






Seasonal recurrence and modular assembly of an Arctic pelagic marine microbiome

Received: 14 May 2024

Accepted: 13 January 2025

Published online: 03 February 2025

 Check for updates

Taylor Priest ^{1,2,8}✉, Ellen Oldenburg ^{3,8}, Ovidiu Popa³, Bledina Dede ^{1,4}, Katja Metfies ⁵, Wilken-Jon von Appen ⁵, Sinhué Torres-Valdés ⁵, Christina Bienhold ^{1,5}, Bernhard M. Fuchs ¹, Rudolf Amann ¹, Antje Boetius ^{1,5,6} & Matthias Wietz ^{1,5,7}✉

Deciphering how microbial communities are shaped by environmental variability is fundamental for understanding the structure and function of ocean ecosystems. While seasonal environmental gradients have been shown to structure the taxonomic dynamics of microbiomes over time, little is known about their impact on functional dynamics and the coupling between taxonomy and function. Here, we demonstrate annually recurrent, seasonal structuring of taxonomic and functional dynamics in a pelagic Arctic Ocean microbiome by combining autonomous samplers and in situ sensors with long-read metagenomics and SSU ribosomal metabarcoding. Specifically, we identified five temporal microbiome modules whose succession within each annual cycle represents a transition across different ecological states. For instance, *Cand. Nitrosopumilus*, Syndiniales, and the machinery to oxidise ammonia and reduce nitrite are signatures of early polar night, while late summer is characterised by *Amylibacter* and sulfur compound metabolism. Leveraging metatranscriptomes from *Tara Oceans*, we also demonstrate the consistency in functional dynamics across the wider Arctic Ocean during similar temporal periods. Furthermore, the structuring of genetic diversity within functions over time indicates that environmental selection pressure acts heterogeneously on microbiomes across seasons. By integrating taxonomic, functional and environmental information, our study provides fundamental insights into how microbiomes are structured under pronounced seasonal changes in understudied, yet rapidly changing polar marine ecosystems.

Bacteria, archaea and microeukaryotes are the dominant life forms in ocean environments, serving essential trophic roles and mediating the biogeochemical cycling of biologically important elements^{1–3}. These microbes drive and respond to changes in their

surrounding environment, which results in the assembly of distinct communities over spatial and temporal scales^{4,5}. Deciphering how environmental variability shapes population dynamics and the structuring of community dynamics is thus fundamental for

¹Max Planck Institute for Marine Microbiology, Bremen, Germany. ²Institute of Microbiology, ETH Zurich, Zurich, Switzerland. ³Institute for Quantitative and Theoretical Biology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ⁴Ecologie Systématique Evolution, CNRS, Université Paris-Saclay, Agro-ParisTech, Gif-sur-Yvette, France. ⁵Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany. ⁶MARUM – Center for Marine Environmental Sciences, University of Bremen, Bremen, Germany. ⁷Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Oldenburg, Germany. ⁸These authors contributed equally: Taylor Priest, Ellen Oldenburg. ✉e-mail: tpriest@ethz.ch; matthias.wietz@awi.de

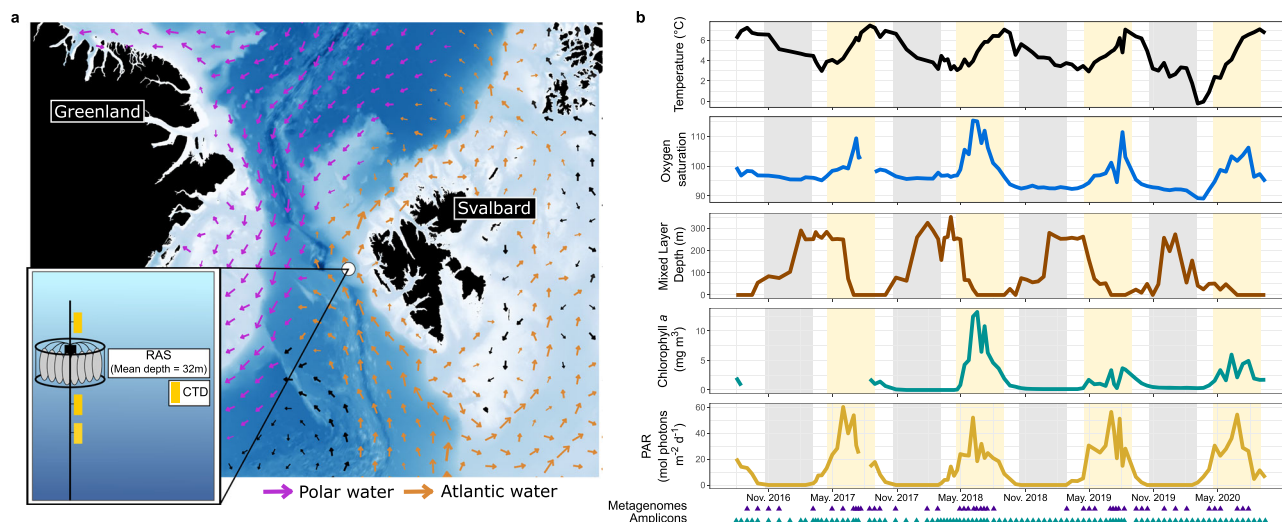


Fig. 1 | The WSC mooring site and environmental conditions between July 2016 and July 2020. a Bathymetric map of the Fram Strait with the mooring location within the West Spitsbergen Current. Arrows represent average current velocities over the four-year sampling period at the average depth of the moored autonomous

sampler (32 m). **b** Water temperature, oxygen saturation, mixed layer depth, chlorophyll concentrations measured from mooring-attached sensors and photo-synthetically active radiation (PAR) derived from AQUA-MODIS satellite data. The shaded grey and yellow represent the periods of polar night and day, respectively.

understanding the structure and function of ecosystems and how they respond to change.

Long-term observations have uncovered recurrent and transient dynamics in microbial communities across daily, seasonal and annual timescales. In temperate ecosystems, communities are predominantly structured by seasonal variability, with broadly recurrent fluctuations of taxa on annual scales^{6–11}. These patterns have led to the conclusion that microbial responses to biological and environmental shifts are predictable¹². However, recent evidence indicates that microeukaryotes exhibit weaker temporal structuring than prokaryotes¹³, suggesting different controlling mechanisms. Furthermore, high-frequency sampling has shown that population dynamics are highly ephemeral, undergoing rapid, short-lived fluctuations that transpire over days^{14–16}. Thus, while microbial communities are structured over time, they also undergo constant flux, reflecting the dynamic nature of ocean ecosystems.

Despite the wealth of knowledge gained from long-term observations, the focus primarily on taxonomic dynamics and on temperate and subtropical ecosystems has left many questions unanswered. In particular, it remains unclear how compositional shifts across environmental gradients translate to changes in the functionality of microbial communities. Since distantly related organisms can perform similar metabolic functions^{17,18}, taxonomic information alone is a relatively weak proxy for ecological landscapes or ecosystem function. Therefore, long-term observations that integrate taxonomic, functional, and environmental information are greatly needed. This is particularly important in the polar oceans, where long-term observations are rare and unprecedented changes are taking place because of climate warming.

To address this, we investigated the dynamics of prokaryotic (here, this term is used operationally to refer to Bacteria and Archaea) and microeukaryotic communities from a taxonomic and functional perspective over a four-year period in an Arctic Ocean ecosystem, the West Spitsbergen Current (WSC). The WSC constitutes the primary entry route for Atlantic water into the Arctic Ocean and is characterised by pronounced seasonal variability in environmental conditions, archetypal of high-latitude ecosystems. However, unlike other Arctic Ocean ecosystems, the WSC remains ice-free year-round, due to the warm North Atlantic water. Therefore, the WSC represents a model system for investigating the dynamics of microorganisms and ecosystem function in the context of pronounced seasonal variability. In

addition, it can provide valuable insights into potential future shifts in Arctic Ocean ecosystems, given the progressive northward expansion of Atlantic water influence¹⁹. By leveraging moorings fitted with autonomous sampling and measuring devices combined with 16S and 18S rRNA gene amplicon and PacBio HiFi metagenome sequencing, we continuously tracked taxonomic, functional and environmental dynamics over time. Using a Fourier transformation-based approach, we demonstrate that prokaryotic and microeukaryotic communities exhibit annually recurrent, seasonally structured dynamics that are characterised by a succession across five distinct temporal modules. Each module features unique taxonomic and metabolic signatures, representing distinct ecological states, and are subject to different environmental selection pressures. Our study provides a multi-year ecosystem catalogue from the Arctic that integrates taxonomic, functional and environmental information, and provides fundamental insights into the dynamics and structuring of microbiomes across pronounced environmental gradients.

Results and discussion

The West Spitsbergen Current shows pronounced temporal structuring

We first investigated how environmental conditions in the WSC are structured over intra- and inter-annual scales. For this, we combined data collected from in situ sensors, including temperature, salinity, oxygen saturation and chlorophyll *a*, with satellite-derived values of photosynthetically active radiation (PAR) (Supplementary Data 1). Our measurements were derived from the epipelagic layer but, due to the movement of the mooring by currents, sampling depths varied from 20 to 100 m. Owing to the inclusion of CTD sensors at different depths, we were also able to determine the lower bound of the mixed layer depth (MLD). Each annual cycle was characterised by pronounced shifts in environmental conditions (Fig. 1). As expected, these shifts followed the transition between polar night and polar day, which is a major force stimulating biological dynamics in the Arctic²⁰. The end of the polar night was marked by an increase in PAR in April, which continued to rise until a maximum of, on average, 38 $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$ in June. The increasing PAR also coincided with rising water temperatures, which peaked at -7°C in August/September before decreasing again to $<4^\circ\text{C}$ between December–May. Changes in MLD were inversely related to temperature (Pearson correlation: $R = -0.47$, $p < 0.05$), with abrupt events of shallowing to $<5\text{ m}$ in June and

deepening to >200 m between December and January. Chlorophyll concentrations showed a lagged association with PAR, peaking between June and August (Fig. 1b). However, the magnitude and timing of the chlorophyll peak varied across years, from $3.35 \mu\text{g L}^{-1}$ in July 2019 to $13.22 \mu\text{g L}^{-1}$ in June 2018, indicating differences in phytoplankton bloom phenologies. The WSC thus exhibits pronounced intra-annual shifts in environmental conditions, presenting an ideal ecosystem to study seasonally driven biological dynamics.

