol Rep 12:473–485. https://doi.org/10.1111/1758-2229.12865.

16. Haro-Moreno JM, López-Pérez M, de la Torre JR, Picazo A, Camacho A, Rodriguez-Valera F. 2018. Fine metagenomic profile of the Mediterranean stratified and mixed water columns revealed by assembly and recruitment. Microbiome 6:128. https://doi.org/10.1186/s40168-018-0513-5.

17. Pelve EA, Fontanez KM, DeLong EF. 2017. Bacterial succession on sinking particles in the ocean's interior. Front Microbiol 8:2269. https://doi.org/ 10.3389/fmicb.2017.02269.

18. Zhao Z, Baltar F, Herndl GJ. 2020. Linking extracellular enzymes to phylogeny indicates a predominantly particle-associated lifestyle of deep-sea prokaryotes. Sci Adv 6:eaaz4354. https://doi.org/10.1126/sciadv.aaz4354.

19. Gašparović B, Plavšić M, Ćosović B, Reigstad M. 2005. Organic matter characterization and fate in the sub-arctic Norwegian fjords during the late spring/summer period. Estuar Coast Shelf Sci 62:95–107. https://doi .org/10.1016/j.ecss.2004.08.008.

20. Storesund JE, Sandaa RA, Thingstad TF, Asplin L, Albretsen J, Erga SR. 2017. Linking bacterial community structure to advection and environmental impact along a coast-fjord gradient of the Sognefjord, western Norway. Prog Oceanogr 159:13–30. https://doi.org/10.1016/j.pocean .2017.09.002.

21. Tobias-Hünefeldt SP, Wing SR, Espinel-Velasco N, Baltar F, Morales SE. 2019. Depth and location influence prokaryotic and eukaryotic microbial community structure in New Zealand fjords. Sci Total Environ 693: 133507. https://doi.org/10.1016/j.scitotenv.2019.07.313.

22. Tiselius P, Belgrano A, Andersson L, Lindahl O. 2016. Primary productivity in a coastal ecosystem: a trophic perspective on a long-term time series. J Plankton Res 38:1092–1102. https://doi.org/10.1093/plankt/fbv094.

23. Lindahl O, Belgrano A, Davidsson L, Hernroth B. 1998. Primary production, climatic oscillations, and physico-chemical processes: the Gullmar Fjord time-series data set (1985–1996). ICES J Mar Sci 55:723–729. https://doi.org/10.1006/jmsc.1998.0379.

24. Vislova A, Sosa OA, Eppley JM, Romano AE, DeLong EF. 2019. Diel oscillation of microbial gene transcripts declines with depth in oligotrophic ocean waters. Front Microbiol 10:2191. https://doi.org/10.3389/fmicb .2019.02191.

25. Shi Y, Tyson GW, Eppley JM, DeLong EF. 2011. Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. ISME J 5:999–1013. https://doi.org/10.1038/ ismej.2010.189.

26. Lindh MV, Sjöstedt J, Andersson AF, Baltar F, Hugerth LW, Lundin D, Muthusamy S, Legrand C, Pinhassi J. 2015. Disentangling seasonal bacterioplankton population dynamics by high-frequency sampling. Environ Microbiol 17:2459–2476. https://doi.org/10.1111/1462-2920.12720.

27. Herlemann DP, Lundin D, Labrenz M, Jürgens K, Zheng Z, Aspeborg H, Andersson AF. 2013. Metagenomic de novo assembly of an aquatic representative of the verrucomicrobial class Spartobacteria. mBio 4:e00569-12. https://doi.org/10.1128/mBio.00569-12.

28. Sichert A, Corzett CH, Schechter MS, Unfried F, Markert S, Becher D, Fernandez-Guerra A, Liebeke M, Schweder T, Polz MF, Hehemann JH. 2020. Verrucomicrobia use hundreds of enzymes to digest the algal polysaccharide fucoidan. Nat Microbiol 5:1026–1039. https://doi.org/10 .1038/s41564-020-0720-2.

29. Karner MB, DeLong EF, Karl DM. 2001. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. Nature 409:507–510. https://doi.org/ 10.1038/35054051.

30. Kitzinger K, Marchant HK, Bristow LA, Herbold CW, Padilla CC, Kidane AT, Littmann S, Daims H, Pjevac P, Stewart FJ, Wagner M, Kuypers MMM. 2020. Single cell analyses reveal contrasting life strategies of the two main nitrifiers in the ocean. Nat Commun 11:767. https://doi.org/10 .1038/s41467-020-14542-3.

31. Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM, Kassabgy M, Huang S, Mann AJ, Waldmann J, Weber M, Klindworth A, Otto A, Lange J, Bernhardt J, Reinsch C, Hecker M, Peplies J, Bockelmann FD, Callies U, Gerdts G, Wichels A, Wiltshire KH, Glöckner FO, Schweder T, Amann R. 2012. Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. Science 336: 608–611. https://doi.org/10.1126/science.1218344.

32. Schneider B, Schlitzer R, Fischer G, Nöthig E-M. 2003. Depth-dependent elemental compositions of particulate organic matter (POM) in the ocean. Global Biogeochem Cycles 17:1032.

33. Muzzarelli RAA, Boudrant J, Meyer D, Manno N, DeMarchis M, Paoletti MG. 2012. Current views on fungal chitin/chitosan, human chitinases, food preservation, glucans, pectins and inulin: a tribute to Henri Braconnot, precursor of the carbohydrate polymers science, on the chitin bicentennial. Carbohydr Polym 87:995–1012. https://doi.org/10.1016/j.carbpol .2011.09.063.

34. Grondin JM, Tamura K, Déjean G, Abbott DW, Brumer H. 2017. Polysaccharide utilization loci: fueling microbial communities. J Bacteriol 199: e00860-16. https://doi.org/10.1128/JB.00860-16.

35. Henrissat B, Coutinho PM, Davies GJ. 2001. A census of carbohydrate-active enzymes in the genome of Arabidopsis thaliana. Plant Mol Biol 47: 55–72. https://doi.org/10.1023/A:1010667012056.

36. Bennke CM, Krüger K, Kappelmann L, Huang S, Gobet A, Schüler M, Barbe V, Fuchs BM, Michel G, Teeling H, Amann RI. 2016. Polysaccharide utilisation loci of Bacteroidetes from two contrasting open ocean sites in the North Atlantic. Environ Microbiol 18:4456–4470. https://doi.org/10 .1111/1462-2920.13429.

37. Brasil B, de Siqueira FG, Salum TFC, Zanette CM, Spier MR. 2017. Microalgae and cyanobacteria as enzyme biofactories. Algal Res 25:76–89. https://doi.org/10.1016/j.algal.2017.04.035.

38. Cano M, Holland SC, Artier J, Burnap RL, Ghirardi M, Morgan JA, Yu J. 2018. Glycogen synthesis and metabolite overflow contribute to energy balancing in cyanobacteria. Cell Rep 23:667–672. https://doi.org/10 .1016/j.celrep.2018.03.083.

39. Kamke J, Sczyrba A, Ivanova N, Schwientek P, Rinke C, Mavromatis K, Woyke T, Hentschel U. 2013. Single-cell genomics reveals complex carbohydrate degradation patterns in poribacterial symbionts of marine sponges. ISME J 7:2287–2300. https://doi.org/10.1038/ismej.2013.111.

40. Lindsay MR, Webb RI, Strous M, Jetten MS, Butler MK, Forde RJ, Fuerst JA. 2001. Cell compartmentalisation in planctomycetes: novel types of structural organisation for the bacterial cell. Arch Microbiol 175:413–429. https://doi.org/10.1007/s002030100280.

41. Carini P, Steindler L, Beszteri S, Giovannoni SJ. 2013. Nutrient requirements for growth of the extreme oligotroph 'Candidatus Pelagibacter ubique' HTCC1062 on a defined medium. ISME J 7:592–602. https://doi .org/10.1038/ismej.2012.122.

42. Giovannoni SJ. 2017. SAR11 bacteria: the most abundant plankton in the oceans. Annu Rev Mar Sci 9:231–255. https://doi.org/10.1146/annurev -marine-010814-015934.

43. Kiene RP, Hoffmann Williams LP, Walker JE. 1998. Seawater microorganisms have a high affinity glycine betaine uptake system which also recognizes dimethylsulfoniopropionate. Aquat Microb Ecol 15:39–51. https://doi .org/10.3354/ame015039.

44. Noell SE, Giovannoni SJ. 2019. SAR11 bacteria have a high affinity and multifunctional glycine betaine transporter. Environ Microbiol 21: 2559–2575. https://doi.org/10.1111/1462-2920.14649.

