

## ARTICLE OPEN



# Atlantic water influx and sea-ice cover drive taxonomic and functional shifts in Arctic marine bacterial communities

Taylor Priest<sup>1</sup>✉, Wilken-Jon von Appen<sup>2</sup>, Ellen Oldenburg<sup>3</sup>, Ovidiu Popa<sup>3</sup>, Sinhué Torres-Valdés<sup>2</sup>, Christina Bienhold<sup>1,4</sup>, Katja Metfies<sup>5</sup>, William Boulton<sup>6,7</sup>, Thomas Mock<sup>6</sup>, Bernhard M. Fuchs<sup>1</sup>, Rudolf Amann<sup>1</sup>, Antje Boetius<sup>1,4,8</sup> and Matthias Wietz<sup>1,4</sup>✉

© The Author(s) 2023

The Arctic Ocean is experiencing unprecedented changes because of climate warming, necessitating detailed analyses on the ecology and dynamics of biological communities to understand current and future ecosystem shifts. Here, we generated a four-year, high-resolution amplicon dataset along with one annual cycle of PacBio HiFi read metagenomes from the East Greenland Current (EGC), and combined this with datasets spanning different spatiotemporal scales (Tara Arctic and MOSAIC) to assess the impact of Atlantic water influx and sea-ice cover on bacterial communities in the Arctic Ocean. Densely ice-covered polar waters harboured a temporally stable, resident microbiome. Atlantic water influx and reduced sea-ice cover resulted in the dominance of seasonally fluctuating populations, resembling a process of “replacement” through advection, mixing and environmental sorting. We identified bacterial signature populations of distinct environmental regimes, including polar night and high-ice cover, and assessed their ecological roles. Dynamics of signature populations were consistent across the wider Arctic; e.g. those associated with dense ice cover and winter in the EGC were abundant in the central Arctic Ocean in winter. Population- and community-level analyses revealed metabolic distinctions between bacteria affiliated with Arctic and Atlantic conditions; the former with increased potential to use bacterial- and terrestrial-derived substrates or inorganic compounds. Our evidence on bacterial dynamics over spatiotemporal scales provides novel insights into Arctic ecology and indicates a progressing Biological Atlantification of the warming Arctic Ocean, with consequences for food webs and biogeochemical cycles.

*The ISME Journal* (2023) 17:1612–1625; <https://doi.org/10.1038/s41396-023-01461-6>

## INTRODUCTION

The Arctic Ocean is experiencing unprecedented changes as a result of climate warming, progressing nearly four times faster than the global average [1]. Of particular significance is the rapid decline in sea-ice extent and thickness [2, 3], with future projections indicating frequent ice-free summers by 2050 [4]. In the Eurasian Arctic, accelerated rates of sea-ice decline are associated with increasing volume and heat content of inflowing Atlantic water (AW) [5]. The expanding influence of AW in the Arctic Ocean, termed Atlantification, not only impacts hydrographic and physicochemical conditions, but also provides avenues for habitat range expansion of temperate organisms [6, 7].

The impact of climate change on biological communities has become increasingly apparent across the Arctic Ocean in recent decades. Elevated primary production in shelf seas has been attributed to declining sea-ice extent and increasing phytoplankton biomass [8], particularly in the Eurasian Arctic where Atlantification is driving a poleward expansion of temperate

phytoplankton [7, 9]. Concurrently, phytoplankton phenologies are also changing, with secondary autumnal blooms now occurring in seasonally ice-covered areas [10]. This will have major consequences for the organic matter pool of the Arctic Ocean. Sea-ice dynamics play an important role in the availability of nutrients and organic matter in surface waters and the transport of carbon to the deep-sea [11–13]. At sea-ice margins, strong melt events result in intense stratification, which traps organic material in surface waters and delays vertical export [11].

Considering their role as primary degraders of organic matter and mediators of biogeochemical cycles, assessing the consequences of such changes for bacterial communities is essential to understand and predict alterations to ecosystem functioning. Recent studies have documented distinctions in bacterial communities between Atlantic- and Arctic-derived waters [14], and between sea-ice and seawater [15]. In addition, sea ice-derived dissolved organic matter (DOM) has been shown to stimulate rapid responses by bacterial taxa and significantly alter communities in incubation experiments [16, 17]. However, in order to gain

<sup>1</sup>Max Planck Institute for Marine Microbiology, Bremen 28359, Germany. <sup>2</sup>Physical Oceanography of the Polar Seas, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven 27570, Germany. <sup>3</sup>Institute for Quantitative and Theoretical Biology, Heinrich Heine University Düsseldorf, Düsseldorf 40225, Germany. <sup>4</sup>Deep-Sea Ecology and Technology, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven 27570, Germany. <sup>5</sup>Polar Biological Oceanography, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven 27570, Germany. <sup>6</sup>School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom. <sup>7</sup>School of Computing Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom. <sup>8</sup>MARUM – Center for Marine Environmental Sciences, University of Bremen, Bremen 28359, Germany. ✉email: [tpriest@mpi-bremen.de](mailto:tpriest@mpi-bremen.de); [matthias.wietz@awi.de](mailto:matthias.wietz@awi.de)

Received: 19 September 2022 Revised: 6 June 2023 Accepted: 15 June 2023

Published online: 8 July 2023

a deeper understanding of potential shifts in Arctic Ocean microbial ecology, communities need to be studied over high-resolution temporal scales and across natural environmental gradients such as the Arctic–Atlantic interface.

The Fram Strait, the main deep-water gateway between the Arctic and Atlantic Oceans, is a key location for conducting long-term ecological research over environmental gradients and under changing conditions [18]. Fram Strait harbours two major current systems; the East Greenland Current (EGC), transporting polar water (PW) southwards, and the West Spitsbergen Current (WSC), transporting AW northward. The EGC accounts for the export of ~50% of freshwater and ~90% of sea-ice from the central Arctic Ocean and carries Arctic hydrographic signatures [19]. Large-scale recirculation of AW into the EGC continuously occurs, although the magnitude varies across latitudes and over time [20, 21]. The mixing of AW and PW in the marginal ice zone (MIZ) creates different hydrographic regimes reflective of Arctic, mixed and Atlantic conditions, which can harbour unique bacterial compositions [14, 22]. It has been predicted that future Atlantification of the Arctic may result in a shift towards temperate, Atlantic-type communities [14]. However, further assessments of microbial population dynamics across spatiotemporal scales are needed to validate such hypotheses.

Here, we performed a high-resolution analysis of the temporal variation of bacterial taxonomy and function in the MIZ (2016–2018) and the core-EGC (2018–2020), covering the full spectrum of ice cover, daylight and hydrographic conditions. Our study is embedded in the “Frontiers in Arctic Marine Monitoring” (FRAM) ocean observing framework that employs mooring-attached sensors and autonomous Remote Access Samplers (RAS) to continuously monitor physicochemical parameters and biological communities in the Fram Strait. We analysed four-year 16S rRNA gene amplicon data supplemented with an annual cycle of PacBio HiFi read metagenomes, expanding a previous assessment of microbial dynamics over a single annual cycle in the EGC [23]. We hypothesise that high AW influx and low sea-ice cover result in communities dominated by chemoheterotrophic populations that taxonomically and functionally resemble those of temperate ecosystems. Our study provides essential insights into the impact of changing conditions on microbial ecology and biogeochemical cycles in the Arctic Ocean.

## METHODS

### Seawater collection and processing

Autonomous sample collection and subsequent processing proceeded as previously described [23]. Briefly, RAS (McLane, East Falmouth, MA) were deployed over four consecutive annual cycles between 2016 and 2020, with deployments and recoveries occurring each summer (2019–2020 mooring recovered in 2021). From 2016 to 2018, RAS were deployed in the MIZ (78.83° N –2.79° E) and from 2018 to 2020 in the core-EGC (79° N –5.4° E), with average sampling depths of 80 and 70 m, respectively. The depths were chosen to prevent contact with moving ice overhead. In weekly to fortnightly intervals (Supplementary Table S1), ~1 L of seawater was pumped into sterile plastic bags and fixed with mercuric chloride (0.01% final concentration). After RAS recovery, water was filtered onto 0.22 µm Sterivex cartridges directly frozen at –20 °C until DNA extraction.

### Amplicon sequencing and analysis

DNA was extracted using the DNeasy PowerWater kit (Qiagen, Germany), followed by amplification of 16S rRNA gene fragments using primers 515F–926R [24]. These primers perform well at recovering marine mock communities, and were recently suggested as optimal for studying Arctic microbial communities [24, 25]. Sequencing was performed on a MiSeq platform (Illumina, San Diego, CA) using 2 × 300 bp paired-end libraries according to the “16S Metagenomic Sequencing Library Preparation protocol” (Illumina). Reads were subsequently processed into amplicon sequence variants (ASVs) using DADA2 and the SILVA v138 database [26–28]. Analysis and plotting were performed in RStudio [29], primarily using the vegan [30], limma [31], mixOmics [32], ggplot2 [33] and

ComplexHeatmap [34] packages. Briefly, community composition was compared using Bray–Curtis dissimilarities and distance-based redundancy analysis (dbRDA) with the functions *decostand* and *dbRDA* in *vegan*, and visualised using *ggplot2*. The influence of environmental variables on community dissimilarity was determined through a stepwise significance test on the dbRDA using the *ordiR2step* and *anova.cca* functions in *vegan*. ASVs were assigned to distribution groups based on the frequency of detection over time.

Co-occurrence networks were calculated for MIZ and core-EGC samples separately using the packages *segmentTier* [35] and *igraph* [36]. Oscillation signals were calculated for each ASV per year based on Fourier transformation of normalised abundances and compared using Pearson's correlations. Only statistically significant positive correlations were retained (adjusted *p*-value < 0.05 after correction using the FDR method [37]). Using a network robustness analysis, a correlation coefficient of 0.7 was determined as a strong co-occurrence. Below this value, removal of a single node would cause network disruption. Networks were constructed using the co-occurrences that passed the above thresholds, and visualised in Cytoscape [38] with the Edge-weighted Spring-Embedded Layout. Values of centrality and node betweenness were calculated using *igraph*.

### PacBio metagenome sequencing

Nine samples from the 2016–2017 annual cycle in the MIZ were selected for metagenomic sequencing, using the same DNA as for amplicon sequencing. Sequencing libraries were prepared following the protocol “Procedure & Checklist – Preparing HiFi SMRTbell Libraries from Ultra-Low DNA Input” (PacBio, Menlo Park, CA) and inspected using a FEMTOpuse. Libraries were sequenced on 8M SMRT cells on a Sequel II platform for 30 h with sequencing chemistry 2.0 and binding kit 2.0. The sequencing was performed together with samples of another project, such that seven samples were multiplexed per SMRT cell. On average, this resulted in 268,000 reads per metagenome, with an N50 of 6.8 kbp.