The intra- and inter-annual temporal structuring of communities

We next investigated the temporal structuring of prokaryotic and microeukaryotic communities from a taxonomic perspective. Using autonomous Remote Access Samplers (RAS), we collected 97 samples at, on average, fortnightly resolution that were used for 16S and 18S rRNA gene sequencing (Supplementary Data 2). From these, we recovered 3629 bacterial, 119 archaeal and 3019 microeukaryotic amplicon sequence variants (ASVs) (Supplementary Data 3, 4).

The alpha diversity of prokaryotic and microeukaryotic communities exhibited distinct trends within each annual cycle (Figs. 2a, b and Supplementary Data 5). For prokaryotic communities, we observed a tight coupling between Species Richness (R), Evenness (E) and Shannon Diversity (H'), evidenced through significant positive Pearson's correlations (Fig. 2c and Supplementary Data 6). The alpha diversity metrics followed a unimodal fluctuation within each annual cycle for prokaryotic communities, reaching a peak during polar night (December–March) with a mean R of 1210 ± 208 , E of 0.78 ± 0.04 and H' of 5.3 ± 0.36 . The higher diversity in winter mirrors previous observations in temperate and polar regions^{21–23} and is associated with the deepening of MLD, which drives the mixing and dilution of previously stratified communities²⁴. The influx of PAR and rapid shallowing of MLD at the end of polar night coincided with a sharp drop in alpha diversity, with the lowest values observed in June ($R = 762 \pm 156$, $H' = 4.50 \pm 0.34$, $E = 0.68 \pm 0.03$). Hence, shifts in prokaryotic alpha diversity were tightly coupled to MLD, supported by strong positive Pearson's correlations; R ($r = 0.71$, 95% CI [0.59, 0.79], $p = 4 \times 10^{-16}$), E ($r = 0.43$, 95% CI [0.25, 0.58], $p = 1.08 \times 10^{-5}$) and H' ($r = 0.58$, 95% CI [0.43, 0.70], $p = 5.7 \times 10^{-10}$) (Supplementary Data 6). This observation is in contrast to previous reports of temperature^{25–27}, ocean currents²⁸ and day length^{10,22} as the primary drivers of epipelagic bacterial diversity. The greater role of stratification in shaping prokaryotic diversity observed here may be a feature unique to polar oceans, where seasonal shifts in MLD are more pronounced²⁹ and can be influenced by sea-ice dynamics³⁰. However, MLD has rarely been measured or incorporated before, and thus its consideration in future studies will help to ascertain whether its influence varies across latitudes.

In contrast to prokaryotes, microeukaryotic communities exhibited a bimodal fluctuation in alpha diversity in each annual cycle. The bimodal pattern was reflected in a peak in H' in both polar night (February–March) and polar day (July–August). However, during polar night, the increased H' was underpinned by a reduced R and increased E , compared to an increased R during polar day. These temporal fluctuations in microeukaryotic diversity are in contrast to observations from the temperate San Pedro time-series (SPOT), where H' was shown to be invariable over time across the top 500 m of the water column¹³. The difference may indicate a stronger seasonal structuring of microeukaryotic communities in high-latitude ocean ecosystems as a consequence of pronounced environmental variability, particularly sunlight and mixing, supported by the negative correlation of MLD and R of microeukaryotes (Pearson's $R = -0.48$, 95% CI [-0.62, -0.31], $p = 7.2 \times 10^{-7}$). Furthermore, the distinct patterns observed in prokaryotes and microeukaryotes provide evidence that the diversity of these communities is shaped by different forces.

We next assessed how the composition of communities was structured over time. We observed a coherent structuring of prokaryotic and microeukaryotic communities based on the month of sampling. That is, communities sampled from the same month across years were typically more similar than from other months in the same year (Fig. 3a, b), reflecting an annual ecosystem clock. For prokaryotic communities, this pattern aligns with previous observations, including month-based clustering at the Banyuls Bay Microbial Observatory in the Mediterranean³¹ and lowest pairwise beta diversity values at 12-month intervals reported from the temperate English Channel L4³² and SPOT¹³ time-series. However, in contrast, microeukaryotic communities showed weaker temporal structuring over multi-annual scales at SPOT¹³. By comparing the within- and between-month dissimilarities across years through the convex hull areas within the NMDS ordination (Fig. 3 and Supplementary Data 7), we demonstrate a clear distinction in the temporal recurrence of prokaryotic and microeukaryotic composition. Prokaryotic communities were more cohesive across years in February–March and more variable during June–July, with convex hull areas of -0.025 and -0.13, respectively. In contrast, microeukaryotic communities were more cohesive during August and more variable during January–March, with convex hull areas of 0.06 and -0.43, respectively. However, maximal inter-annual differences in microeukaryotic communities were observed in April, in the phase before the spring bloom. This indicates that the assembly of the spring bloom is not predictable, and only later in the summer, the increased richness of microeukaryotic populations (high R) assembles into a cohesive, annually recurrent structure. In previous years, a high inter-annual variability during polar day has been observed in microeukaryotic communities in this region^{33,34}, so this pattern may change with time and reflect climate-induced or natural decadal variations³⁴. The recurrent structuring of microeukaryotic communities during summer could be anticipated to stimulate predictable dynamics in prokaryotes, owing to specialised substrate niches and specific interactions^{35,36} as observed in temperate systems^{7,8}. Indeed, it triggers a pronounced response of a minority of prokaryotic populations, evidenced through reduced R and E , but their emergent population structure shows high inter-annual variability, which may be a result of selection on a functional level and stochastic processes. In contrast, polar night manifests species-rich and compositionally cohesive prokaryotic communities, suggesting that vertical mixing drives the replenishment of the microbiome back to a “standing stock”.

Temporal dynamics of ASVs and community gene content

To gain a deeper understanding of ecosystem dynamics, we complemented the amplicon dataset with long-read metagenomes and investigated temporal fluctuations of ASVs and community gene content. We generated PacBio HiFi metagenomes from 47 time points spanning the 4 years (Supplementary Data 2), resulting in 8.7 million HiFi reads with an average length of 6 kbp. Domain-level classification of the HiFi reads revealed that the majority, 7.9 million, were associated with Bacteria, while 0.47 million were associated with Archaea and 0.27 million with Eukarya. Owing to the length and quality of HiFi reads and the loss of information that typically occurs during metagenome assembly, we elucidated the genetic content of the WSC microbiome by extracting genes directly from the reads. Through this, we recovered 48 million gene sequences that were grouped into 306,088 non-singleton clusters based on a 95% average nucleotide identity cut-off. Combining the gene cluster collection with the ASV data provides a rich multi-year ecosystem catalogue to elucidate community assembly processes and temporal dynamics on a taxonomic and gene content level.

To unravel the temporal dynamics of ASVs and gene clusters, we employed an approach based on Fourier transformations and the determination of oscillation signals. Fourier transformations convert

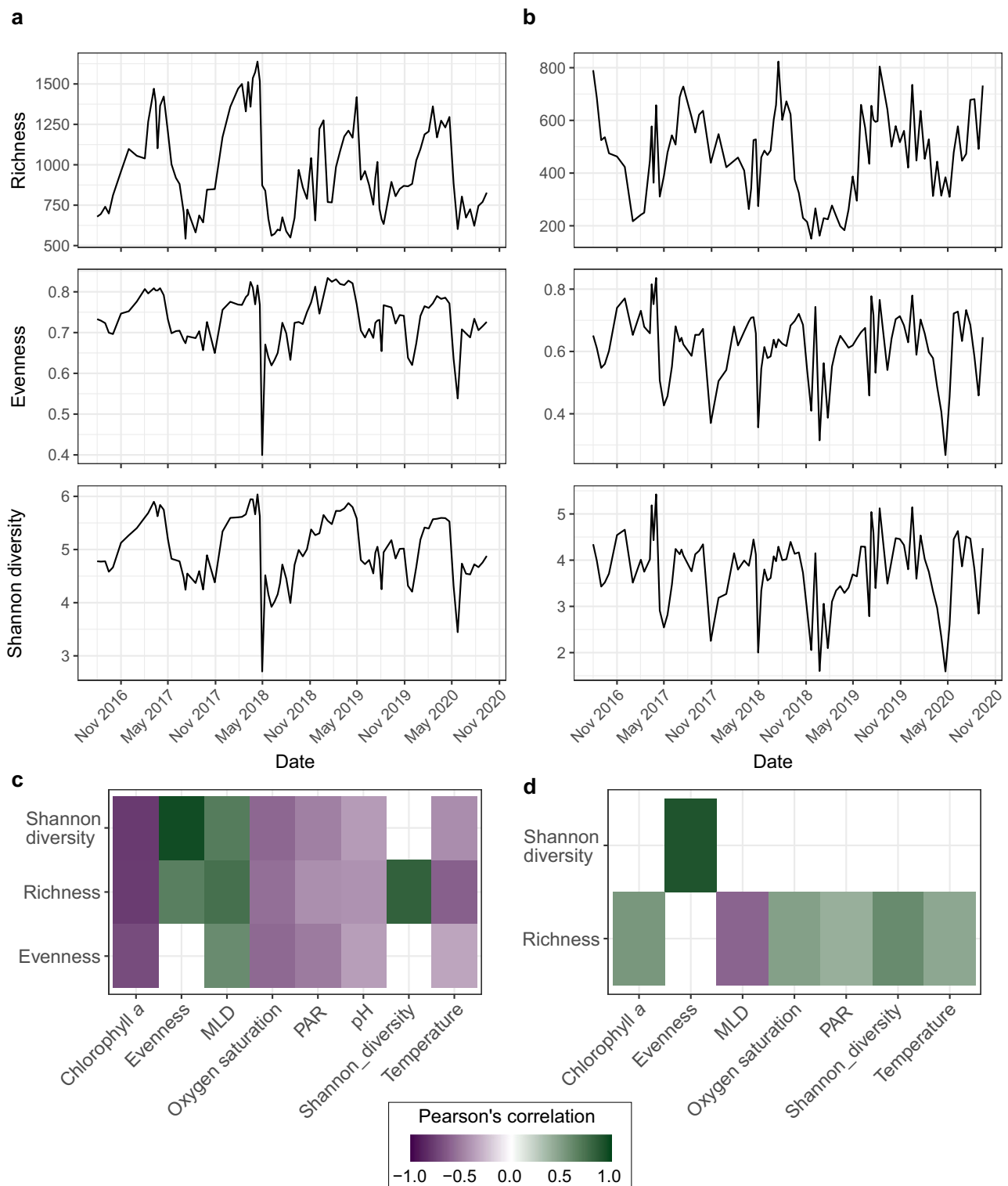


Fig. 2 | The diversity of prokaryotic and microeukaryotic communities is structured differently over time. All values of diversity shown represent the mean value after performing 100 iterations of rarefying and metric calculation. Richness, Evenness and Shannon diversity of **a** prokaryotic and **b** microeukaryotic

communities. Associations between diversity indices and environmental conditions were tested through Pearson's correlations. Statistically significant associations ($p < 0.05$) after applying Benjamini–Hochberg multiple testing correction are visualised for **c** prokaryotic and **d** microeukaryotic alpha diversity indices.