45. Bergauer K, Fernandez-Guerra A, Garcia JAL, Sprenger RR, Stepanauskas R, Pachiadaki MG, Jensen ON, Herndl GJ. 2018. Organic matter processing by microbial communities throughout the Atlantic water column as revealed by metaproteomics. Proc Natl Acad Sci U S A 115:E400–E408. https://doi.org/10.1073/pnas.1708779115.

46. Fuhrman JA, Schwalbach MS, Stingl U. 2008. Proteorhodopsins: an array of physiological roles? Nat Rev Microbiol 6:488–494. https://doi.org/10.1038/nrmicro1893.

47. Pinhassi J, DeLong EF, Béjà O, González JM, Pedrós-Alió C. 2016. Marine bacterial and archaeal ion-pumping rhodopsins: genetic diversity, physiology, and ecology. Microbiol Mol Biol Rev 80:929–954. https://doi.org/10.1128/MMBR.00003-16.

48. Olson DK, Yoshizawa S, Boeuf D, Iwasaki W, DeLong EF. 2018. Proteorhodopsin variability and distribution in the North Pacific Subtropical Gyre. ISME J 12:1047–1060. https://doi.org/10.1038/s41396-018-0074-4.

49. Steindler L, Schwalbach MS, Smith DP, Chan F, Giovannoni SJ. 2011. Energy starved Candidatus Pelagibacter ubique substitutes light-mediated ATP production for endogenous carbon respiration. PLoS One 6: e19725. https://doi.org/10.1371/journal.pone.0019725.

50. Akram N, Palovaara J, Forsberg J, Lindh MV, Milton DL, Luo H, Gonzalez JM, Pinhassi J. 2013. Regulation of proteorhodopsin gene expression by nutrient limitation in the marine bacterium Vibrio sp. AND4. Environ Microbiol 15:1400–1415. https://doi.org/10.1111/1462-2920.12085.

51. Camacho A. 2006. On the occurrence and ecological features of deep chlorophyll maxima (DCM) in Spanish stratified lakes. Limnetica 25: 453–478.

52. Vermaas WFJ. 2001. Photosynthesis and respiration in cyanobacteria. In Encyclopedia of life sciences. https://doi:10.1038/npg.els.0001670.

53. Wylie JL, Worobec EA. 1995. The OprB porin plays a central role in carbohydrate uptake in Pseudomonas aeruginosa. J Bacteriol 177:3021–3026. https://doi.org/10.1128/jb.177.11.3021-3026.1995.

54. Martiny AC, Coleman ML, Chisholm SW. 2006. Phosphate acquisition genes in Prochlorococcus ecotypes: evidence for genome-wide adaptation. Proc Natl Acad Sci U S A 103:12552–12557. https://doi.org/10.1073/pnas.0601301103.

55. Reistetter EN, Krumhardt K, Callnan K, Roache-Johnson K, Saunders JK, Moore LR, Rocap G. 2013. Effects of phosphorus starvation versus limitation on the marine cyanobacterium Prochlorococcus MED4 II: gene expression. Environ Microbiol 15:2129–2143. https://doi.org/10.1111/1462-2920.12129.

56. Tetu SG, Brahamsha B, Johnson DA, Tai V, Phillippy K, Palenik B, Paulsen IT. 2009. Microarray analysis of phosphate regulation in the marine cyanobacterium Synechococcus sp. WH8102. ISME J 3:835–849. https://doi.org/10.1038/ismej.2009.31.

57. Warr SRC, Reed RH, Stewart WDP. 1985. Carbohydrate accumulation in osmotically stressed Cyanobacteria (Blue-Green Algae): interactions of temperature and salinity. New Phytol 100:285–292. https://doi.org/10.1111/j.1469-8137.1985.tb02779.x.

58. Muñoz-Marín MC, Gómez-Baena G, López-Lozano A, Moreno-Cabezuelo JA, Díez J, García-Fernández JM. 2020. Mixotrophy in marine picocyanobacteria: use of organic compounds by Prochlorococcus and Synechococcus. ISME J 14:1065–1073. https://doi.org/10.1038/s41396-020-0603-9.

59. Yelton AP, Acinas SG, Sunagawa S, Bork P, Pedrós-Alió C, Chisholm SW. 2016. Global genetic capacity for mixotrophy in marine picocyanobacteria. ISME J 10:2946–2957. https://doi.org/10.1038/ismej.2016.64.

60. Noinaj N, Guillier M, Barnard TJ, Buchanan SK. 2010. TonB-dependent transporters: regulation, structure, and function. Annu Rev Microbiol 64: 43–60. https://doi.org/10.1146/annurev.micro.112408.134247.