### Taxonomic and functional annotation of HiFi reads

The 2.4 million generated HiFi reads were processed through a custom taxonomic classification and functional annotation pipeline. The classification pipeline followed similar steps to previously published tools, but with some modifications. A local database was constructed based on protein sequences from all species-representatives in the GTDB r202 database [39]. Prodigal v2.6.3 [40] was used to predict open reading frames (ORFs) on HiFi reads, which were subsequently aligned to the GTDB-based database using Diamond blastp v2.0.14 [41] with the following parameters: `-id 50 -query-cover 60 -top 5 -fast`. After inspection of the hits, a second filtering step was performed: percentage identity of >65% and an *e*-value of <1<sup>-10</sup>. Using Taxonkit v0.10.1 [42], the last common ancestor (LCA) algorithm was performed, resulting in a single taxonomy for each ORF. A secondary LCA was subsequently performed for all ORFs from the same HiFi read, generating a single taxonomy for each read. Functional annotation of HiFi reads was performed using Prokka [43] followed by a series of specialised databases. This included using blastp v2.11.0 [44] or HMMscan (HMMER v3.2.1) [45] against dbCAN v10 [46], CAZy (release 09242021) [47], SulfAtlas v1.3 [48], the Transporter Classification [49], MEROPS [50] and KEGG [51] databases along with sets of Pfam HMM family profiles for SusD and TonB-dependent transporter genes. Functional gene counts were normalised by the average sequencing depth of 16 universal, single-copy ribosomal protein genes per sample [52] – providing “per genome” counts. Genes enriched under high- and low-ice cover conditions were identified using ALDEx2 [53].

### Metagenome-assembled genome recovery

In order to maximise the recovery of metagenome-assembled genomes (MAGs), metagenomes were clustered into two groups based on dissimilarity in ASV composition of the corresponding amplicon samples. Samples were individually assembled using metaFlye v2.8.3 (parameters: `--meta --pacbio-hifi --keep-haplotypes --hifi-error 0.01`). Contigs with a length of <10 kbp were removed and the remaining contigs were renamed to reflect the sample of origin. Contigs from each group were concatenated into a single file. Coverage information, necessary for binning, was acquired through read recruitment of raw reads from all metagenomes to the contigs using Minimap2 v2.1 [54], using the ‘map-hifi’ preset. Contigs were binned using Vamb v3.0.2 [55] in multisplit mode using three different sets of parameters (set1: `-l 32 -n 512 512`, set2: `-l 24 -n 384 384`, set3: `-l 40 -n 768 768`). Completeness and contamination estimates of bins were determined using CheckM v1.1.3 [56], and those with >50% completeness were

manually refined using the interactive interface of Anvi'o v7 [57]. A consensus set of refined MAGs with non-redundant contigs was obtained using DASTool v1.1.1 [58]. The consensus MAGs were de-replicated at 99% average nucleotide identity using dRep v3.2.2 [59] (parameters: -comp 50 -con 5 -nc 0.50 -pa 0.85 -sa 0.98), resulting in 47 population-representative MAGs. A phylogenetic tree was reconstructed that also incorporated MAGs recently published from the Fram Strait [22], following a procedure outlined previously [52]. Briefly, 16 single-copy universal ribosomal protein genes were identified in each MAG using HMMsearch against the individual Pfam HMM family profiles and aligned using Muscle v3.8.15 [60]. Alignments were trimmed using TrimAl v1.4.1 [61], concatenated, and submitted to FastTree v2.1.0 [62]. The tree was visualised and annotated in iTOL [63].

### Classification, abundance and distribution of MAGs

A dual taxonomic classification of MAGs was performed using single-copy marker and 16S rRNA genes. Firstly, MAGs were assigned a taxonomy using the GTDBtk tool v1.7.0 [64] with the GTDB r202 database. Secondly, extracted 16S rRNA gene sequences were imported into ARB [65], aligned with SINA [66] and placed into the SILVA SSU 138 Ref NR99 reference tree using ARB parsimony. Those containing a 16S rRNA gene were linked to ASV sequences through competitive read recruitment using BBMap of the BBtools programme v35.14, with an identity threshold of 100%.

The distribution of MAGs across the Arctic Ocean were determined through recruitment of reads from the herein generated metagenomes and published datasets from the Tara Arctic and MOSAiC expeditions (Supplementary Table S11). Counts of competitively mapped reads were converted into the 80% truncated average sequencing depth, TAD80 [67]. Relative abundance was then determined as the quotient between the TAD80 and the average sequencing depth of 16 single-copy ribosomal protein genes. Ribosomal proteins were identified following the same procedure outlined above, and their sequencing depth estimated using read recruitment with minimap2 (for PacBio-derived metagenomes) and BBMap (for Illumina-derived metagenomes).

### Mooring and satellite data

Bacterial community data was placed into context using in situ measured environmental parameters (Supplementary Table S1). Temperature, depth, salinity and oxygen concentrations were measured using Seabird SBE37-ODO CTD sensors and chlorophyll *a* concentration was measured using a WET Labs ECO Triplet sensor, all attached to the RAS. Sensor measurements were averaged over 4 h around each sampling event. The relative proportions of AW and PW were determined as described previously [23]. Physical sensors were manufacturer-calibrated and processed in accordance with <https://epic.awi.de/id/eprint/43137>. Mooring-derived data are published under PANGAEA accession 904565 [68], 941159 [69], and 946539 [70]. Sea-ice concentrations, derived from the AMSR-2 satellite, were downloaded from <https://seaice.uni-bremen.de/sea-ice-concentration-amrs2>, and averaged over a 15 km radius around the moorings.

## RESULTS

The amplicon dataset incorporates samples (>0.2 µm fraction) collected at weekly to fortnightly intervals in the MIZ (2016–2018) and central EGC (core-EGC; 2018–2020) between 70 and 90 m depth (Supplementary Table S1). The two locations were selected in order to capture the full spectrum of water mass and sea-ice conditions. The core-EGC was characterised by year-round dense ice cover (hereon abbreviated as “high ice”) and PW conditions. In contrast, the MIZ featured variable, generally lower ice cover (hereon abbreviated as “low ice”) and periodic AW influx (Fig. 1). To visually portray this variability, animated GIFs were created for current velocities (Supplementary Fig. S1) and sea-ice cover (Supplementary Fig. S2) over the four-year period. Combining the high-resolution data from both mooring locations allowed for the assessment of bacterial community dynamics over time and in relation to Arctic- and Atlantic-dominated conditions.

### Bacterial community and population dynamics over time

The amplicon dataset encompasses 12.5 million quality-filtered reads in 84 samples, with an average of 134,588 reads per sample. A total of 4083 ASVs (Supplementary Table S2) were recovered,

which were initially used in a taxonomy-independent approach to assess community dynamics over environmental gradients (Fig. 2). A dbRDA with stepwise significance testing identified AW proportion, daylight and past ice cover (average ice cover of the days preceding the sampling event) as the significant factors constraining compositional variation (model  $R^2 = 0.23$ ,  $p = 0.001$ ). AW proportion explained 13% of the variation in bacterial community dissimilarity, compared to 6% for daylight and 4% for past ice cover.

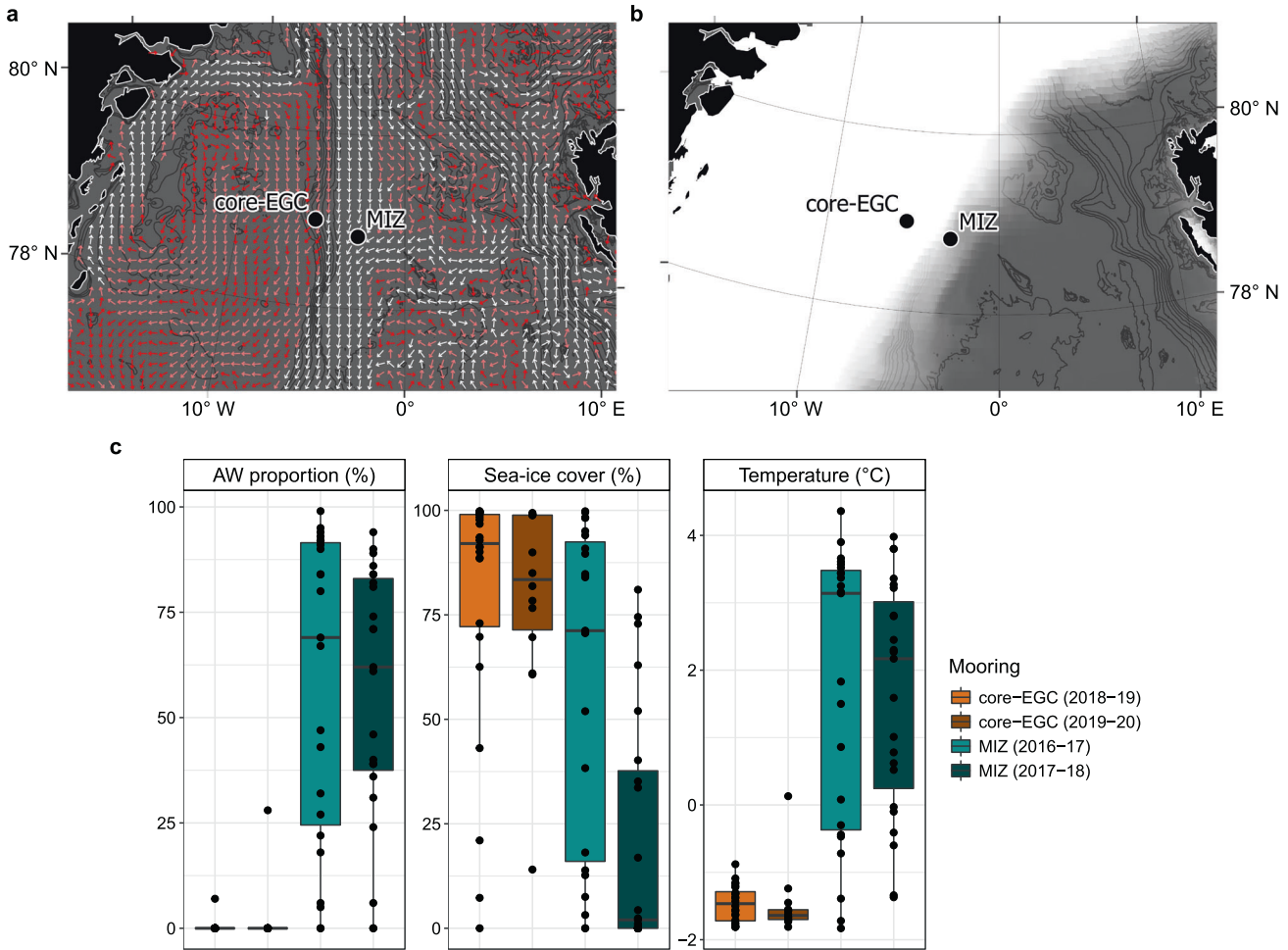
Assessing ASV dynamics at the two mooring locations over time revealed several distinct patterns. In total, 75% of the ASVs were detected at both mooring sites (i.e. shared), whilst 16% and 9% were unique to the MIZ and core-EGC respectively. The frequency of detection and maximum relative abundance of shared ASVs exhibited a strong positive linear relationship, i.e., those identified in more samples also reached higher maximum relative abundances (Fig. 3a). To better understand the structuring of communities and distinguish between ecologically different fractions, we categorised ASVs into three groups: (a) Resident (Res-ASVs), present in >90% of samples, (b) Intermittent (Int-ASVs), present in 25–90% of samples, and (c) Transient (Trans-ASVs), present in <25% of samples (Supplementary Table S3). Res-ASVs represented a small fraction of the diversity (231 ASVs) but the largest proportion of the sampled bacterial communities (43–87% relative abundance). In comparison, the 1943 Int-ASVs constituted 12–53% and the 1909 Trans-ASVs 0.4–9.3% of relative abundances. Presence of a dominant resident microbiome, represented by a minority of ASVs, is consistent with multiannual observations in the Western English Channel and Hawaiian Ocean time-series [71, 72].