abundance data into frequencies, resulting in wave-like signals that can be evaluated in terms of peak/trough dynamics, hereon termed oscillation signals. We determined that 18% of prokaryotic ASVs, 15% of microeukaryotic ASVs and 58% of gene clusters exhibited a single oscillation each year, reflecting a unimodal fluctuation in abundance with a single peak and trough (Fig. 3c). Although only capturing a

fraction of the diversity, these annually oscillating ASVs and gene clusters comprised a large proportion of communities across all time points, with an average relative abundance of 67% for prokaryotic ASVs, 55% for microeukaryotic ASVs, and 28% for gene clusters. These findings expand on previous reports of seasonally recurrent patterns in temperate prokaryotic communities at more coarse-grained resolutions,

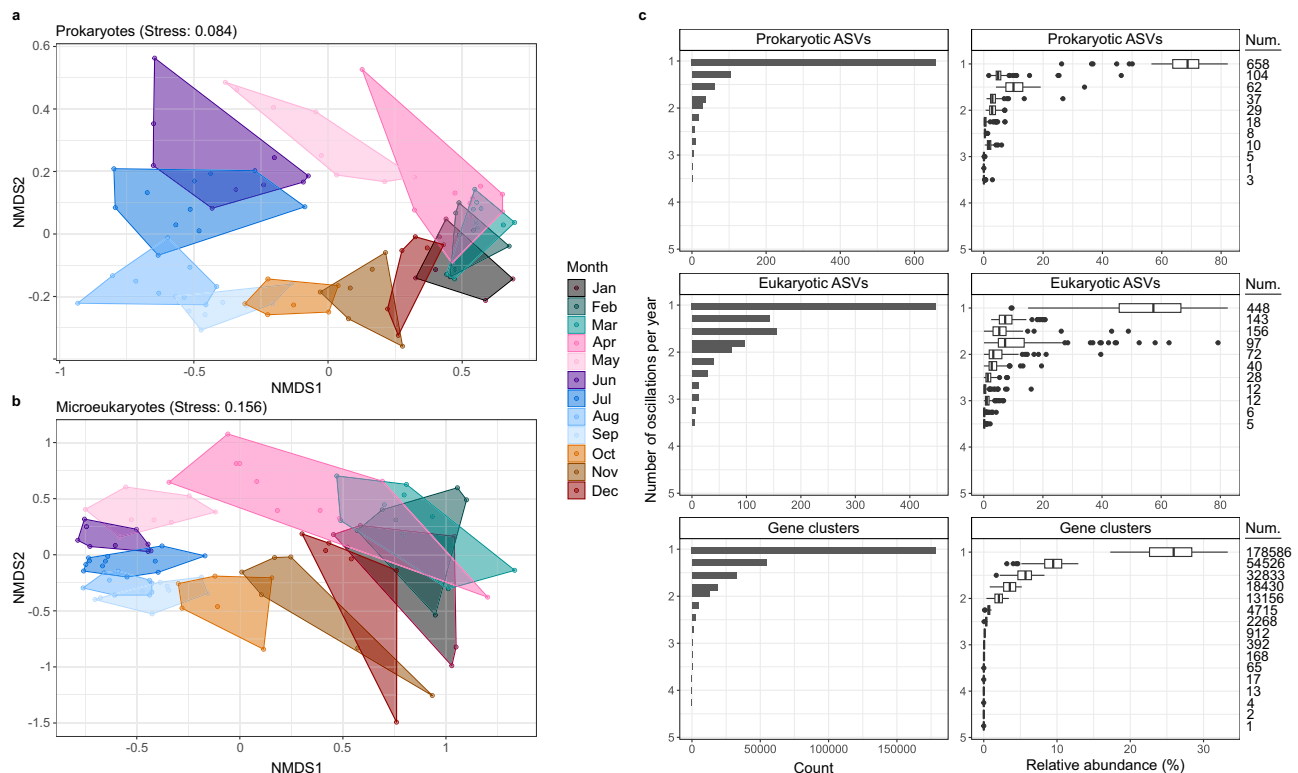


Fig. 3 | Structuring of community composition and the oscillations of ASVs and genes across years. Non-metric multi-dimensional scaling of Bray-Curtis dissimilarities calculated after rarefying and Hellinger transforming ASV count data for **a** prokaryotic and **b** microeukaryotic communities over 4 years, wherein colours indicate months. **c** The count (left) and relative abundance (right) of prokaryotic ASVs, microeukaryotic ASVs and gene clusters for each oscillation signal. The relative abundances of ASVs and genes were subject to Fourier

transformation, and oscillation signals were determined based on the amplitude of peak/trough dynamics within each annual cycle. An oscillation of 1 indicates a single peak/trough per year, reflecting a unimodal annual fluctuation in abundance. Boxplots illustrate the median, 25 and 75% percentile of the number and relative abundance of ASVs and gene clusters exhibiting each oscillation signal. The number of ASVs and gene clusters observed for each oscillation signal are included next to each boxplot.

such as OTUs¹⁰ and higher taxonomic levels¹³, and are in line with those on microdiversity temporal dynamics from a coastal temperate region⁸—suggesting comparable ecological forcing in polar oceans.

The dominance of annually oscillating ASVs raises the question, “To what extent are these patterns deterministic?”. More than 15 years ago, Fuhrman and colleagues proposed that “annually recurrent microbial communities can be predicted from ocean conditions”¹². However, only more recently, with technological advancements and the growing quantity and resolution of data, is predictive ecology becoming more feasible. To contribute to this, we assessed the timing and order of ASV oscillations across each annual cycle. We found that 20% of prokaryotic ASVs consistently oscillated in the same order each year, while 51% reached their peak within the same 30-day window. Therefore, while recurrent dynamics of ASVs are largely bound within temporal windows, the composition of co-occurring populations can vary across years, likely reflecting the influence of trophic interactions and the ephemeral nature of population dynamics^{14,37}. A similar pattern was recently described from prokaryotic communities in the NW Mediterranean, where seasonally recurrent taxa exhibited changes in their connections across years³¹. Despite these variations, the periodic timing of recurrent population dynamics within narrow temporal windows provides support towards deterministic patterns in ocean ecosystems—at least in regions that experience marked environmental variability over seasonal timescales, e.g. polar and temperate regions.

Seasonal recurrence is underpinned by transitions across distinct ecological states

Upon identifying annually oscillating ASVs and gene clusters, we next sought to elucidate how these patterns translate to ecological shifts

within the WSC microbiome. To achieve this, integrated taxonomic and functional information is essential. Although taxonomy and function are typically connected through metagenome-assembled genomes (MAGs) in metagenomic studies, this often only captures a fraction of the sequenced material and thus can limit our perspective on microbiome dynamics. This is particularly true in this study, with the recovered 91 species-representative MAGs only capturing between 2.5–35.2% of sequenced genomes (Supplementary Fig. 1 and Supplementary Data 8), largely due to low sequencing depth. An alternative strategy, which is enabled by long-read sequencing, is to directly assess taxonomy and function on a read level. To this end, we performed taxonomic classification on PacBio HiFi reads using a blast and last-common ancestor approach with the GTDB as a reference, and leveraged this information to assign taxonomic labels to gene clusters. Through this, 88% of annually oscillating gene clusters were assigned a phylum-level taxonomy, 84% were assigned a class and 56% were assigned a genus. To translate gene clusters to functional dynamics, we annotated the annually oscillating gene clusters against the EGGNOG database. Of the 178,586 annually oscillating gene clusters, 68% were assigned to reference orthologs and were subsequently grouped into clusters based on the functional annotation of the matching references. This resulted in 7629 unique functional clusters whose oscillations were determined over time.

To investigate temporal structuring on a taxonomic and functional level, we combined the oscillations of functional clusters and ASVs to build a correlation-based network. Correlation co-occurrence networks are powerful for disentangling community dynamics and organismal interactions over spatial and temporal scales³⁸. Here, we compared the oscillation signals of ASVs and gene functions through