61. Pajares S, Varona-Cordero F, Hernández-Becerril DU. 2020. Spatial distribution patterns of bacterioplankton in the oxygen minimum zone of the tropical mexican pacific. Microb Ecol 80:519–536. https://doi.org/10.1007/s00248-020-01508-7.

62. Reji L, Tolar BB, Chavez FP, Francis CA. 2020. Depth-differentiation and seasonality of planktonic microbial assemblages in the Monterey Bay upwelling system. Front Microbiol 11:1075. https://doi.org/10.3389/fmicb.2020.01075.

63. Francis CA, Roberts KJ, Beman JM, Santoro AE, Oakley BB. 2005. Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. Proc Natl Acad Sci U S A 102:14683–14688. https://doi.org/10.1073/pnas.0506625102.

64. Herndl GJ, Reinthaler T, Teira E, van Aken H, Veth C, Pernthaler A, Pernthaler J. 2005. Contribution of archaea to total prokaryotic production in the deep Atlantic Ocean. Appl Environ Microbiol 71:2303–2309. https://doi.org/10.1128/AEM.71.5.2303-2309.2005.

65. Martens-Habbena W, Berube PM, Urakawa H, de la Torre JR, Stahl DA. 2009. Ammonia oxidation kinetics determine niche separation of nitrifying archaea and bacteria. Nature 461:976–979. https://doi.org/10.1038/nature08465.

66. Bayer B, Vojvoda J, Offre P, Alves RJ, Elisabeth NH, Garcia JA, Volland JM, Srivastava A, Schleper C, Herndl GJ. 2016. Physiological and genomic characterization of two novel marine thaumarchaeal strains indicates niche differentiation. ISME J 10:1051–1063. https://doi.org/10.1038/ismej.2015.200.

67. Offre P, Kerou M, Spang A, Schleper C. 2014. Variability of the transporter gene complement in ammonia-oxidizing archaea. Trends Microbiol 22: 665–675. https://doi.org/10.1016/j.tim.2014.07.007.

68. Alonso-Sáez L, Waller AS, Mende DR, Bakker K, Farnelid H, Yager PL, Lovejoy C, Tremblay JE, Potvin M, Heinrich F, Estrada M, Riemann L, Bork P, Pedrós-Alió C, Bertilsson S. 2012. Role for urea in nitrification by polar marine archaea. Proc Natl Acad Sci U S A 109:17989–17994. https://doi.org/10.1073/pnas.1201914109.

69. Ingalls AE, Shah SR, Hansman RL, Aluwihare LI, Santos GM, Druffel ER, Pearson A. 2006. Quantifying archaeal community autotrophy in the mesopelagic ocean using natural radiocarbon. Proc Natl Acad Sci U S A 103:6442–6447. https://doi.org/10.1073/pnas.0510157103.

70. Seitz EM, Brockman JP, Sandler SJ, Clark AJ, Kowalczykowski SC. 1998. RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange. Genes Dev 12:1248–1253. https://doi.org/10.1101/gad.12.9.1248.

71. Pratscher J, Dumont MG, Conrad R. 2011. Ammonia oxidation coupled to CO_2 fixation by archaea and bacteria in an agricultural soil. Proc Natl Acad Sci U S A 108:4170–4175. https://doi.org/10.1073/pnas.1010981108.

72. Santoro AE, Richter RA, Dupont CL. 2019. Planktonic marine archaea. Annu Rev Mar Sci 11:131–158. https://doi.org/10.1146/annurev-marine-121916-063141.

73. Norman JS, Lin L, Barrett JE. 2015. Paired carbon and nitrogen metabolism by ammonia-oxidizing bacteria and archaea in temperate forest soils. Ecosphere 6:176. https://doi.org/10.1890/ES14-00299.1.

74. Zhang LM, Offre PR, He JZ, Verhamme DT, Nicol GW, Prosser JI. 2010. Autotrophic ammonia oxidation by soil thaumarchaea. Proc Natl Acad Sci U S A 107:17240–17245. https://doi.org/10.1073/pnas.1004947107.

75. Bayer B, Vojvoda J, Reinthaler T, Reyes C, Pinto M, Herndl GJ. 2019. Nitrosopumilus adriaticus sp. nov. and Nitrosopumilus piranensis sp. nov., two ammonia-oxidizing archaea from the Adriatic Sea and members of the class Nitrososphaeria. Int J Syst Evol Microbiol 69:1892–1902. https://doi.org/10.1099/ijsem.0.003360.