Temporal dynamics of the three community fractions was linked to changes in AW proportion, evidenced by negative correlations for the resident (Pearson's coefficient:  $-0.29$ ,  $p < 0.05$ ) and transient fractions (Pearson's coefficient:  $-0.36$ ,  $p < 0.01$ ) compared to positive correlations for the intermittent fraction (Pearson's coefficient:  $0.37$ ,  $p < 0.01$ ). This is reflected in the more stable temporal dynamics at the core-EGC with less AW influence, compared to the MIZ (Fig. 3c). In addition, the transient fraction was positively correlated with ice cover (Pearson's coefficient:  $0.26$ ,  $p < 0.05$ ).

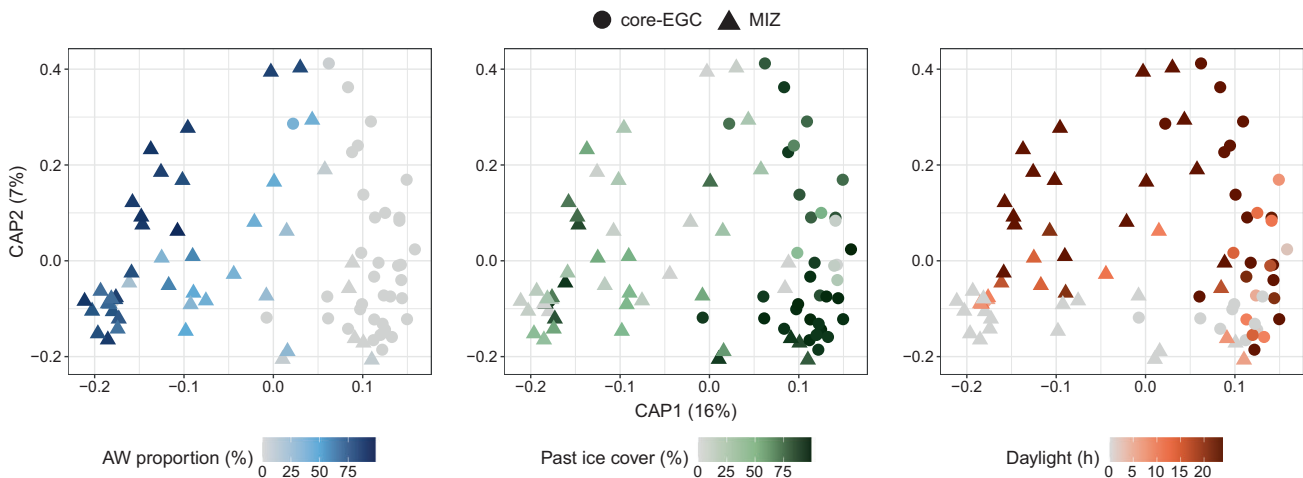
The dynamics of the three community fractions were supported by co-occurrence networks computed at ASV level (Supplementary Fig. S3). The MIZ network contained more ASVs and more significant co-occurrences compared to the core-EGC, primarily driven by Int-ASVs. There were 283 more Int-ASVs in the MIZ than in the core-EGC network, and the number of connections per ASV was nine-fold higher. In contrast, Trans-ASVs were threefold more numerous and exhibited threefold more connections per ASV in the core-EGC compared to the MIZ network (Fig. 3b and Supplementary Information). Res-ASVs were comparable in number in both networks.

The resident microbiome was phylogenetically diverse, incorporating both abundant and rare community members. Res-ASVs were assigned to 61 families and 79 genera, with the *Flavobacteriaceae* ( $n = 15$ ), *Magnetospiraceae* ( $n = 13$ ), *Marinimicrobia* ( $n = 11$ ), SAR11 Clade I ( $n = 21$ ) and SAR11 Clade II ( $n = 17$ ) harbouring the largest diversity. Maximum relative abundances of Res-ASVs ranged from 0.04 to 13.9%, with the most prominent being affiliated with SAR11 Clade Ia (asv1; 14%), *Polaribacter* (asv6; 14%), *Aurantivirga* (asv7; 12%), SUP05 (asv2; 12%), SAR92 (asv16; 11%) and SAR86 (asv3; 9%). Pronounced fluctuations of the intermittent community coincided with AW influx in the MIZ. Int-ASVs were more phylogenetically diverse than Res-ASVs, encompassing 254 genera, and included rare and abundant populations that reached 0.004–36% maximum relative abundance. The most diverse taxa included the SAR11 Clade II ( $n = 148$ ), *Marinimicrobia* ( $n = 129$ ), NS9 Marine Group ( $n = 78$ ), AEGEAN-169 ( $n = 73$ ), and SAR86 ( $n = 47$ ). Those with largest relative abundances were

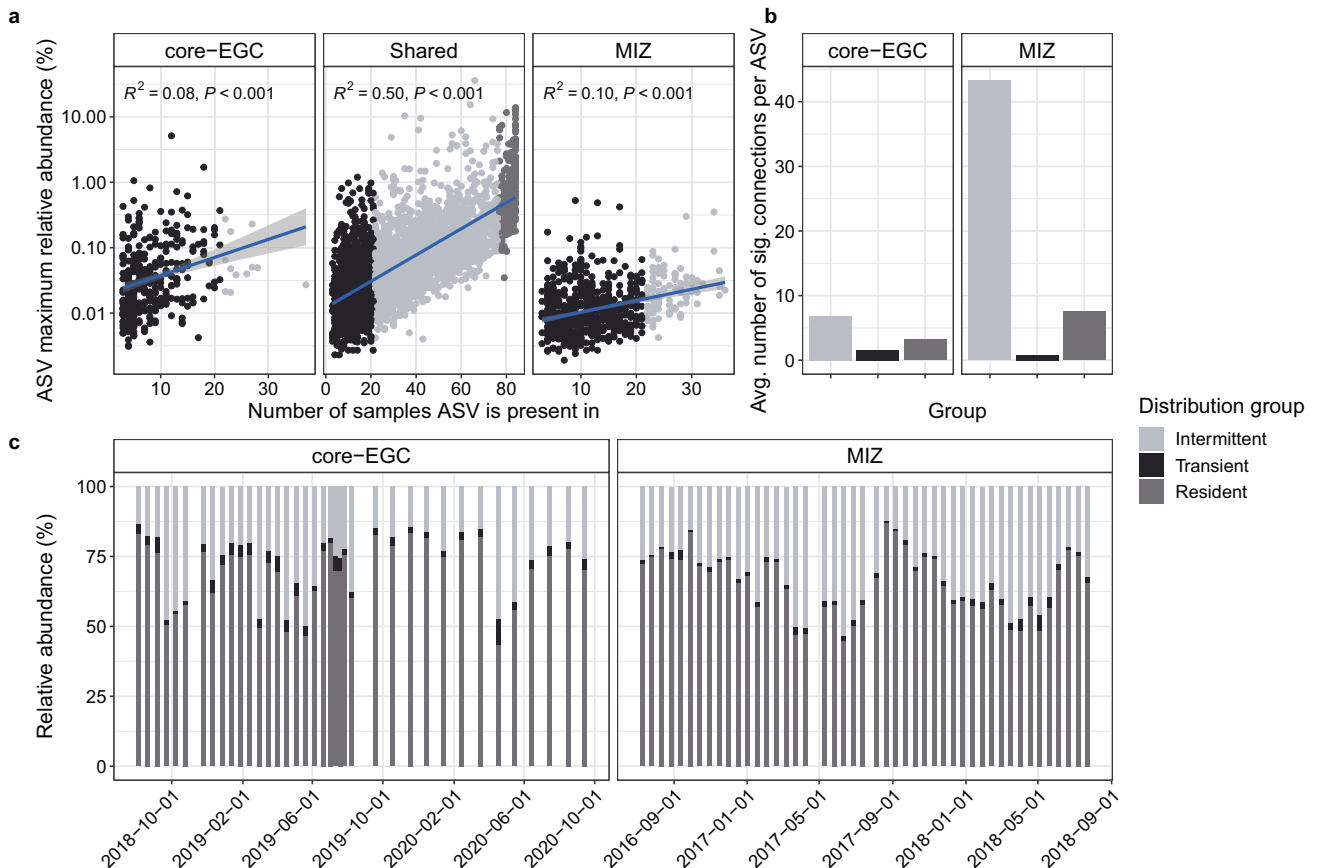




**Fig. 1** Location of seafloor moorings and environmental conditions in the MIZ (2016–2018) and core-EGC (2018–2020). **a** Example representation of monthly average (January 2020) current velocities at the approximate depth of sampling (78 m). White and dark red arrows indicate strongest and weakest velocities, respectively. **b** Example representation (December 2019) of sea-ice cover. Increasing opacity of white colour reflects increasing sea-ice cover (pure white = 100%). Current and sea-ice data were obtained from copernicus.eu under 'ARCTIC\_ANALYSIS\_FORECAST\_PHY\_002\_001\_a'. **c** Variation in AW proportion, ice cover and water temperature at the two moorings. The bathymetric map was made using data from [GEBCO](#).



**Fig. 2** Community structure across water mass, sea-ice and daylight conditions. Distance-based redundancy analysis based on Bray-Curtis dissimilarities of community composition along with AW proportion (blue), past ice cover (green) and daylight (orange) as constraining factors. The factors were selected using a stepwise significance test and combined into a single model ( $R^2 = 0.1$ ,  $p = 0.01$ ) that constrains 14% of the total variation. For ease of interpretation, the environmental conditions are visualised individually on the same ordination.



**Fig. 3 Distribution dynamics and co-occurrence of ASVs.** **a** Occurrence of ASVs across samples in relation to their maximum relative abundances, along with categorisation into resident, intermittent and transient. **b** Average number of connections within the co-occurrence networks for resident, intermittent and transient ASVs. **c** Relative abundance dynamics of resident, intermittent and transient ASVs over time.

affiliated with *Colwellia* (asv10; 36%), *Luteolibacter* (asv24; 15%), *Flavobacterium* (asv140; 10%), and *Polaribacter* (asv206; 10%). The resident and intermittent community fractions shared 71 genera, constituting 90% of the genus-level diversity of the resident microbiome. Hence, compositional changes over temporal scales relate to dynamics on the (sub-)species level.