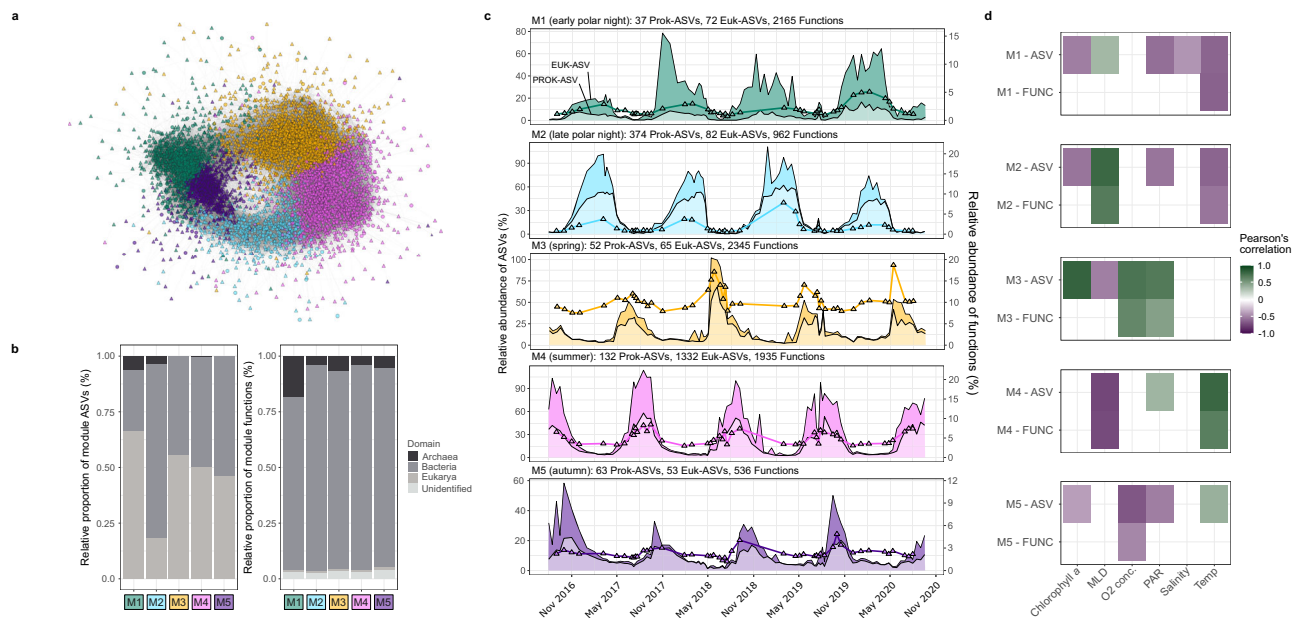


Fig. 4 | Ecosystem modules, their temporal dynamics and association with environmental conditions. **a** Co-occurrence network constructed from significant positive Pearson correlations (>0.7) between prokaryotic, microeukaryotic ASVs and functional cluster oscillation signals. Circle nodes = ASVs, Triangle nodes = functional clusters. **b** The relative proportion of module ASVs and functions associated with the three domains of life. **c** The temporal dynamics of module

prokaryotic ASVs (lighter-coloured area), microeukaryotic ASVs (darker-coloured area) and functional clusters (line). **d** Pearson's correlations between the combined relative abundance of module ASVs and module functional clusters against measured environmental factors (only those with a $p < 0.05$ after Benjamini–Hochberg multiple testing correction are shown).

Pearson's correlation and retained only those with a strong, positive coefficient ($R > 0.7$, FDR-based adjusted $p < 0.05$). We subsequently built a network using these correlation coefficients as the edges and the respective ASVs and functions as the nodes. Using the Louvain algorithm, we partitioned the network into five modules comprised of co-oscillating ASVs and functions (Fig. 4a) that represent distinct temporal states in the WSC. Thus, each annual cycle was characterised by a succession across these temporal modules (Fig. 4c).

To uncover the ecological shifts associated with module succession, we compared the composition of modules and the environmental conditions under which they prevail. The modules differed markedly in their diversity and composition of taxonomic and functional components (Fig. 4b), from a low number of ASVs (prokaryotic = 37, eukaryotic = 72) and high number of functions ($n = 2165$) in the early polar night module M1 to a high number of both ASVs (prokaryotic = 132, eukaryotic = 1332) and functions ($n = 1935$) in the summer module M3. Exploring these patterns further revealed clear differences in the distribution of ASVs and functions attributed to *Archaea*, *Bacteria* and *Eukarya* across the modules. The functional composition of modules was dominated by *Bacteria* year-round, although a threefold enrichment of archaeal functions (18% of module functions) was observed during early polar night (module M1). From a taxonomic perspective, eukaryotic ASVs dominated the early polar night (M1), spring (M3) and summer modules (M4), while recurrent bacterial ASVs were more prominent during late polar night (M2) and autumn (M5). To ensure that these patterns were reflective of biological signals and not a consequence of uneven sequencing depth across metagenomes or the frequency of sampling across seasonal periods, we assessed whether the composition of modules changed after subsampling to only one sample per month and rarefying the ASV and functional cluster datasets. The subsampling had a negligible impact on the diversity and composition of modules (Supplementary Fig. 2). Although this does not exclude that additional sampling may further resolve these patterns, it supports the taxonomic and functional differences observed between modules.

The temporal succession of modules represents a transition across distinct ecological states within each annual cycle. These ecological shifts are evidenced through the unique taxonomic (Fig. 5a) and metabolic signatures of each module (Fig. 5b). The early polar night module M1 was taxonomically distinguished by *Cand. Nitrosopumilus*, SAR86, *Porticoccus* and *Syndiniales* and functionally distinguished by nitrogen metabolism, specifically ammonia oxidation (*amoABC*) and nitrite reduction (*nirK* and *norBC*), as well as carbon fixation machinery of the hydroxypropionate/hydroxybutyrate cycle³⁹. Read taxonomic classification supported that the ammonia oxidation and carbon fixation machinery was attributed to members of *Nitrosopumilaceae*. The dominance of *Cand. Nitrosopumilus* ASVs during this period and their large degree of connectedness in the network to other ASVs and functions within module M1 (Supplementary Fig. 3) suggests their integral role in polar night microbiomes. This finding expands on previous observations of higher *Cand. Nitrosopumilus* abundances during polar night in southward flowing Arctic waters⁴⁰ and increased ammonia oxidation rates in Antarctic coastal waters⁴¹. A shift in dominance from module M1 to M2 marked an ecological transition during mid- to late-polar night, underpinned by the emergence of *Thalassobius*, *Arenicellaceae* and Radiolarians group C. Despite this taxonomic shift, few distinguishing functional features could be elucidated – resolving the functional transition from early to late-polar night may require increased sampling resolution and the incorporation of additional strategies to improve functional annotations of genes, such as protein structure-based comparisons. The onset of solar radiation and rapid transition from module M2 to M3 marks the end of the polar night. Module M3 represents the early spring phytoplankton bloom, supported by correlations with PAR and chlorophyll (Fig. 4c and Supplementary Data 9) combined with the emergence of photosynthetic organisms (*Pseudo-nitzschia* and *Micromonas*). Thus, this module was dominated both by phototrophs and heterotrophic bacteria, with the latter recognised as primary responders to phytoplankton blooms and their carbohydrate exudates in temperate regions^{7,35}, including *Polaribacter*, *Aurantivirga* and SAR92.

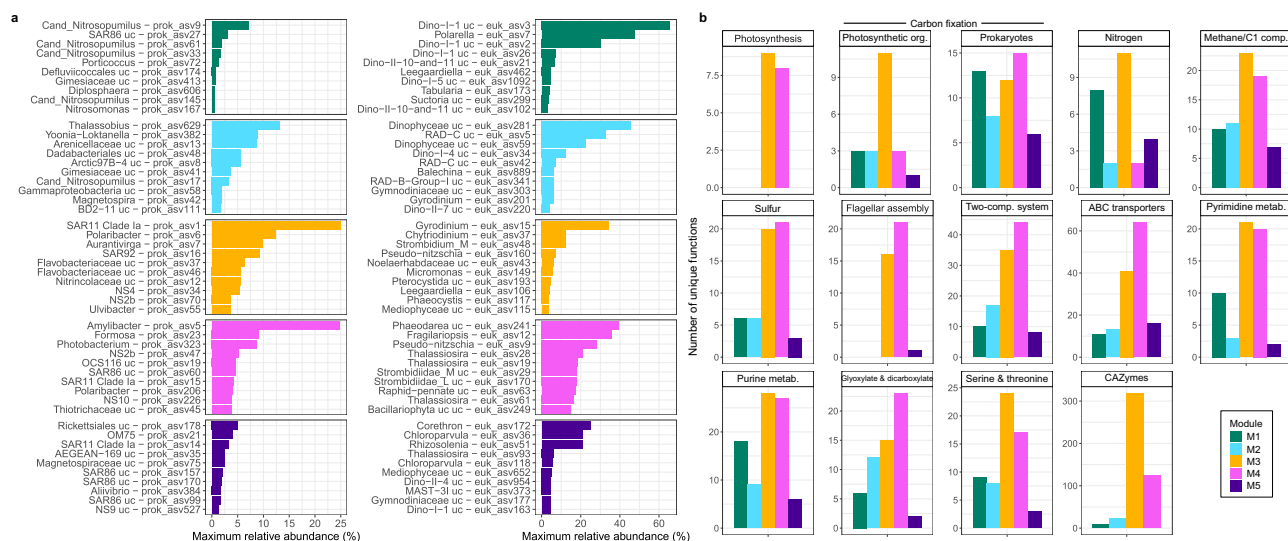


Fig. 5 | Modules are phylogenetically and functionally distinct. a The maximum relative abundance of the ten most abundant prokaryotic and microeukaryotic ASVs in each module. **b** Composition of KEGG metabolic pathways that exhibited

Functionally, module M3 was enriched in photosynthesis machinery, nitrogen metabolism (glutamate synthesis and degradation and the *nrtABC* nitrate/nitrite uptake system), carbohydrate-active enzymes and organosulfur compound utilisation, including DMSP demethylation (*dmdBCD*), DMSO to DMS conversion (*dmsBC*), taurine uptake (*tauABC*) and sulfite reduction (*soeABC*). Taurine uptake machinery was almost exclusively attributed to *Alphaproteobacteria*, particularly *Pelagibacteraceae* and *Rhodobacteraceae*, while the machinery to convert DMSO to DMS was attributed to *Flavobacteriaceae* and *Opi-tales* and the sulfite reduction machinery to *Rhodobacteraceae* and *Pseudomonadales*.

Sulfur metabolism persisted as a key feature in the summer module M4 but transitioned to both inorganic and organic sulfur utilisation. In particular, summer was characterised by the machinery for sulfur oxidation (*soxABCXYZ*), DMSP demethylation (*dmdA*), methanethiol oxidation (*MeSH*), and alkane sulfonate utilisation (*msuD*), which was accompanied by an enrichment in flagellar assembly, two-component system machinery and ABC transporters. The prokaryotic community in module M4 was dominated by *Amylibacter*, along with other heterotrophic bacteria affiliated with *Formosa* and NS2b, while the eukaryotic community was comprised of diverse functional groups, including *Fragilariopsis* (photosynthetic diatoms) and *Phaeodarea* (heterotrophic protists). *Amylibacter* are key contributors to organosulfur compound metabolism in temperate coastal waters during summer^{42,43}. However, we could also attribute the machinery to oxidise sulfur (SOX system), demethylate DMSP (*dmdA*) and reduce sulfite (*soeAB*) to *Amylibacter*, based on read taxonomic classification against GTDB representatives, suggesting their capacity to utilise diverse sulfur sources⁴⁴.