76. Arneborg L. 2004. Turnover times for the water above sill level in Gullmar Fjord. Cont Shelf Res 24:443–460. https://doi.org/10.1016/j.csr.2003.12.005.

77. Bunse C, Israelsson S, Baltar F, Bertos-Fortis M, Fridolfsson E, Legrand C, Lindehoff E, Lindh MV, Martínez-García S, Pinhassi J. 2019. High frequency multi-year variability in Baltic Sea microbial plankton stocks and activities. Front Microbiol 9:3296. https://doi.org/10.3389/fmicb.2018.03296.

78. Gasol JM, Del Giorgio PA. 2000. Using flow cytometry for counting natural planktonic bacteria and understanding the structure of planktonic bacterial communities. Sci Mar 64:197–224. https://doi.org/10.3989/scimar.2000.64n2197.

79. Poretsky RS, Gifford S, Rinta-Kanto J, Vila-Costa M, Moran MA. 2009. Analyzing gene expression from marine microbial communities using environmental transcriptomics. J Vis Exp 24:1086. https://doi.org/10.3791/1086.

80. Pontiller B, Martínez-García S, Lundin D, Pinhassi J. 2020. Labile dissolved organic matter compound characteristics select for divergence in marine bacterial activity and transcription. Front Microbiol 11:588778. https://doi.org/10.3389/fmicb.2020.588778.

81. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200.

82. Li D, Liu CM, Luo R, Sadakane K, Lam TW. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31:1674–1676. https://doi.org/10.1093/bioinformatics/btv033.

83. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform 11:119. https://doi.org/10.1186/1471-2105-11-119.

84. Seemann T. 2013. BARNAP: basic rapid ribosomal RNA predictor. https://github.com/tseemann/barrnap.

85. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60. https://doi.org/10.1038/nmeth.3176.

86. Huson DH, Beier S, Flade I, Górska A, El-Hadidi M, Mitra S, Ruscheweyh HJ, Tappu R. 2016. MEGAN community edition—interactive exploration and analysis of large-scale microbiome sequencing data. PLoS Comput Biol 12:e1004957. https://doi.org/10.1371/journal.pcbi.1004957.

87. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.

88. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing S, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

89. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. 2018. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res 46:W95–W101. https://doi.org/10.1093/nar/gky418.

90. Eddy SR. 2008. A probabilistic model of local sequence alignment that simplifies statistical significance estimation. PLoS Comput Biol 4:e1000069. https://doi.org/10.1371/journal.pcbi.1000069.

91. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340.

92. Price MN, Dehal PS, Arkin AP. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol 26:1641–1650. https://doi.org/10.1093/molbev/msp077.

93. Berger SA, Stamatakis A. 2011. Aligning short reads to reference alignments and trees. Bioinformatics 27:2068–2075. https://doi.org/10.1093/bioinformatics/btr320.

94. Barbera P, Kozlov AM, Czech L, Morel B, Darriba D, Flouri T, Stamatakis A. 2019. EPA-ng: massively parallel evolutionary placement of genetic sequences. Syst Biol 68:365–369. https://doi.org/10.1093/sysbio/syy054.

95. Czech L, Barbera P, Stamatakis A. 2020. Genesis and gappa: processing, analyzing and visualizing phylogenetic (placement) data. Bioinformatics 36:3263–3265. https://doi.org/10.1093/bioinformatics/btaa070.

96. Letunic I, Bork P. 2021. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res https://doi.org/10.1093/nar/gkab301.

97. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37:1530–1534. https://doi.org/10.1093/molbev/msaa015.

98. Legendre P, Gallagher ED. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129:271–280. https://doi.org/10.1007/s004420100716.

99. Beule L, Karlovsky P. 2020. Improved normalization of species count data in ecology by scaling with ranked subsampling (SRS): application to microbial communities. PeerJ 8:e9593. https://doi.org/10.7717/peerj.9593.

100. RStudio Team. 2019. RStudio: integrated development for R, v1.2.5019. RStudio, Inc., Boston, MA. https://www.rstudio.com/.

101. Wickham H. 2017. Tidyverse: easily install and load the 'tidyverse', v1.2.1. https://CRAN.R-project.org/package=tidyverse.

102. Oksanen J, Guillaume Blanchet F, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2019. Vegan: community ecology package, v2.5–5. https://CRAN.R-project.org/package=vegan.