#### Taxonomic signatures of distinct environmental conditions

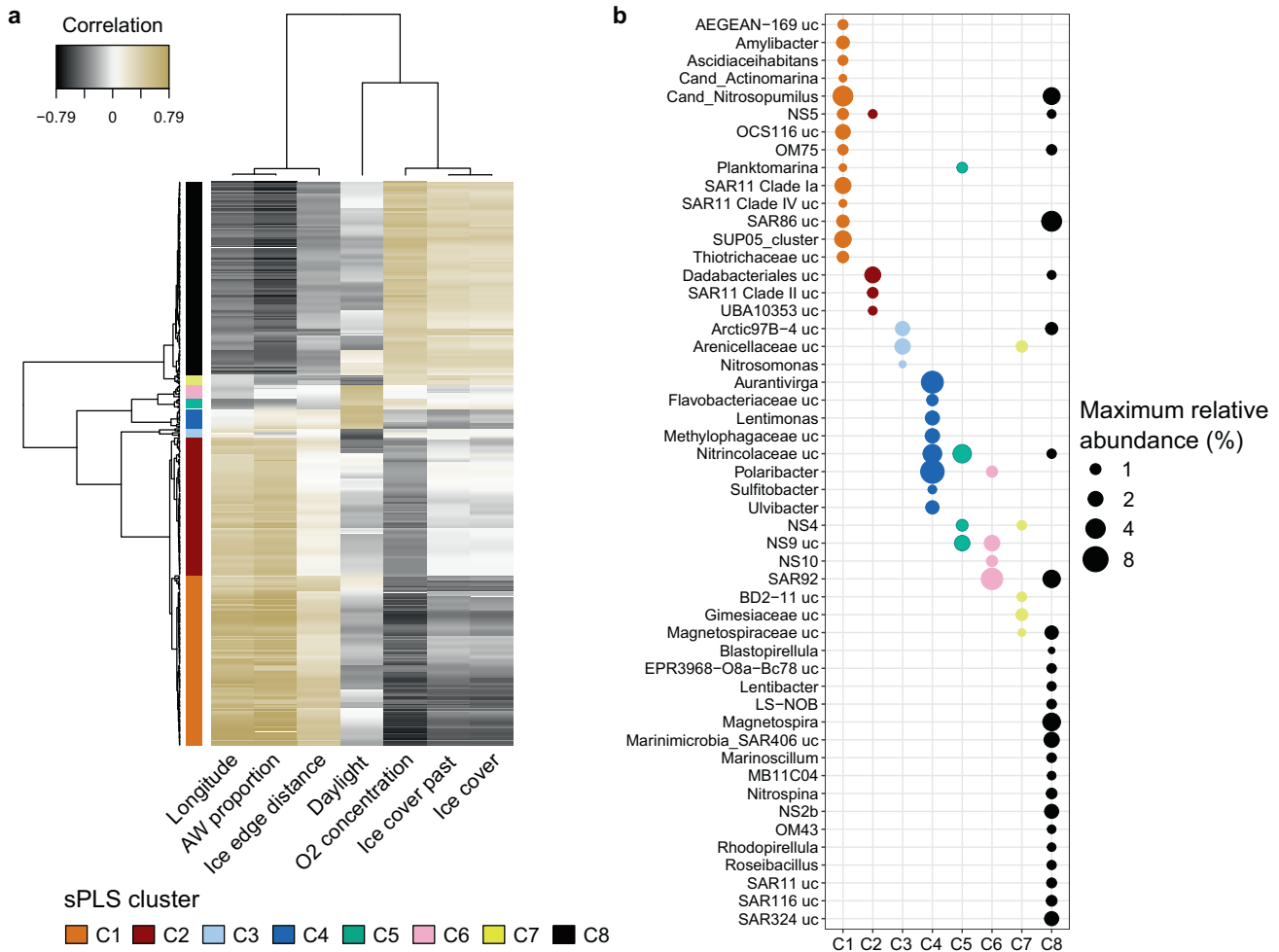
A sparse partial least squares regression analysis (sPLS) identified 430 ASVs that were associated with distinct environmental conditions. Based on similar, significant correlations (Pearson's coefficient  $> 0.4, p < 0.05$ ) to environmental parameters, the ASVs were grouped into eight distinct clusters (Fig. 4a and Supplementary Table S3), each comprising unique taxonomic signatures (Fig. 4b). The three largest clusters encompassed 88% of the ASVs, and were distinguishable based on their associations to different water mass and ice cover conditions. Clusters C1 and C2 represent AW conditions, with C1 also being associated with low-ice cover. In contrast, cluster C8 represents PW conditions under high-ice cover. In accordance with the distribution dynamics described above, the AW-associated clusters comprised a higher proportion of Int-ASVs, 51–88%, compared to ~50% Res-ASVs in PW-associated clusters. Five smaller clusters (C3–C7) correspond to polar day and night under different ice cover and water mass conditions. Comparing the most prominent ASVs ( $>1\%$  relative abundance) of each cluster revealed unique taxonomic signatures at the genus level (Fig. 4b). For instance, *Amylibacter*, SUP05 and AEGEAN-169 are signatures of the AW-associated, low-ice cluster C1, whereas SAR324, NS2b and *Magnetospira* are signatures of the PW-associated, high-ice cluster C8. Overall, this pattern underlines

that water mass and ice cover have the largest influence on microbial community structure, with a smaller number of ASVs being influenced by daylight and seasonality.

#### MAGs and comparison to other Arctic datasets

Nine PacBio HiFi read metagenomes spanning one annual cycle in the MIZ yielded 43 manually refined, population-representative MAGs, delineated at 99% ANI (Supplementary Table S4). The MAGs were of medium- and high-quality according to MIMAG standards [73], exhibited low fragmentation (average number of contigs = 33), and  $>80\%$  contained at least one complete rRNA gene operon. MAGs covered a broad phylogenetic diversity, including 35 genera, 27 families and nine classes (Supplementary Fig. S4). For deeper ecological insights, we contextualised ASV dynamics with MAGs to link distribution with metabolic potential. Of the 27 ASVs linked to a MAG through competitive read recruitment (100% identity threshold), 18 were associated with sPLS clusters and thus distinct environmental conditions – these are hereon referred to as “signature populations” (Supplementary Table S5). Signature populations included some of the most abundant ASVs, such as asv6-*Polaribacter* and asv7-*Aurantivirga* from cluster C4 (polar day-associated) and asv18-SAR86 from cluster C8 (high-ice, PW-associated).

To corroborate the associations of signature populations with distinct environmental conditions, we assessed their spatiotemporal dynamics across the Arctic Ocean. This comparison included an additional 59 metagenomes as well as 1184 MAGs and metagenomic bins from the Fram Strait [22], the Tara Arctic expedition (TARA) [74], and the MOSAiC expedition [75]. Combined, these datasets provide an extensive geographical and seasonal coverage, from above the continental shelf in summer to the central basin in



**Fig. 4** Sparse partial least square regression (sPLS) linking community structure and environmental parameters. **a** Heatmap showing eight major sPLS clusters, encompassing 430 ASVs with significant correlations to environmental conditions. **b** Representation of the most prominent genera per cluster. ASVs with <1% relative abundance were excluded, whilst the remaining were grouped by genus and the maximum abundance of each genus shown. Due to high collinearity with AW proportion, temperature and salinity were excluded. Thresholds: coefficients > 0.4,  $p < 0.05$ .

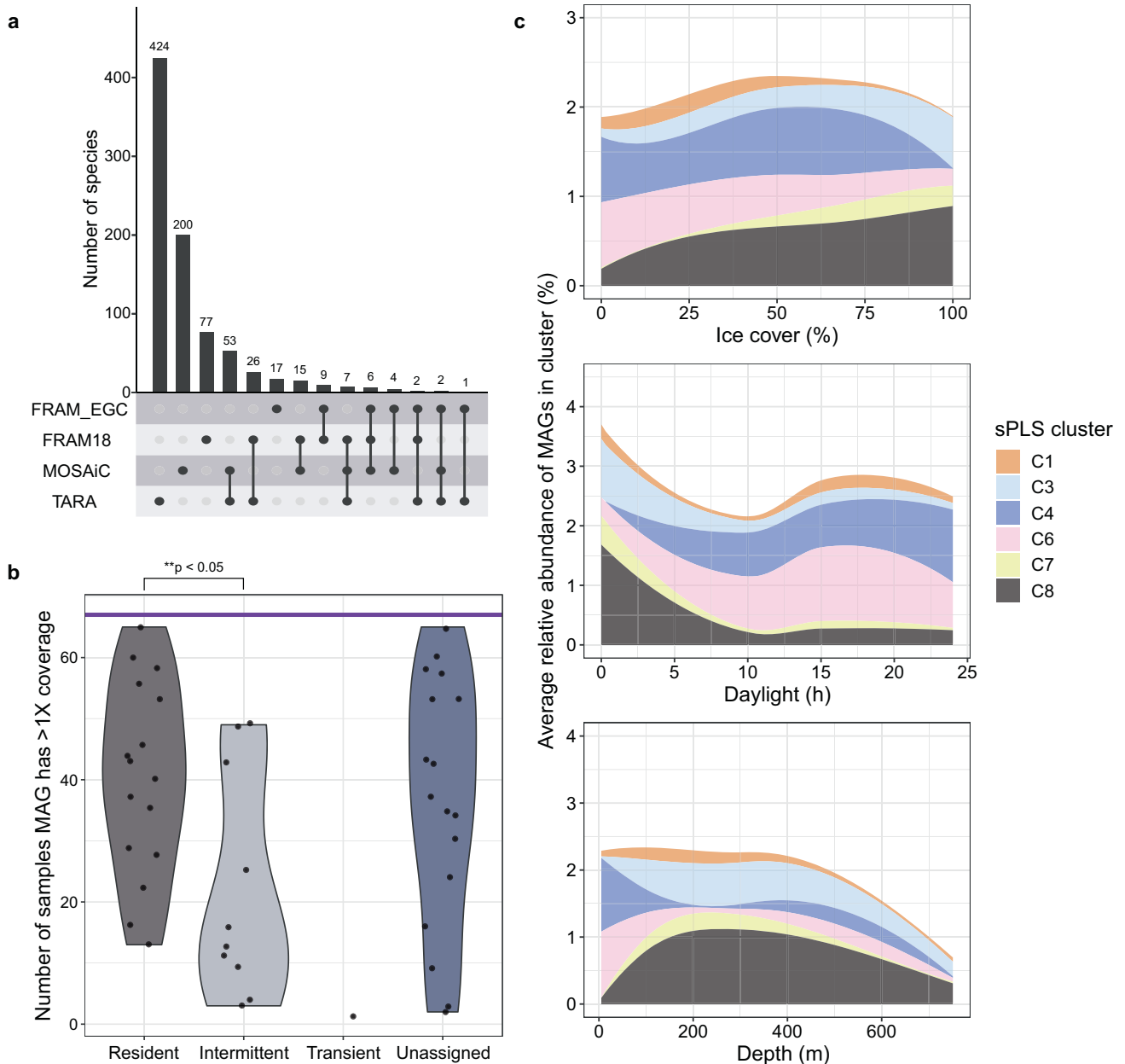
winter. Combining the MAG datasets resulted in 843 species-level clusters at 95% ANI (Supplementary Table S6). Each dataset comprised a mixture of unique and shared species (Fig. 5a), but there were no cosmopolitan species. Of the MAGs recovered in this study, hereon termed FRAM\_EGC MAGs, 42% were unique species. However, these results are influenced by differences in dataset size, sequencing platforms and analysis pipelines, e.g. co-assembly (TARA) vs. single sample assembly (FRAM and MOSAiC).