The taxonomic and metabolic signatures of each module demonstrate that the WSC microbiome is structured by a succession across five distinct ecological states within each annual cycle that is largely coupled to changes in environmental conditions. Previous taxonomic-centred analyses of prokaryotic communities from temperate and tropical ecosystems have also reported recurrent dynamics structured by seasonality. However, the proportion of ASVs exhibiting seasonal recurrence and the number of temporal modules appears to be higher in the WSC. For instance, in the NW Mediterranean, only 4% of prokaryotic ASVs, constituting a relative abundance of 47%, showed seasonal recurrence and could be grouped into three distinct seasonal clusters⁴⁵. In a coastal temperate region that lacks pronounced

the largest variance in the number of assigned functions between modules along with the number of carbohydrate-active enzyme (CAZyme) gene families.

phytoplankton blooms but experiences large environmental variability, recurrent dynamics of prokaryotes primarily partitioned into summer and winter groups⁴⁶. The higher prevalence of recurrent dynamics and their organisation into narrower temporal modules in the WSC may indicate stronger selective pressure arising from the more pronounced seasonal environmental variability.

Functional dynamics in the WSC are reflective of microbiomes across the wider Arctic Ocean

We next assessed whether the ecological states in the WSC are representative of microbiomes across the wider Arctic Ocean. After passing through the Fram Strait, the WSC turns east and integrates into the Arctic circumpolar boundary current, which circulates along the continental slope at intermediate water depths (>200 m)⁴⁷. The Atlantic water within the boundary current later separates from the slope due to pronounced bathymetric features and subsequently enters the deep central basins. As a result, Atlantic water derived from the WSC has a significant influence across the wider Arctic Ocean—an influence that is currently expanding¹⁹. Given this, we hypothesised that the WSC can serve as a valuable model system for assessing ecological dynamics within Arctic Ocean microbiomes. To test this, we investigated the activity of module functions across the Arctic Ocean using metatranscriptomes generated from the *Tara Ocean Polar Circle* (TOPC) expedition. We observed that functions within the spring module M3 and autumn module M5 exhibited the highest transcription levels in surface and deep chlorophyll maximum waters (Fig. 6a, b), followed by those of the summer module M4. In contrast, the transcription of functions in mesopelagic depths was dominated by those of the autumn module M5, while signals of spring and summer functions were diminished (Fig. 6b). The only exception to these patterns was the station within the Fram Strait (TARA_163), whose surface waters were dominated by the transcription of functions of the late polar night module M2. The distinct functional composition observed at TARA_163 likely stems from the influence of southward flowing polar water that occurs in this central Fram Strait location (14°E), supported by the lower surface water temperatures (<2 °C) compared to the WSC. Exploring the composition of module functions further, we identified that transcription signals of module M3 in surface and deep chlorophyll maximum waters of the Arctic Ocean were dominated by functions associated with photosynthesis and secondary metabolism, in line with previous findings from the same dataset⁴⁸, while module

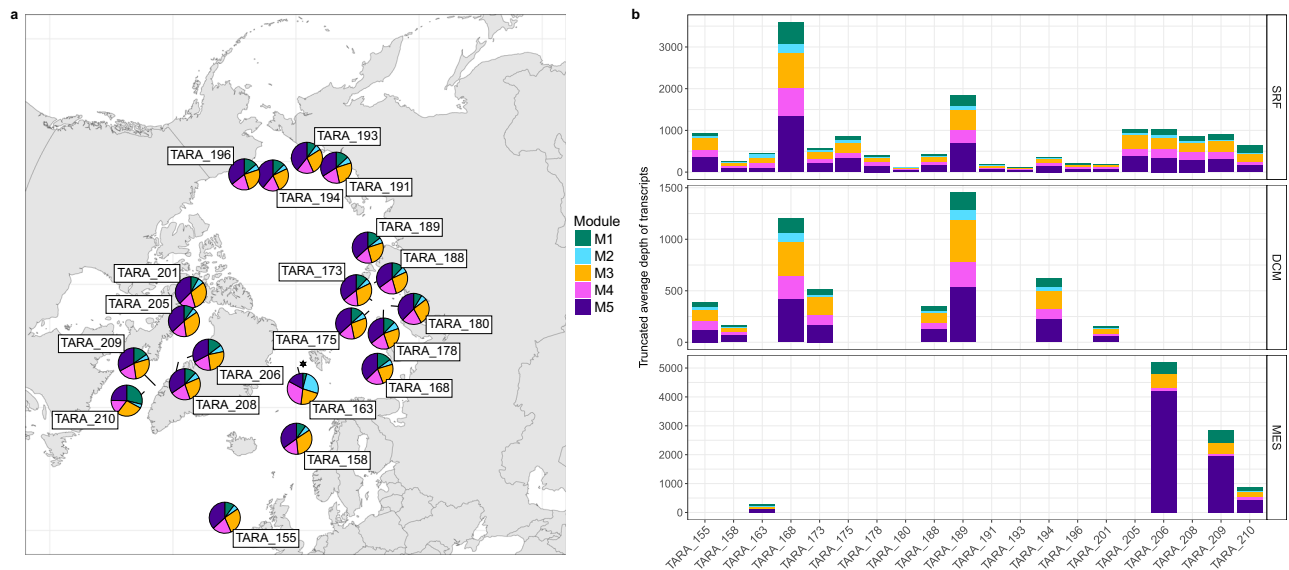


Fig. 6 | Transcription of module functions across surface, deep-chlorophyll maximum and mesopelagic depth layers of the Arctic Ocean. Transcription of module functions was determined by read recruitment from metatranscriptomic data generated from surface (SRF), deep-chlorophyll maximum (DCM) and mesopelagic (MES) depth layers of the Arctic Ocean during the *Tara Ocean* Polar Circle expedition. For each module function, the 80% truncated average depth (TAD80) of transcripts was calculated. The mean of the TAD80 values for all functions in

each module was determined, providing a mean of transcription at the module level. **a** Map illustrating the relative proportion of module functions transcription across surface waters at TOPC sampling stations. The black star indicates the location of the RAS mooring. **b** The mean TAD80 of transcripts of module functions across depth layers and stations. Bars not present in DCM and MES layers are where no metatranscriptomic data was available.

M3 functions also dominated the transcription of energy metabolism processes (Supplementary Fig. 5). In contrast, the transcription signals associated with module M5 were primarily underpinned by ribosomal synthesis and transcription/translation machinery. Given that the TOPC expedition took place between May and October, these patterns not only support that the annually oscillating functions in the WSC are actively transcribed across the wider Arctic Ocean but also confirm that they largely proliferate during similar temporal periods. These observations support that the WSC can serve as a valuable system for gaining insights into and inferring the ecological dynamics of Arctic Ocean microbiomes.

Selection pressure is heterogeneous across seasonal ecological states

The recurrent dynamics of ASVs, genes and functions and their assembly into cohesive, ecological modules suggest that the WSC microbiome is predominantly shaped by deterministic processes. The primary mechanisms that contribute to driving the deterministic structuring of microbiomes are environmental selection and organismal interactions. Environmental selection is considered the primary force shaping the distribution of populations and the assembly of microbial communities in ocean ecosystems⁴⁹. However, we lack a mechanistic understanding of how environmental selection operates. In particular, it is unclear whether the environment primarily selects for a function or for a specific organism with a function. Recent evidence has shown that the taxonomic composition of a microbiome can change while the gene content remains conserved^{50,51}, reflecting the prevalence of metabolic redundancy across microbial taxa. In addition, studies on surface ocean microbial communities have demonstrated incongruencies in structuring on a taxonomic and functional level, with strong selection pressure on functional groups but weak selection pressure on the taxonomic composition within functional groups¹⁸. These observations provide evidence for the decoupling of taxonomy and function within microbiomes and suggest that selection may act on a functional level.

To assess the influence of selection on shaping microbiome dynamics, we investigated the composition and structuring of gene clusters within functions over time. We focus on the spring module M3 and polar night module M2, as they represent two contrasting scenarios. We first assessed the proportion of functions comprised of multiple, as opposed to single, gene clusters as a proxy for functional redundancy. In module M3, 85% of functions are comprised of multiple gene clusters, compared to only 67% in module M2, indicating more functional redundancy in spring compared to late polar night (Fig. 7a). Along with increased redundancy, a higher proportion of multi gene-cluster functions exhibited a strong positive linear relationship between function abundance and the diversity of the contained gene clusters in module M3 than module M2, 57 and 32%, respectively (Fig. 7b). That is, functional dynamics in spring are primarily driven by the concurrent growth of metabolically overlapping organisms. Exploring the dynamics of redundant functions further, we observed the same gene cluster dominating each year in 73% of functions during late polar night compared to only 55% in spring (Fig. 7c). Taken together, these observations suggest that selection pressure is stronger on a functional level in spring.

Integrating the observations of functional dynamics with those of ASVs further supports seasonal variations in environmental selection pressure. The late polar night period is characterised by high species richness and evenness (Fig. 2) with a large number of recurrent populations that assemble into a consistent community structure across years (Fig. 3). Amongst the recurrent populations of late polar night, we also observe taxa that are typically associated with deeper water layers, such as *Dadabacteriales*. Combined with the gene cluster and functional dynamics, these observations suggest that deeper vertical mixing during polar night drives a recurrent homogenisation of epipelagic communities, with either a weak environmental selection pressure or selection acting on an organismal level. In contrast, spring is characterised by low species richness and evenness and high inter-annual variability in beta diversity, driven by only a small assemblage of recurrent populations. The recurrent populations in spring are

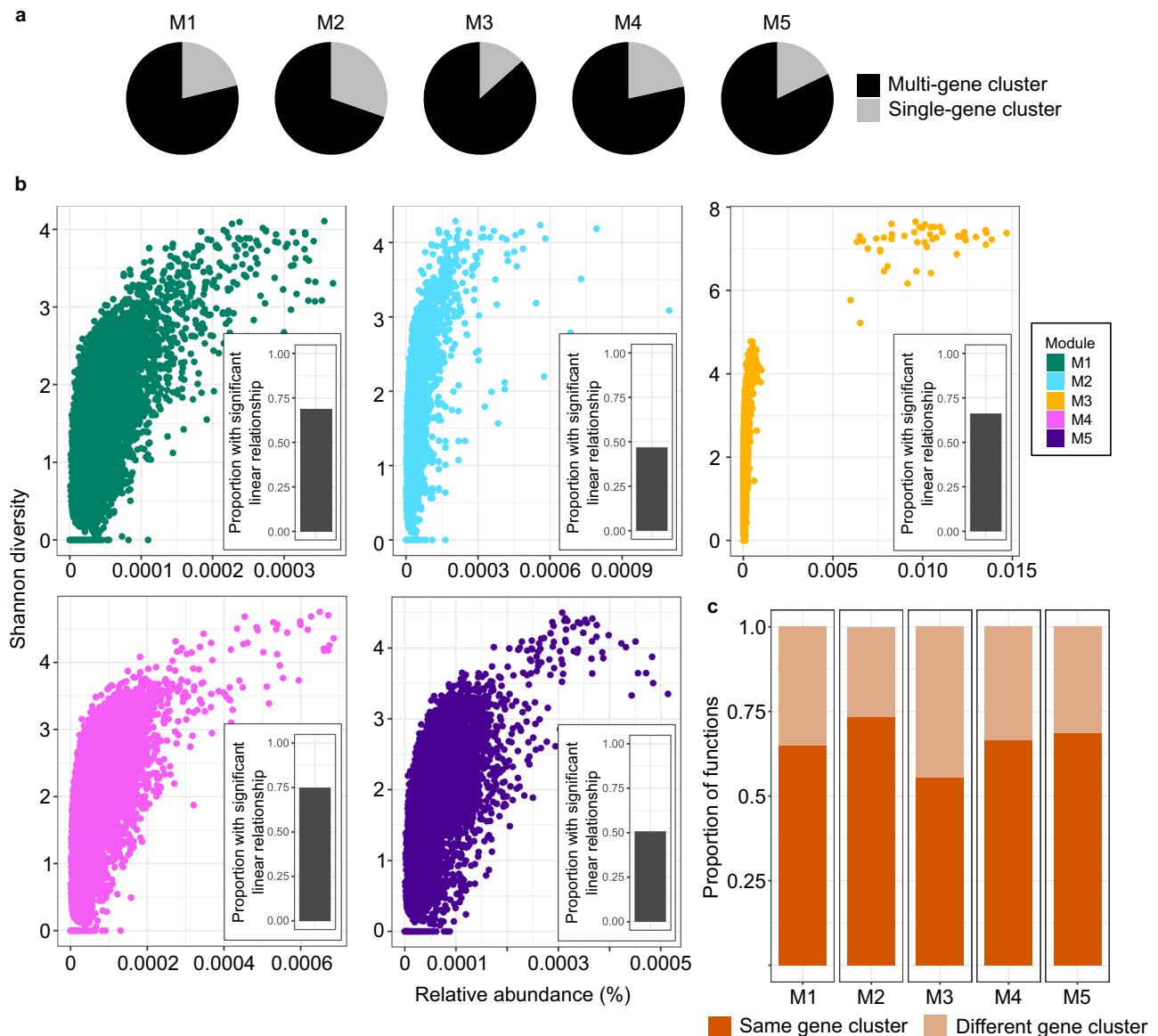


Fig. 7 | Diversity, abundance and structuring of gene clusters within functions. **a** Proportion of functions in each module comprised of multiple or single gene clusters. For those functions comprised of multiple gene clusters, we calculated the **b** Shannon diversity of the gene clusters within each function and compared this to the relative abundance of the function. Each point represents an individual

function. The inserted barplots illustrate the proportion of functions comprised of multiple gene clusters that exhibit a significant positive linear relationship between the gene cluster Shannon diversity and function abundance. **c** Proportion of functions that are dominated by the same or different gene cluster when the function reaches peak abundance each year.

functionally redundant, and the same functions are often dominated by different populations each year. Therefore, selection pressure during spring appears to act primarily on the functional level.

To further advance our understanding of environmental selection and the mechanisms that shape the structuring of ocean microbiomes, it is paramount to integrate both taxonomic and functional profiling along with more extensive environmental information. Of particular importance is information on the availability of organic carbon and energy substrates, which play a fundamental role in shaping the dynamics of microbial communities. For instance, phytoplankton blooms in coastal temperate regions that are of a larger magnitude than in the WSC have been shown to drive a predictable composition of microbes^{7,8}, in contrast to our observations here. The disparity between the regions may be related to the quantity of organic matter released during such phenomena, i.e. higher organic matter concentrations may drive more deterministic taxonomic responses. However, whether the availability of organic matter, or other here

unmeasured environmental factors, plays a role in shaping the deterministic and recurrent dynamics within microbiomes requires further investigation.

Our study provides fundamental insights into the seasonal and inter-annual structuring of prokaryotic and microeukaryotic communities in an Arctic pelagic ocean ecosystem. We demonstrate the prevalence of annually recurrent dynamics of populations and community gene content, which are organised into five distinct seasonal modules. Each of the modules represents a unique ecological state, underpinned by specific prokaryotic and microeukaryotic organisms and functions, and is connected to certain environmental conditions. We further provide evidence that environmental selection is heterogeneous across these ecological states, with differential pressures on an organismal and functional level. Our findings provide fundamental insights into understudied yet rapidly changing ocean regions and advance our understanding of how microbial communities are structured across pronounced environmental gradients.

Methods

Sample collection and processing

Moorings carrying autonomous water samplers (Remote Access Samplers; RAS) were deployed between 2016 and 2020 at a single location in the eastern Fram Strait (mooring F4: 79.0118N 6.9648E). Moorings were deployed for 12-month intervals, with collection and redeployment occurring in summer, typically in August. Owing to ocean currents, the vertical positioning of the RAS fluctuated between 20 and 110 m over the 4-year period. At weekly to fortnightly intervals, 2 × 500 ml of seawater was collected in sterile plastic bags and fixed with mercuric chloride (0.01% final concentration). Following mooring recovery, fixed seawater samples from each time point were filtered onto 0.22 µm Sterivex cartridges and directly frozen at −20 °C until DNA extraction.

Mooring and satellite data

Seabird SBE37-ODO CTD sensors measured temperature, depth, salinity and oxygen, as detailed in a preceding study²³. Employing multiple CTD sensors along the mooring enabled the determination of the minimum mixed layer depth (MLD) at each sampling time point. For instance, if two CTDs showed the same temperature and salinity measurements, the MLD was at least the depth of the deeper CTD. Chlorophyll concentrations were measured via Wetlab Ecotriplet sensors. Surface water photosynthetically active radiation (PAR) data, with a 4 km grid resolution, was obtained from AQUA-MODIS (Level-3 mapped; SeaWiFS, NASA) and extracted in QGIS v3.14.16 (<http://www.qgis.org>).

SSU rRNA gene amplicon and metagenome sequencing

Filtered seawater samples from 97 time points were subjected to DNA extraction using the DNeasy PowerWater Kit (QIAGEN, Hilden, Germany). 16S and 18S rRNA gene fragments were PCR-amplified using the primers 515F–926R⁵² and 528iF–964iR⁵³, respectively. Sequencing libraries were constructed from rRNA gene products according to the “16S Metagenomic Sequencing Library Preparation” protocol (Illumina, San Diego, CA) and sequenced on an Illumina MiSeq platform in 2 × 300 bp, paired-end mode. Amplicon sequencing took place at the Alfred Wegener Institute. The extracted DNA from 47 time points was additionally used to generate PacBio HiFi metagenomes. Sequencing libraries were prepared following the protocol “Procedure & Checklist – Preparing HiFi SMRTbell Libraries from Ultra-Low DNA Input” (PacBio, Menlo Park, CA), followed by inspection with a FEMTOpulse. The libraries were multiplexed and sequenced on 8 M SMRT cells (7–8 samples per cell) on a PacBio Sequel II platform for 30 h with sequencing chemistry 2.0 and binding kit 2.0. Metagenomes were sequenced at the Max Planck Genome Centre, Cologne, Germany.

Recovery of metagenome-assembled genomes

PacBio HiFi reads from each metagenome were assembled using metaFlye v2.9.1⁵⁴ (parameters: --meta --pacbio-hifi --hifi-error 0.01 --keep-haplotypes). The depth of coverage of contigs across all 47 metagenomes was determined through read recruitment using minimap2 v2.1⁵⁵ (parameters: -ax map_hifi --sam-hit-only) followed by the jgi_summarize_bam_contig_depths script from MetaBat2 v2.15⁵⁶. Contigs from each metagenome were binned using MetaBat2 with default settings. The completeness and contamination of bins was determined using the lineage_wf workflow of CheckM v1.2.0 and those with a completeness <50% and contamination >10% were removed. The quality-filtered bins from the 47 metagenomes were dereplicated into species-level clusters based on a 95% average nucleotide identity using the dereplicate function of dRep v3.2.2⁵⁷ (parameters: -sa 0.95 -nc 0.50 -comp 50 -con 10). The taxonomic classification of species-representative metagenome-assembled genomes (MAGs) was performed against the GTDB r220 database using the classify_wf workflow of gtdbtk v2.3.2⁵⁸.

The abundance of species-representative MAGs was determined based on the average depth of coverage of four ribosomal proteins (large subunit 3, large subunit 4, large subunit 6 and small subunit 8). The average depth of the four ribosomal protein genes was used to estimate the number of genomes sequenced in each metagenome. The same sequences, extracted from the raw PacBio HiFi reads, were competitively aligned against the species-representative MAGs using minimap2 v2.1⁵⁵ (parameters: -ax map_hifi --sam-hit-only). For each MAG, we calculated the average depth of aligned ribosomal protein gene sequences and took the quotient between this value and the average depth on the metagenomic read level, resulting in values that represent the proportions of genomes sequenced.