FRAM\_EGC MAGs were among the most abundant and widely detected (Supplementary Table S7) across the Fram Strait and Arctic Ocean, constituting 0.02–58% of bacterial communities. Their distribution across the wider Arctic supported the dynamics observed in the EGC; e.g. residents (associated with a Res-ASV) were more widely detected than intermittent or transient populations (Fig. 5b, Supplementary Figs. S5 and S6). Three of the resident FRAM\_EGC MAGs, one assigned to OM182 (UBA9659) and two to *Thioglobus*, were detected in >90% of all metagenomes. One of these species did not have a MAG representative in the other Arctic datasets, highlighting that our study contributes novel genomic information towards a better understanding of Arctic Ocean microbial ecology. Furthermore, the dynamics of signature-population MAGs across the Arctic supported their association with distinct environmental conditions. MAGs from cluster C8 (high-ice and PW) and C7 (high-ice

and polar night) reached higher relative abundances in mesopelagic depths (TARA) and during polar night (MOSAiC) (Fig. 5c). In contrast, a higher relative abundance of C4 and C6 (polar day) MAGs occurred in surface water collected during summer (TARA).

### Functional potential of Atlantic and Arctic signature populations

Connecting ASV temporal dynamics and MAG functional potential facilitated predictions on the ecology of signature populations within the context of environmental conditions. Of particular interest were the signature populations of Atlantic (cluster C1) and Arctic (cluster C8) conditions, as they could provide insights into how bacterial community structure and function may shift in the future Arctic Ocean. Comparing the functional potential of MAGs revealed that Atlantic and Arctic signature populations clearly differ in substrate metabolism. In short, signature populations of Arctic conditions harboured genes for autotrophy and the utilisation of bacterial- and/or terrestrial-derived compounds, compared to Atlantic signature populations that were functionally connected to phytoplankton-derived organic substrates. Ecological descriptions of all signature populations and functional gene tables are provided in Supplementary Information and Supplementary Files S1, respectively.

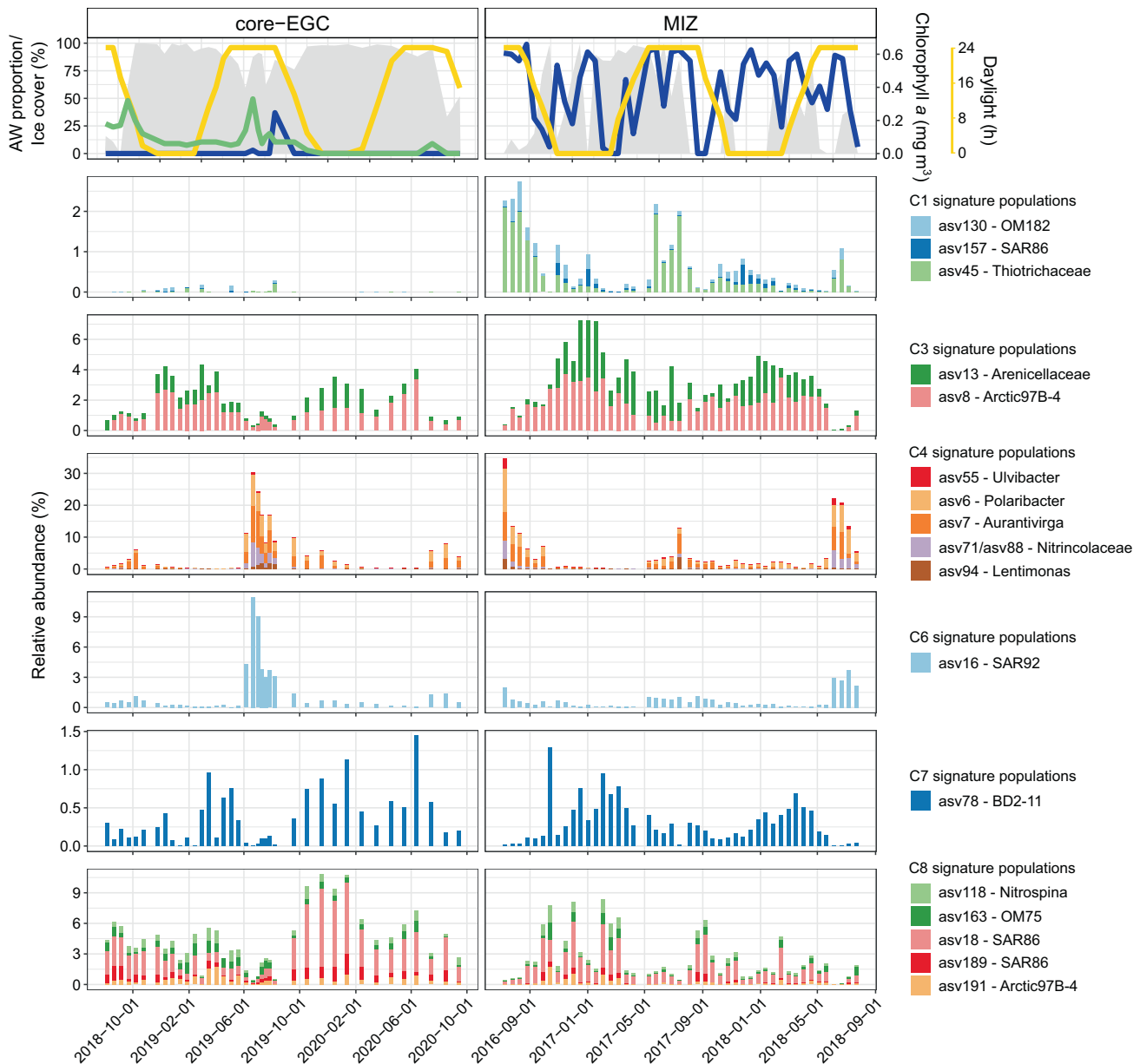


**Fig. 5 Comparison and dynamics of MAGs across the Fram Strait and Arctic Ocean.** We compared metagenome-assembled genomes (MAGs) generated in this study (FRAM\_EGC), and from samples previously collected in the Fram Strait (FRAM18) [23], in the Arctic Ocean during summer (TARA) [34], and in the Arctic Ocean during winter (MOSAIC) [35]. **a** Number of shared and unique species across the four MAG datasets, determined by comparisons at 95% average nucleotide identity threshold. **b** Number of metagenomes in which FRAM\_EGC MAGs were detected with at least 1× coverage. The horizontal purple line represents the total number of samples ( $n = 67$ ). **c** Average relative abundance of sPLS clusters (Fig. 4) across different ice cover, daylight and depth values determined by read recruitment from Arctic Ocean and Fram Strait metagenomes to the respective FRAM\_EGC MAGs. The ~4000 m sample from MOSAIC was not included.

**Atlantic signature populations.** Atlantic signature populations included *Thiotrichaceae* (asv45), OM182 (asv130) and SAR86 (asv157) from the *Gammaproteobacteria* (Fig. 6). Although all three populations were more abundant in the MIZ, differences were observed in their temporal dynamics (asv45 peaking during polar day, asv157 peaking during polar night, and asv130 showing minimal seasonality). The asv45 and asv130 populations both harboured genes for the degradation of phytoplankton-derived organic compounds. For asv45-*Thiotrichaceae*, this included the capacity to oxidise methanethiol (MTO gene) and the downstream reaction products, sulfide (*dsrAB* and *soeABC*) and formaldehyde (H4-MPT-dependent oxidation pathway), which could provide

carbon, sulfur and energy. The asv130-OM182 population encoded a more diverse substrate metabolism, with the capacity to use dissolved organic sulfur (DOS) and nitrogen (DON) compounds, such as taurine and methylamine, as well as carbon monoxide (CO) as supplemental energy source. The capacity to store and use elemental sulfur was evidenced by a polysulfide reductase and flavocytochrome *c*-sulfide dehydrogenase. Together with its flagellar machinery, this suggests a motile, heterotrophic, carboxydovorous lifestyle.

**Arctic signature populations.** Arctic signature populations included *Nitrospina* (asv118), OM75 (asv163), SAR86 (asv18) and



**Fig. 6 Temporal dynamics of signature populations.** Signature populations were identified as ASV representatives from sPLS clusters that a corresponding MAG was recovered for (based on 100% identity threshold competitive read recruitment). The temporal dynamics visualised are derived from ASV data. The missing chlorophyll data in 2016–2018 is due to the lack of a sensor on the MIZ mooring.

asv189) and Arctic97B-4 (asv191) affiliated with cluster C8, as well as BD2-11 (asv78) affiliated with cluster C7 (high-ice and polar night) (Fig. 6). Their metabolic potential and predicted ecological role varied considerably. The most prominent population (asv18-SAR86), reaching 8% relative abundance, harboured a heterotrophic metabolism with the capacity to gain supplemental energy through a green-light proteorhodopsin. Although similar to other SAR86 members [76, 77], asv18-SAR86 has an enriched repertoire of peptidases ( $n = 19$ ) compared to carbohydrate-active enzymes ( $n = 7$ ), as well as genes for D-amino acid metabolism.

Two of the Arctic signature populations were affiliated with enigmatic taxa, including the Arctic97B-4 (*Verrucomicrobiae*; *Pedospharaceae*) and BD2-11 (*Gemmatimonadota*). Arctic97B-4 was shown to be enriched in the particle-attached fraction in the Southern Ocean [78] and in subsurface waters [79, 80]. In comparison, BD2-11 has largely been observed in terrestrial and freshwater environments or in deep-sea sediments [81]. The

genomic content of the Arctic97B-4 population indicated a motile chemomixotrophic lifestyle with the capacity to fix carbon, assimilate sulfate, and synthesise the vitamins riboflavin and biotin. This population encoded a high number of CAZymes (23 genes) and sulfatases (84 genes). The most numerous CAZyme gene families are involved in animal glycan degradation, such as sialic acids (GH33). The BD2-11 population encodes genes for inorganic and organic compound metabolism, including aerobic denitrification (*nap*, *nirk*) and the metabolism of taurine, hypotaurine, D-amino acids, dicarboxylic acids and halogenated haloaliphatic compounds.

#### Whole-community functional shifts with contrasting environmental conditions

The raw HiFi reads contained 17.6 million ORFs (Supplementary Table S8), with 54% being assigned a function and 92% a taxonomy. Expectedly, taxonomic classifications varied in



resolution, with 92% of genes assigned to a kingdom and 37% to a genus (Supplementary Fig. S7). Evident taxonomic shifts over the annual cycle included higher proportions of *Bacteroidia* during polar day and low-ice cover; compared to *Verrucomicrobiae*, BD2-11 and *Marinimicrobia* under polar night and high-ice cover, in agreement with ASV dynamics. A dissimilarity analysis of community functionality separated samples into two distinct clusters, with ice cover being the only statistically significant factor between the two (F-statistic = 12.6,  $p = 0.009$ ) (Supplementary Fig. S8). A total of 1088 differentially abundant genes were identified between the two clusters, with 328 and 845 genes enriched under high- and low-ice conditions, respectively.

**Enriched functions under different ice-cover regimes.** In agreement with Arctic and Atlantic signature populations, the enrichment of genes under high- and low-ice cover suggested differences in substrate utilisation (Supplementary Fig. S9). Low-ice communities were enriched in genes involved in the utilisation of phytoplankton-derived carbohydrates as well as DON and DOS compounds, including dimethylsulfoniopropionate (DMSP), taurine, sulfoquinovose and methylamine (Fig. 7). In addition, glycoside hydrolase families involved in the degradation of laminarin,  $\alpha$ -galactose- and  $\beta$ -galactose-containing polysaccharides (GH16, GH36, GH42 and GH8), and genes related to the metabolism of mono- and disaccharides, such as D-xylose, glucose and rhamnose, were enriched (Fig. 7). All of these compounds have been related to phytoplankton production [82] and can act as carbon, nitrogen and sulfur sources for heterotrophic microbes [83, 84].

Under high-ice cover, 50% fewer genes were enriched, and they were mostly related to the recycling of bacterial cell wall carbohydrates, proteins, amino acids, aromatics and ketone compounds (Fig. 7). Reduced phytoplankton productivity under high-ice cover and during polar night [23] limits the availability of fresh labile organic matter, which would necessitate alternative growth strategies. For instance, the enrichment of an assimilatory nitrate reductase gene (*nap*) indicates a need for utilising inorganic nitrogen compounds. Enrichment of GH109 and GH18 involved in peptidoglycan and chitin degradation [85], along with genes for D-amino acid degradation, indicate an increased reliance on recycling of bacterial-derived organic matter. Furthermore, we observed an enrichment in genes for the degradation of aromatic and ketone compounds, such as phenylpropionate (Fig. 7).

## DISCUSSION

In recent decades, the Atlantic influence in the Arctic Ocean has expanded, a process termed Atlantification [5, 6]. Atlantification encompasses the multi-faceted physicochemical impacts of northward-flowing AW, such as accelerated sea-ice decline, weakened water column stratification and altered nutrient availability. Although its impact on microbial communities has been postulated [14], we provide the first high-resolution analysis over a natural mixing zone between outflowing PW and inflowing AW in Fram Strait to assess potential ecological implications. We show that sea-ice cover and AW influx have a considerable impact on the composition, structure and functionality of bacterial communities. Densely ice-covered PW harboured a temporally stable, resident microbiome capable of using versatile substrates, with an enriched potential to degrade bacterial- and terrestrial-derived substrates as well as inorganic compounds. In contrast, low ice cover and high AW influx coincided with seasonally fluctuating populations that are functionally linked to phytoplankton-derived organic matter. We further identified bacterial signatures of distinct environmental conditions in the EGC (Fig. 8), showed the consistency of these patterns across the wider Arctic Ocean, and assessed ecological roles through MAG-based functional gene content. Our combined

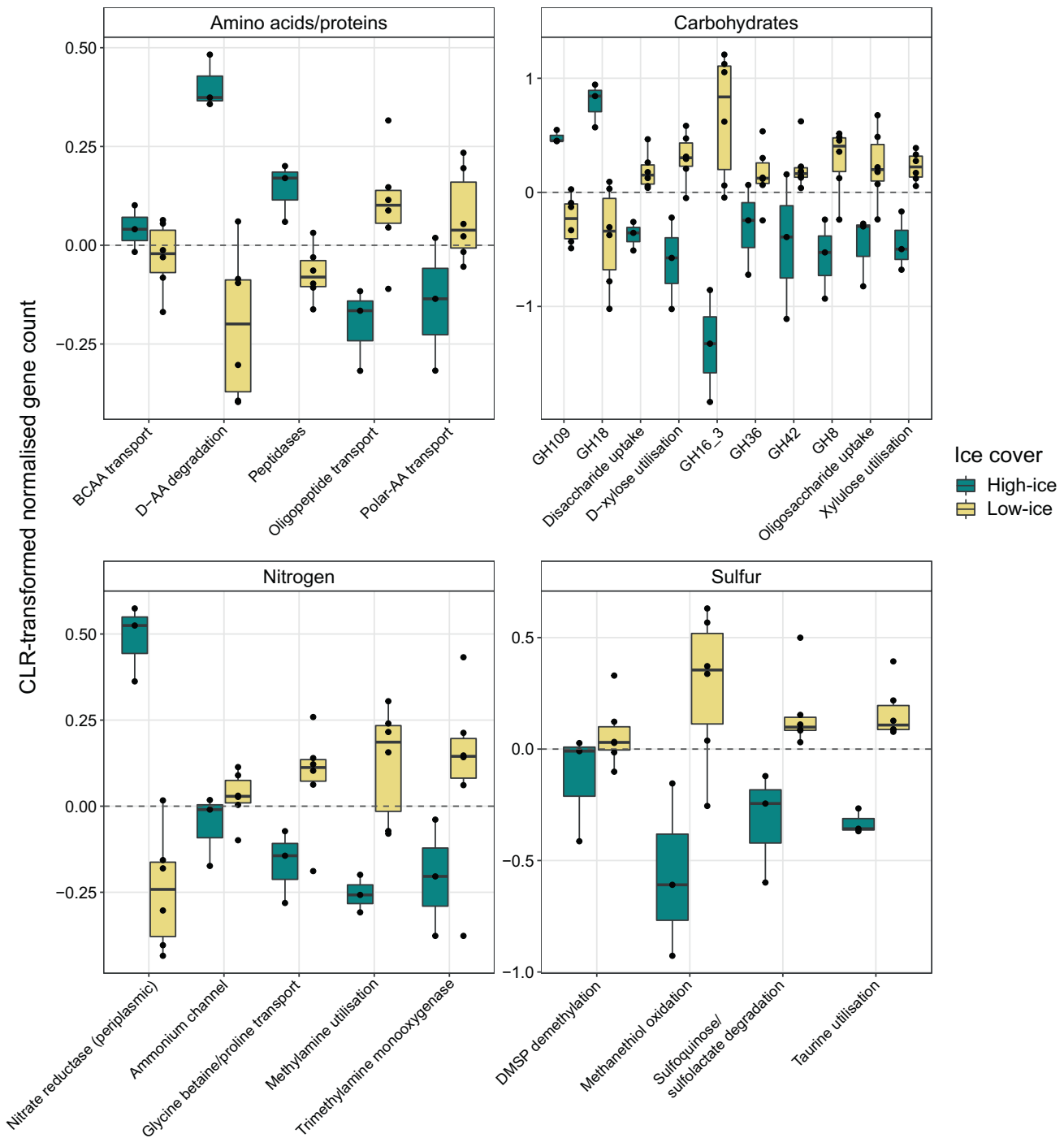
population- and community-level evidence suggests a future “Biological Atlantification” of the Arctic Ocean.

## Bacterial communities under different water mass and ice cover regimes

The pronounced impact of AW influx reflects the role of water masses as physical barriers to and conduits of dispersion for planktonic organisms. Influx events thus result not only in physicochemical changes, but also the mixing of microbial communities. How the microbiomes are reshaped under these events is a function of the degree of influx as well as the size (in number), competitive fitness and physiological adaptations of individual populations. Our dataset reveals that large influx events over short timescales can lead to the “replacement” of populations, evidenced by the dominance of AW-derived populations (Int-ASVs) in the MIZ. In contrast, the core-EGC, with rare occurrences of AW influx, harboured a temporally stable resident community that is adapted to polar conditions and constantly seeded from southward-flowing PW. However, the continual detection of the resident community in the MIZ indicates that even large influx events do not result in complete community turnover. Although the hydrological dynamics assessed here are more rapid than the gradually proceeding Atlantification of other Arctic regions, northward advection of organisms and subsequent replacement has already been documented for phyto- and zooplankton [7, 9, 86].

In addition to AW influx, bacterial communities were significantly impacted by sea-ice cover, which reflects its integral role in shaping Arctic Ocean ecosystems. Of particular significance is the influence of sea ice on water column stratification and organic matter availability. Sea ice supports rich biological communities that contribute significantly to Arctic Ocean primary production and the pool of organic matter [87, 88]. The melting of sea ice results in the release of dissolved and particulate organic matter, which heterotrophic bacteria can be highly responsive to [16, 17, 89]. However, ice-derived meltwater also induces rapid and strong stratification of the water column, which can reduce the mixed layer depth to as little as 5 m [11]. This shallow mixed layer can support prolonged phytoplankton blooms, but also trap the produced organic carbon, delaying vertical export [11]. In contrast, ice-free conditions result in a deeper mixed layer, shorter but more pronounced phytoplankton blooms and a higher response of grazers [11], potentially contributing to an increased availability of organic carbon to communities below. Considering the sampling depth in this study (70–80 m), the bacterial communities likely experienced an indirect influence from sea ice, through its impact on mixed layer depth, mixing and the vertical export of surface water production.

AW influx and sea-ice cover are intrinsically linked in the Eurasian Arctic. Consequently, the majority of signature populations were associated with either Arctic (high-ice and low-AW) or Atlantic (low-ice and high-AW) conditions. Furthermore, these populations were metabolically distinguishable, with Arctic signature populations harbouring genes for chemoautotrophy and the utilisation of bacterial and/or terrestrial-derived compounds. In the Beaufort Sea and Canadian Arctic, heterotrophic *Alphaproteobacteria* (*Rhodobacterales* and *Rhodospirillales*) and SAR324 (*Chloroflexi*) were shown to encode [90] and transcribe [91] pathways for the degradation of terrestrial-derived aromatic compounds. Similarly, Royo-Llonch et al. [74] described a number of bacteria as Arctic habitat specialists with versatile metabolisms, including the potential for autotrophy and denitrification. In this study, we found further examples of specific adaptations to Arctic Ocean conditions. For instance, the asv18-SAR86 population appears adapted towards proteinaceous and bacterial-derived compounds, with a reduced capacity for carbohydrate degradation compared to other SAR86 [76, 77]. In addition, the asv78-BD2-11 population encodes the capacity to use diverse inorganic and



**Fig. 7 Selected genes involved in the metabolism of organic and inorganic compounds enriched under high- and low-ice conditions.** Enrichment is displayed as centred-log ratio transformed normalised gene counts. Where several genes of a single pathway or mechanism were identified as enriched, they were grouped into one and the term ‘utilisation’ used (e.g. “taurine utilisation” indicates the uptake and degradation of taurine). When single genes were identified, the corresponding gene names are included. AA amino acids, BCAA branched-chain amino acids, GH glycoside hydrolase.

organic substrates, indicating a high degree of metabolic flexibility. The metabolic distinctions of Arctic signature populations illustrate evolutionary adaptations to the unique hydrological and physicochemical conditions. The Arctic Ocean is characterised by a comparatively large terrestrial and riverine influence [92, 93] and experiences a short productive season with a single phytoplankton bloom, compared to biannual bloom events in temperate oceans. This results in an organic matter pool rich in terrestrial-derived material, up to 33% in the case of DOM

[94], which has likely contributed to the enrichment of distinct metabolic potentials.