Gene and functional analysis

A total of 48 Gbp of PacBio HiFi reads were generated, with an average of 1 Gbp per sample. Gene sequences were predicted on HiFi reads using FragGenescan (v1.31; parameters: -complete=1 -train=sanger_5)⁵⁹. Reads and their contained gene sequences were taxonomically classified at the domain-level using Tiara v1.0.3, followed by a diamond blastp and last-common ancestor approach for those identified as *Archaea* and *Bacteria*, as described previously⁶⁰. In brief, a Diamond v2.10⁶¹ blastp search (minimum coverage of 40% and identity of 60%) was performed against a database containing gene sequences predicted from the GTDB species-representatives. The last-common ancestor approach was applied, using the tool taxonkit⁶², to provide a conservative taxonomic classification at the gene level and then repeated at the read level. A gene cluster catalogue was constructed through clustering of gene nucleotide sequences at a 95% identity threshold with the *easy-cluster* workflow from MMSeqs2 v15.6⁶³ (parameters: --min-seq-id 0.95 -c 0.6 --cluster-reassign). To facilitate comparisons between metagenomes, gene cluster counts were normalised by the estimated number of prokaryotic genomes in each sample, determined from the average sequencing depth of 16 single-copy ribosomal proteins, as described previously⁶⁰, before converting to relative proportions. The longest sequence from each cluster was used as the representative for a functional assignment against the EGGNOG v5.0 database⁶⁴ using the eggno-mapper tool v2.1.12⁶⁵ with a minimum coverage of 60% and identity of 40% (parameters: --query_cover 60 --pid 40 --sensmode more-sensitive --pfam_realign). To build the ‘functional clusters’ used in our analysis, we grouped gene clusters based on functional annotations of matching seed orthologs. As the seed orthologs of EGGNOG have been functionally annotated using numerous sources, we first grouped the gene clusters that had an annotation from the carbohydrate-active enzyme database, followed by KEGG and then PFAM. The abundance of functional clusters was determined from the sum of the contained gene clusters.

Taxonomic diversity analyses

The 16S and 18S rRNA gene sequences were processed into Amplicon Sequence Variants (ASV) using DADA2⁶⁶ in R v4.1.3⁶⁷. ASVs were taxonomically classified using the SILVA SSU v138 (16S) and PR2 v4.12 (18S) databases using the assignTaxonomy function of DADA2. We only considered ASVs with >3 counts in >3 samples. For alpha diversity, the ASV count table was subject to 100 iterations of rarefying followed by the calculation of Richness, Shannon diversity (*diversity* function in *vegan*) and Evenness. The mean and standard deviation was calculated for each metric, with the mean being used for visualisations and statistical comparisons to environmental metadata. For beta diversity analysis, the ASV count table was rarefied a single time, followed by Hellinger transformation before the calculation of Bray-Curtis dissimilarities (using *vegan*). Dissimilarities were ordinated using non-metric multi-dimensional scaling.

Time series and network analysis

The temporal analysis of ASVs and gene clusters, as well as the construction of co-occurrence networks, was performed using the same workflow as that presented recently by authors of this paper⁶⁸. In brief, for each ASV and gene cluster, we employed a Fourier transformation approach to calculate a time-series signal from relative abundances using the *segmenTier/segmenTools* packages⁶⁹. The Fourier Transform is a technique for decomposing functions or signals into the sum of their frequency components, characterised by sine and cosine functions. In practice, this involves converting temporal variations in relative abundances into frequencies, whereby each frequency is comprised of two parts, the real components and the imaginary components. The imaginary components are derived from the calculation of the sine and cosine trajectories. Together, these two components describe the amplitude and phase of the oscillation of the frequency. To determine the oscillation signals, i.e. the number of peaks in the frequencies during each annual cycle, we reconstruct the time series through a limited number of significant Fourier components. For this, we retained only the first few Fourier components, which capture the frequencies that contribute the most to the signal. By focusing on these significant components, it minimises noise and minor fluctuations that do not inform about major temporal dynamics. Through this, we obtain the number of oscillations in the frequencies within each annual cycle, which we term oscillation signals. A single oscillation would refer to a single peak in abundance of an ASV or gene cluster within a single annual cycle. The oscillation signals were extracted for each ASV and gene cluster using the *fprocessTimeseries* function from the *segmenTier* package. The ASVs and genes with an oscillation signal of one, i.e. a single peak and trough in each annual cycle, were deemed as annually oscillating and retained for further analysis. Phase-Rectified Signal Averaging (PRSA) was used to visualise periodic patterns of ASVs, using phase-rectified data to remove phase variability. Then, the phase-rectified data was averaged for the final PRSA plot. The oscillation signals of ASVs and gene clusters were compared through pairwise Pearson's correlation, with multiple testing corrections using the FDR method. Those with a statistically significant ($p < 0.05$) positive correlation coefficient of >0.7 were used to build a co-occurrence network, with edges as correlation coefficients. By using oscillation signals, focusing on ASVs and gene clusters with a defined oscillation, and using only positive correlations, we minimise the noise in the dataset and prevent potential network topology distortion of negative correlations. Co-occurrence correlation networks were constructed using the *igraph* package⁷⁰ in R and visualised in Cytoscape v3.7.2⁷¹ using the edge-weighted spring-embedded layout. ASVs and gene clusters in the network were clustered using the Louvain algorithm⁷². All steps outlined above were performed in R v4.1.3.

Statistical analysis of microbiome modules

To compare the diversity of module components while accounting for variations in sequencing depth, we iteratively calculated the diversity of ASVs and gene clusters in each module after subsampling. That is, each iteration involved subsampling the ASV and gene cluster tables to the lowest count and then calculating the diversity of components that were retained in each module. To ensure that the diversity comparisons were not influenced by the frequency of metagenome samples over the time series, we repeated the same subsampling and diversity calculation while only considering one sample per calendar month. The dynamics of ASVs and gene clusters in each module were compared with environmental conditions through Pearson's correlations using the *rcorr* package in R.

Assessment of module function transcription across the wider Arctic Ocean

To assess the dynamics and activity of module functions across the wider Arctic Ocean, we leveraged a read recruitment approach using

metatranscriptomes generated from the *Tara Oceans* Polar Circle expedition. Metatranscriptomic reads were mapped to gene cluster representatives using BBMAP v39.06⁷³ using a minimum identity threshold of 95%. The coverage of gene clusters in each metatranscriptome was determined from mapped read counts using the 'genomcov' function of the Bedtools v2.31.1 programme⁷⁴. Coverage values were further processed to determine the 80% truncated average depth (TAD80) value for each gene cluster in each metatranscriptome using the BedGraph.tad.rb script from the Enveomics collection⁷⁵.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Mooring data are available at PANGAEA under 904565 (2016–2017), 904534 (2017–2018), 941126 (2018–2019), and 946508 (2019–2020). 16S rRNA amplicon reads are available at ENA under PRJEB43890 (2016–2017), PRJEB43889 (2017–2018), PRJEB67813 (2018–2019), PRJEB66202 (2019–2020). 18S rRNA amplicon reads are available under PRJEB43504 (2016–2017), PRJEB43885 (2017–2018), PRJEB66212 (2018–2019), PRJEB66220 (2019–2020). Raw metagenomic reads, assemblies and metagenome-assembled genomes are available under PRJEB67368. Accessions for each amplicon and metagenome sample are listed in Supplementary Data 2. Source Data necessary to reproduce the analysis and figures presented in this manuscript are available at <https://zenodo.org/records/14586634>.

Code availability

The code for reproducing the analysis and recreating the figures presented in this manuscript are available at https://github.com/tpriest0/Fram_Strait_WSC_time_series_2016-2020.

References

- Falkowski, P. G., Barber, R. T. & Smetacek, V. Biogeochemical controls and feedbacks on ocean primary production. *Science* **281**, 200–206 (1998).
- Azam, F. et al. The ecological role of water-column microbes in the sea. *Mar. Ecol. Prog. Ser.* **10**, 257–263 (1983).
- Behrenfeld, M. J. & Boss, E. S. Resurrecting the ecological underpinnings of ocean plankton blooms. *Annu. Rev. Mar. Sci.* **6**, 167–194 (2014).
- Fuhrman, J. A., Cram, J. A. & Needham, D. M. Marine microbial community dynamics and their ecological interpretation. *Nat. Rev. Microbiol.* **13**, 133–146 (2015).
- Rodriguez-Brito, B. et al. Viral and microbial community dynamics in four aquatic environments. *ISME J.* **4**, 739–751 (2010).
- Eiler, A., Hayakawa, D. & Rappé, M. Non-random assembly of bacterioplankton communities in the subtropical North Pacific Ocean. *Front. Microbiol.* **2**, 140 (2011).
- Teeling, H. et al. Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *eLife* **5**, e11888 (2016).
- Chafee, M. et al. Recurrent patterns of microdiversity in a temperate coastal marine environment. *ISME J.* **12**, 237–252 (2018).
- Cram, J. A. et al. Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *ISME J.* **9**, 563–580 (2015).
- Gilbert, J. A. et al. Defining seasonal marine microbial community dynamics. *ISME J.* **6**, 298–308 (2012).
- Käse, L. et al. Rapid succession drives spring community dynamics of small protists at Helgoland Roads, North Sea. *J. Plankton Res.* **42**, 305–319 (2020).
- Fuhrman, J. A. et al. Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc. Natl Acad. Sci. USA* **103**, 13104–13109 (2006).

13. Yeh, Y.-C. & Fuhrman, J. A. Contrasting diversity patterns of prokaryotes and protists over time and depth at the San-Pedro Ocean time series. *ISME Commun.* **2**, 1–12 (2022).
14. Needham, D. M. & Fuhrman, J. A. Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nat. Microbiol.* **1**, 16005 (2016).
15. Martin-Platero, A. M. et al. High resolution time series reveals cohesive but short-lived communities in coastal plankton. *Nat. Commun.* **9**, 266 (2018).
16. Berdjeb, L., Parada, A., Needham, D. M. & Fuhrman, J. A. Short-term dynamics and interactions of marine protist communities during the spring–summer transition. *ISME J.* **12**, 1907–1917 (2018).
17. Martiny, J. B. H., Jones, S. E., Lennon, J. T. & Martiny, A. C. Microbiomes in light of traits: a phylogenetic perspective. *Science* **350**, aac9323 (2015).
18. Louca, S., Parfrey, L. W. & Doebeli, M. Decoupling function and taxonomy in the global ocean microbiome. *Science* **353**, 1272–1277 (2016).
19. Polyakov, I. V. et al. Greater role for Atlantic inflows on sea-ice loss in the Eurasian Basin of the Arctic Ocean. *Science* **356**, 285–291 (2017).
20. Berge, J. et al. In the dark: a review of ecosystem processes during the Arctic polar night. *Prog. Oceanogr.* **139**, 258–271 (2015).
21. Ghiglione, J. F. & Murray, A. E. Pronounced summer to winter differences and higher wintertime richness in coastal Antarctic marine bacterioplankton. *Environ. Microbiol.* **14**, 617–629 (2012).
22. Raes, E. J. et al. Seasonal patterns of microbial diversity across the world oceans. *Limnol. Oceanogr. Lett.* **9**, 512–523 (2024).
23. Wietz, M. et al. The polar night shift: seasonal dynamics and drivers of Arctic Ocean microbiomes revealed by autonomous sampling. *ISME Commun.* **1**, 76 (2021).
24. Whitney, F. A. & Freeland, H. J. Variability in upper-ocean water properties in the NE Pacific Ocean. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **46**, 2351–2370 (1999).
25. Sunagawa, S. et al. Structure and function of the global ocean microbiome. *Science* **348**, 1261359 (2015).
26. Ibarbalz, F. M. et al. Global trends in marine plankton diversity across kingdoms of life. *Cell* **179**, 1084–1097.e21 (2019).
27. Fuhrman, J. A. et al. A latitudinal diversity gradient in planktonic marine bacteria. *Proc. Natl Acad. Sci. USA* **105**, 7774–7778 (2008).
28. Hernando-Morales, V., Ameneiro, J. & Teira, E. Water mass mixing shapes bacterial biogeography in a highly hydrodynamic region of the Southern Ocean. *Environ. Microbiol.* **19**, 1017–1029 (2017).
29. Zhang, Y., Xu, H., Qiao, F. & Dong, C. Seasonal variation of the global mixed layer depth: comparison between Argo data and FIO-ESM. *Front. Earth Sci.* **12**, 24–36 (2018).
30. von Appen, W.-J. et al. Sea-ice derived meltwater stratification slows the biological carbon pump: results from continuous observations. *Nat. Commun.* **12**, 7309 (2021).
31. Lambert, S., Lozano, J.-C., Bouget, F.-Y. & Galand, P. E. Seasonal marine microorganisms change neighbours under contrasting environmental conditions. *Environ. Microbiol.* **23**, 2592–2604 (2021).
32. Hatosy, S. M. et al. Beta diversity of marine bacteria depends on temporal scale. *Ecology* **94**, 1898–1904 (2013).
33. Kubiszyn, A. M., Piwosz, K., Wiktor, J. M. Jr. & Wiktor, J. M. The effect of inter-annual Atlantic water inflow variability on the planktonic protist community structure in the West Spitsbergen waters during the summer. *J. Plankton Res.* **36**, 1190–1203 (2014).
34. Nöthig, E.-M. et al. Summertime plankton ecology in Fram Strait—a compilation of long- and short-term observations. *Polar Res.* **34**, 23349 (2015).
35. Sarmento, H., Morana, C. & Gasol, J. M. Bacterioplankton niche partitioning in the use of phytoplankton-derived dissolved organic carbon: quantity is more important than quality. *ISME J.* **10**, 2582–2592 (2016).
36. Seymour, J. R., Amin, S. A., Raina, J.-B. & Stocker, R. Zooming in on the phycosphere: the ecological interface for phytoplankton–bacteria relationships. *Nat. Microbiol.* **2**, 17065 (2017).
37. Choi, D. H. et al. Daily variation in the prokaryotic community during a spring bloom in shelf waters of the East China Sea. *FEMS Microbiol. Ecol.* **94**, fiy134 (2018).
38. Lima-Mendez, G. et al. Determinants of community structure in the global plankton interactome. *Science* **348**, 1262073 (2015).
39. Könneke, M. et al. Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO₂ fixation. *Proc. Natl Acad. Sci. USA* **111**, 8239–8244 (2014).
40. Priest, T. et al. Atlantic water influx and sea-ice cover drive taxonomic and functional shifts in Arctic marine bacterial communities. *ISME J.* **17**, 1612–1625 (2023).
41. Tolar, B. B. et al. Contribution of ammonia oxidation to chemoautotrophy in Antarctic coastal waters. *ISME J.* **10**, 2605–2619 (2016).
42. Liu, Y., Blain, S., Crispi, O., Rembauville, M. & Obernosterer, I. Seasonal dynamics of prokaryotes and their associations with diatoms in the Southern Ocean as revealed by an autonomous sampler. *Environ. Microbiol.* **22**, 3968–3984 (2020).
43. O'Brien, J. et al. The microbiological drivers of temporally dynamic dimethylsulfoniopropionate cycling processes in Australian coastal shelf waters. *Front. Microbiol.* **13**, 894026 (2022).
44. González, J. M., Hernández, L., Manzano, I. & Pedrós-Alió, C. Functional annotation of orthologs in metagenomes: a case study of genes for the transformation of oceanic dimethylsulfoniopropionate. *ISME J.* **13**, 1183–1197 (2019).
45. Auladell, A. et al. Seasonal niche differentiation among closely related marine bacteria. *ISME J.* **16**, 178–189 (2022).
46. Ward, C. S. et al. Annual community patterns are driven by seasonal switching between closely related marine bacteria. *ISME J.* **11**, 1412–1422 (2017).
47. Aksenov, Y. et al. The Arctic circumpolar boundary current. *J. Geophys. Res. Oceans* **116**, C0917 (2011).
48. Salazar, G. et al. Gene expression changes and community turnover differentially shape the global ocean metatranscriptome. *Cell* **179**, 1068–1083.e21 (2019).
49. Aalto, N. J., Schweitzer, H. D., Krsmanovic, S., Campbell, K. & Bernstein, H. C. Diversity and selection of surface marine microbiomes in the Atlantic-influenced Arctic. *Front. Microbiol.* **13**, 892634 (2022).
50. Turnbaugh, P. J. et al. A core gut microbiome in obese and lean twins. *Nature* **457**, 480–484 (2009).
51. Louca, S. et al. High taxonomic variability despite stable functional structure across microbial communities. *Nat. Ecol. Evol.* **1**, 0015 (2016).
52. Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* **18**, 1403–1414 (2016).
53. Fadeev, E. et al. Microbial communities in the east and west Fram Strait during sea ice melting season. *Front. Mar. Sci.* **5**, 429 (2018).
54. Kolmogorov, M. et al. metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat. Methods* **17**, 1103–1110 (2020).
55. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
56. Kang, D. D. et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**, e7359 (2019).
57. Olm, M. R., Brown, C. T., Brooks, B. & Banfield, J. F. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **11**, 2864–2868 (2017).
58. Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**, 1925–1927 (2020).

59. Rho, M., Tang, H. & Ye, Y. FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res.* **38**, e191 (2010).
60. Priest, T., Vidal-Melgosa, S., Hehemann, J.-H., Amann, R. & Fuchs, B. M. Carbohydrates and carbohydrate degradation gene abundance and transcription in Atlantic waters of the Arctic. *ISME Commun.* **3**, 1–13 (2023).
61. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
62. Shen, W. & Ren, H. TaxonKit: a practical and efficient NCBI taxonomy toolkit. *J. Genet. Genomics* **48**, 844–850 (2021).
63. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
64. Huerta-Cepas, J. et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314 (2019).
65. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* **38**, 5825–5829 (2021).
66. Callahan, B. J. et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
67. RStudio Team. *RStudio: Integrated Development of R* (RStudio Inc., 2015).
68. Oldenburg, E. et al. Sea-ice melt determines seasonal phytoplankton dynamics and delimits the habitat of temperate Atlantic taxa as the Arctic Ocean atlantifies. *ISME Commun.* **4**, ycae027 (2024).
69. Machné, R., Murray, D. B. & Stadler, P. F. Similarity-Based Segmentation of Multi-Dimensional Signals. *Sci. Rep.* **7**, 12355 (2017).
70. Csárdi, G. et al. *graph: Network Analysis and Visualization in R* (2024).
71. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
72. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008 (2008).
73. Bushnell, B. BBTools software package. <http://btools.jgi.doe.gov> (2014).
74. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
75. Rodriguez-R, L. M. & Konstantinidis, K. T. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Prepr.* **4**, e1900v1 (2016).

Acknowledgements

We thank Jana Bäger, Theresa Hargesheimer, Rafael Stiens and Lili Hufnagel for RAS operations; Daniel Scholz for RAS and sensor operations; Normen Lochthofen, Janine Ludszuweit, Lennard Frommhold and Jonas Hagemann for mooring operations; Jakob Barz, Swantje Ziemann and Anja Batzke for DNA extraction, library preparation and sequencing; and Bruno Huettel, Christian Woehle and the technicians at the Max Planck Genome Centre in Cologne for sequencing. The captain, crew and scientists of RV Polarstern cruises PS99.2, PS107, PS114, PS121 and PS126 are gratefully acknowledged. This project has received funding from Polarstern grants AWI_PS99_00, AWI_PS107_05, AWI_PS114_01, AWI_PS121_07, AWI_PS126_05, and AWI_PS126_07. Further support came

from the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007-2013) research project ABYSS (Grant Agreement no. 294757) to AB, from the Helmholtz Association, specifically for the FRAM infrastructure, and from the Max Planck Society.

Author contributions

T.P. designed and executed ASV and metagenomic analyses and wrote the manuscript, with input from M.W., E.O. and B.D. E.O. and O.P. performed time series and network analyses. M.W. processed amplicon raw data into ASVs, co-designed the sampling and mooring strategy, and coordinated data analysis. W.-J.v.A. contributed quality-controlled oceanographic data and coordinated the mooring operations. K.M. coordinated the processing of samples and sequencing and provided 18S rRNA gene sequence data. S.T.-V. provided quality-controlled chlorophyll sensor data. C.B., K.M. and A.B. co-designed the sampling and mooring strategy, and contributed to the interpretation. B.D., B.F. and R.A. contributed to the study design and interpretation. All authors contributed to the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-56203-3>.

Correspondence and requests for materials should be addressed to Taylor Priest or Matthias Wietz.

Peer review information *Nature Communications* thanks Ramiro Logares and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

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

© The Author(s) 2025, corrected publication 2025