Atlantic signature populations featured a closer relation to labile, phytoplankton-derived organic matter. For example, the *asv45-Thiotrichaceae* population harbours genes for the degradation of methanethiol and its downstream reaction products. Methanethiol originates from DMSP demethylation [95], an osmoprotectant produced by phytoplankton. DMSP concentrations in the Arctic Ocean are spatially heterogeneous and

influenced by water mass and sea ice, with highest concentrations in areas with AW inflow [96] where its availability is tightly coupled to chlorophyll [96, 97]. Methanethiol concentrations would thus be elevated in AW during polar day. Similarly, the concentration of CO and its production by phytoplankton is also elevated in temperate compared to Arctic water masses [98]. The asv130-OM182 population encodes genes for CO degradation, as well as a capacity to use DOS compounds and store sulfur that may contribute to sustaining its more stable temporal dynamics in the MIZ. Given that previous reports of such a metabolism are restricted to members of the *Roseobacter* clade [99], the asv130-OM182 population may contribute to connecting carbon and sulfur cycles.

### Connectivity and structuring of bacterial communities across the Arctic

The identification of signature populations not only highlights ecological distinctions associated with different environmental regimes, but also aids in elucidating patterns in dispersal and connectivity in the Arctic Ocean. Although microbial species have been previously associated with certain depth layers and regions in the Arctic [74], this is the first study to evidence tight associations of populations with specific conditions over seasonally and geographically resolved scales. For example, polar day signatures found in the MIZ between 2016 and 2018 were also abundant in surface waters above continental shelves during the summer of 2013 (TARA). In addition, the here identified polar night and high-ice signature populations, which are of particular significance due to limited sampling of these conditions, were also abundant in the central Arctic during winter (MOSAIC). The consistency in dynamics of signature populations over space and time indicates a strong connectivity between Arctic regions, which is in agreement with the relatively short residence times of upper water layers [100]. Consequently, local environmental forcing is likely the key process shaping microbial communities. Furthermore, the prevalence of Arctic winter signatures in mesopelagic depths during summer suggests that solar- and meltwater-induced stratification contribute to shaping bacterial distribution.

In the core-EGC, where conditions are temporally stable, we identified a persistent, resident community fraction. The temporal stability of resident populations in the core-EGC, their variable dynamics in the MIZ, and low detection rate across summer Arctic samples suggests an adaptation to high-ice and PW conditions. However, in order to persist, the populations must be continually seeded from southward-flowing PW, underlining high dispersal and connectivity in the Arctic. Although the presence of a persistent

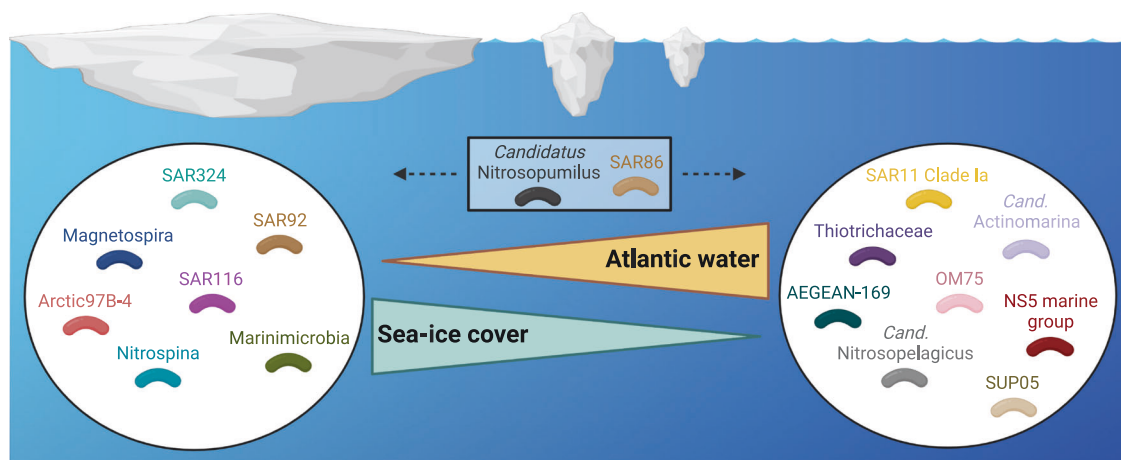
community fraction has been reported from the Western English Channel and Hawaiian Ocean time-series [71, 72], this is the first such description from the Arctic. It is also a feature likely restricted to the central Arctic and core-EGC, as bacterial communities of continental shelf and peripheral regions are exposed to more dynamic conditions and stronger seasonal forcing.

### Biological Atlantification and future Arctic Ocean bacterial communities

Considering population- and community-level dynamics in concert with contrasting environmental conditions, we predict a Biological Atlantification of Arctic Ocean bacterial communities. Biological Atlantification will be driven by northward advection of populations, coupled with shifting physicochemical conditions from expanding AW influence as well as its associated effects on primary producers and higher trophic levels. There are two underlying mechanisms; “replacement” through advection, mixing and species sorting (as outlined above), and physiological or evolutionary adaptation. We hypothesise that replacement will be more commonplace for bacteria with narrow ecological niches due to their sensitivity to change. In addition, replacement is more likely to occur in the central Arctic and above Eurasian shelves. With ice-free summers predicted by 2050, the central Arctic is shifting to a seasonally dynamic environment. This will reduce the niche space of bacteria that are adapted to permanent ice cover, while benefitting those adapted to conditions of the shelf and peripheral regions. Similarly, the Eurasian shelves will experience the immediate impact of Atlantification along with the northward expansion of temperate species. In short, we envision a net shift in bacterial distribution from shelf regions to the central Arctic, and from the North Atlantic onto the Eurasian Arctic shelves. However, adaptation will also play a role in the reshaping of communities, but will likely be more commonplace among bacteria with wider ecological niches and higher competitive fitness that are less vulnerable to changing conditions (Fig. 8).

### DATA AVAILABILITY

The 16S rRNA gene sequences are available at EBI-ENA under PRJEB43890 (2016–17), PRJEB43889 (2017–18), PRJEB54562 (2018–19), and PRJEB54586 (2019–20). Individual sample accessions are provided in Supplementary Table S9. The metagenomic sequence data and MAGs generated are available at EBI-ENA under PRJEB52171 (accessions provided in Supplementary Table S10). Tara Arctic data are available under PRJEB9740. MOSAIC accession numbers are shown in Supplementary Table S11. Functional gene annotations for all signature populations are provided in Supplementary Files S1. Physicochemical parameters are available under PANGAEA accessions 904565 [68], 941159 [69], and 946539 [70].



**Fig. 8 Bacterial communities under contrasting AW influx and ice cover conditions.** Illustration showing the ten taxonomic groups with highest average relative abundances under Atlantic vs. Arctic conditions, derived from the relative abundances of Int-ASVs (sPLS cluster C1) and Res-ASVs (sPLS cluster C8), respectively. Figure was generated using Biorender.com.

## CODE AVAILABILITY

Bioinformatic code for reproducing analyses and generating figures, along with necessary data files, is available at [https://github.com/tpriest0/FRAM\\_EGC\\_2016\\_2020\\_data\\_analysis](https://github.com/tpriest0/FRAM_EGC_2016_2020_data_analysis).

## REFERENCES

- Rantanen M, Karpechko AY, Lipponen A, Nordling K, Hyvärinen O, Ruosteenoja K, et al. The Arctic has warmed nearly four times faster than the globe since 1979. *Commun Earth Environ*. 2022;3:1–10.
- Kwok R. Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). *Environ Res Lett*. 2018;13:105005.
- Arctic Monitoring and Assessment Programme (AMAP). Snow, water, ice and permafrost in the Arctic (SWIPA). Oslo; 2017. <https://www.amap.no/documents/doc/snow-water-ice-and-permafrost-in-the-arctic-swipa-2017/1610>.
- Notz D, SIMIP Community. Arctic sea ice in CMIP6. *Geophys Res Lett*. 2020;47:e2019GL086749.
- Årthun M, Eldevik T, Smedsrud LH, Skagseth Ø, Ingvaldsen RB. Quantifying the influence of Atlantic heat on Barents Sea ice variability and retreat. *J Clim*. 2012;25:4736–43.
- Polyakov IV, Pnyushkov AV, Alkire MB, Ashik IM, Baumann TM, Carmack EC, et al. Greater role for Atlantic inflows on sea-ice loss in the Eurasian Basin of the Arctic Ocean. *Science*. 2017;356:285–91.
- Oziel L, Baudena A, Ardyna M, Massicotte P, Randelhoff A, Sallée J-B, et al. Faster Atlantic currents drive poleward expansion of temperate phytoplankton in the Arctic Ocean. *Nat Comm*. 2020;11:1–8.
- Lewis KM, Dijken GL, van, Arrigo KR. Changes in phytoplankton concentration now drive increased Arctic Ocean primary production. *Science*. 2020;369:198–202.
- Neukermans G, Oziel L, Babin M. Increased intrusion of warming Atlantic water leads to rapid expansion of temperate phytoplankton in the Arctic. *Glob Change Biol*. 2018;24:2545–53.
- Ardyna M, Arrigo KR. Phytoplankton dynamics in a changing Arctic Ocean. *Nat Clim Chang*. 2020;10:892–903.
- von Appen W-J, Waite AM, Bergmann M, Bienhold C, Boebel O, Bracher A, et al. Sea-ice derived meltwater stratification slows the biological carbon pump: results from continuous observations. *Nat Commun*. 2021;12:7309.
- Rapp JZ, Fernández-Méndez M, Bienhold C, Boetius A. Effects of ice-algal aggregate export on the connectivity of bacterial communities in the central Arctic Ocean. *Front Microbiol*. 2018;9:1035.
- Fadeev E, Rogge A, Ramondenc S, Nöthig E-M, Wekerle C, Bienhold C, et al. Sea ice presence is linked to higher carbon export and vertical microbial connectivity in the Eurasian Arctic Ocean. *Commun Biol*. 2021;4:1–13.
- Carter-Gates M, Balestreri C, Thorpe SE, Cottier F, Baylay A, Bibby TS, et al. Implications of increasing Atlantic influence for Arctic microbial community structure. *Sci Rep*. 2020;10:19262.
- Yergeau E, Michel C, Tremblay J, Niemi A, King TL, Wyglinski J, et al. Metagenomic survey of the taxonomic and functional microbial communities of sea-water and sea ice from the Canadian Arctic. *Sci Rep*. 2017;7:42242.
- Niemi A, Meisterhans G, Michel C. Response of under-ice prokaryotes to experimental sea-ice DOM enrichment. *Aquat Micro Ecol*. 2014;73:17–28.
- Underwood GJC, Michel C, Meisterhans G, Niemi A, Belzile C, Witt M, et al. Organic matter from Arctic sea-ice loss alters bacterial community structure and function. *Nat Clim Change*. 2019;9:170–6.
- Soltwedel T, Bauerfeind E, Bergmann M, Bracher A, Budaeva N, Busch K, et al. Natural variability or anthropogenically-induced variation? Insights from 15 years of multidisciplinary observations at the Arctic marine LTER site HAUS-GARTEN. *Ecol Indic*. 2016;65:89–102.
- Serreze MC, Barrett AP, Slater AG, Woodgate RA, Aagaard K, Lammers RB, et al. The large-scale freshwater cycle of the Arctic. *J Geophys Res Oceans*. 2006;111:C11010.
- de Steur L, Hansen E, Mauritzen C, Beszczynska-Möller A, Fahrbach E. Impact of recirculation on the East Greenland Current in Fram Strait: results from moored current meter measurements between 1997 and 2009. *Deep Sea Res Part I Oceanogr Res Pap*. 2014;92:26–40.
- Hofmann Z, von Appen W-J, Wekerle C. Seasonal and mesoscale variability of the two Atlantic water recirculation pathways in Fram Strait. *J Geophys Res Oceans*. 2021;126:e2020JC017057.
- Priest T, Orellana LH, Huettel B, Fuchs BM, Amann R. Microbial metagenome-assembled genomes of the Fram Strait from short and long read sequencing platforms. *PeerJ*. 2021;9:e11721.
- Wietz M, Bienhold C, Metfies K, Torres-Valdés S, von Appen W-J, Salter I, et al. The polar night shift: seasonal dynamics and drivers of Arctic Ocean microbes revealed by autonomous sampling. *ISME Commun*. 2021;1:1–12.
- Parada AE, Needham DM, Fuhrman JA. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol*. 2016;18:1403–14.
- Fadeev E, Cardozo-Mino MG, Rapp JZ, Bienhold C, Salter I, Salman-Carvalho V, et al. Comparison of two 16S rRNA primers (V3–V4 and V4–V5) for studies of Arctic microbial communities. *Front Microbiol*. 2021;12:637526.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41:590–6.
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13:581–3.
- Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and “All-species Living Tree Project (LTP)” taxonomic frameworks. *Nucleic Acids Res*. 2014;42:643–8.
- RStudio Team. RStudio: Integrated development of R. 2020. PBC, Boston, MA.
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. Vegan community ecology package version 2.5-7. 2020. <https://github.com/vegandevs/vegan>.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
- Le Cao K-A, Rohart F, Gonzalez I, Dejean S. mixOmics: omics data integration project. 2016. R package version 6.1.1. <https://www.bioconductor.org/packages/release/bioc/html/mixOmics.html>.
- Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer; 2016.
- Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;32:2847–9.
- Machné R, Murray DB, Stadler PF. Similarity-based segmentation of multidimensional signals. *Sci Rep*. 2017;7:12355.
- Csárdi G, Nepusz T, Traag V, Horvát S, Zanini F, Noom D, Müller K. igraph: Network analysis and visualization in R. R package version 1.5.0, 2023. <https://CRAN.R-project.org/package=igraph>.
- Benjamini Y, Hochberg Y. Controlling the false discovery rates: a practical and powerful approach to multiple testing. *J R Stat Soc, B Stat Methodol*. 1995;57:289–300.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13:2498–504.
- Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil P-A, Hugenholtz P. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res*. 2022;50:785–94.
- Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform*. 2010;11:119.
- Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015;12:59–60.
- Shen W, Ren H. TaxonKit: a practical and efficient NCBI taxonomy toolkit. *J Genet Genomics*. 2021;48:844–50.
- Seemann T. Prokka: rapid prokaryotic genome annotation. *J Bioinform*. 2014;30:2068–9.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinform*. 2009;10:421.
- Eddy SR. Accelerated profile HMM searches. *PLoS Comput Biol*. 2011;7:e1002195.
- Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, et al. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res*. 2018;46:95–101.
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 2014;42:490–5.
- Barbeyron T, Brillet-Guéguen L, Carré W, Carrière C, Caron C, Czjzek M, et al. Matching the diversity of sulfated biomolecules: creation of a classification database for sulfatases reflecting their substrate specificity. *PLoS ONE*. 2016;11:e0164846.
- Saier MH Jr, Reddy VS, Moreno-Hagelsieb G, Hendargo KJ, Zhang Y, Iddamsetty V, et al. The transporter classification database (TCDB): 2021 update. *Nucleic Acids Res*. 2021;49:D461–D467.
- Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res*. 2018;46:624–32.
- Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, et al. KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *J Bioinform*. 2020;36:2251–2.
- Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, et al. A new view of the tree of life. *Nat Microbiol*. 2016;1:1–6.



53. Fernandes AD, Reid J, Macklaim JM, McMurrough TA, Edgell DR, Gloor GB. Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. *Microbiome* 2014;2:15.
54. Li H. Minimap2: pairwise alignment for nucleotide sequences. *J Bioinform.* 2018;34:3094–3100.
55. Nissen JN, Johansen J, Allesøe RL, Sønderby CK, Armenteros JJA, Grønbech CH, et al. Improved metagenome binning and assembly using deep variational autoencoders. *Nat Biotechnol.* 2021;39:555–60.
56. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 2015;25:1043–55.
57. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, et al. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ.* 2015;3:e1319.
58. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, et al. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol.* 2018;3:836–43.
59. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* 2017;11:2864–8.
60. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32:1792–7.
61. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *J Bioinform.* 2009;25:1972–3.
62. Price MN, Dehal PS, Arkin AP. FastTree 2 – approximately maximum likelihood trees for large alignments. *PLoS ONE.* 2010;5:e9490.
63. Letunic I, Bork P. Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 2019;47:256–9.
64. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *J Bioinform.* 2020;36:1925–7.
65. Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadukumar, et al. ARB: a software environment for sequence data. *Nucleic Acids Res.* 2004;32:1363–71.
66. Pruesse E, Peplies J, Glöckner FO. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics.* 2012;28:1823–9.
67. Orellana LH, Francis TB, Ferraro M, Hehemann J-H, Fuchs BM, Amann RL. Verucomicrobiota are specialist consumers of sulfated methyl pentoses during diatom blooms. *ISME J.* 2022;16:630–41.
68. von Appen W-J. Physical oceanography and current meter data (including raw data) from FRAM moorings in the Fram Strait, 2016–2018. PANGAEA; 2019. <https://doi.org/10.1594/PANGAEA.904565>.
69. Hoppmann M, von Appen W-J, Monsees M, Lochthofen N, Bäger J, Behrendt A, et al. Raw physical oceanography, ocean current velocity, bio-optical and biogeochemical data from mooring HG-EGC-5 in the Fram Strait, July 2018 - August 2019. PANGAEA; 2022. <https://doi.org/10.1594/PANGAEA.941159>.
70. Hoppmann M, von Appen W-J, Barz J, Bäger J, Bienhold C, Frommhold L et al. Raw physical oceanography, ocean current velocity, bio-optical and biogeochemical data from mooring HG-EGC-6 in the Fram Strait, August 2019 - June 2021. PANGAEA; 2022. <https://doi.org/10.1594/PANGAEA.946539>.
71. Gilbert JA, Field D, Swift P, Newbold L, Oliver A, Smyth T, et al. The seasonal structure of microbial communities in the Western English Channel. *Environ Microbiol.* 2009;11:3132–9.
72. Eiler A, Hayakawa D, Rappé M. Non-random assembly of bacterioplankton communities in the subtropical North Pacific Ocean. *Front Microbiol.* 2011;2:140.
73. Bowers RM, Kyrpidis NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol.* 2017;35:725–31.
74. Royo-Llonch M, Sánchez P, Ruiz-González C, Salazar G, Pedrós-Alió C, Sebastián M, et al. Compendium of 530 metagenome-assembled bacterial and archaeal genomes from the polar Arctic Ocean. *Nat Microbiol.* 2021;6:1561–74.
75. Mock T, Boulton W, Balmonte J-P, Barry K, Bertilsson S, Bowman J, et al. Multionics in the central Arctic Ocean for benchmarking biodiversity change. *PLoS Biol.* 2022;20:e3001835.
76. Hoarfrost A, Nayfach S, Ladau J, Yooshep S, Arnosti C, Dupont CL, et al. Global ecotypes in the ubiquitous marine clade SAR86. *ISME J.* 2020;14:178–88.
77. Francis B, Ulrich T, Mikolasch A, Teeling H, Amann R. North Sea spring bloom-associated Gammaproteobacteria fill diverse heterotrophic niches. *Environ Microbiome.* 2021;16:15.
78. Milici M, Vital M, Tomasch J, Badewien TH, Giebel H-A, Plumeier I, et al. Diversity and community composition of particle-associated and free-living bacteria in mesopelagic and bathypelagic Southern Ocean water masses: evidence of dispersal limitation in the Bransfield Strait. *Limnol Oceanogr.* 2017;62:1080–95.
79. Pajares S. Unraveling the distribution patterns of bacterioplankton in a mesoscale cyclonic eddy confined to an oxygen-depleted basin. *Aquat Micro Ecol.* 2021;87:151–66.
80. Chun S-J, Cui Y, Baek SH, Ahn C-Y, Oh H-M. Seasonal succession of microbes in different size-fractions and their modular structures determined by both macro- and micro-environmental filtering in dynamic coastal waters. *Sci Total Environ.* 2021;784:147046.
81. Mujakić I, Piwosz K, Koblížek M. Phylum Gemmatimonadota and its role in the environment. *Microorganisms.* 2022;10:151.
82. Ksionzek KB, Lechtenfeld OJ, McCallister SL, Schmitt-Kopplin P, Geuer JK, Geibert W, et al. Dissolved organic sulfur in the ocean: biogeochemistry of a petagram inventory. *Science.* 2016;354:456–9.
83. Landa M, Burns AS, Roth SJ, Moran MA. Bacterial transcriptome remodeling during sequential co-culture with a marine dinoflagellate and diatom. *ISME J.* 2017;11:2677–90.
84. Moran MA, Durham BP. Sulfur metabolites in the pelagic ocean. *Nat Rev Microbiol.* 2019;17:665–78.
85. Beier S, Bertilsson S. Bacterial chitin degradation—mechanisms and ecophysiological strategies. *Front Microbiol.* 2013;4:149.
86. Ershova EA, Kosobokova KN, Banas NS, Ellingsen I, Niehoff B, Hildebrandt N, et al. Sea ice decline drives biogeographical shifts of key *Calanus* species in the central Arctic Ocean. *Glob Change Biol.* 2021;27:2128–43.
87. Arrigo KR. Sea ice as a habitat for primary producers. In: Thomas DN (ed). *Sea ice*, 3rd ed. Wiley & Sons, New York, 2017. p. 352–69.
88. Fernández-Méndez M, Wenzhöfer F, Peeken I, Sørensen HL, Glud RN, Boetius A. Composition, buoyancy regulation and fate of ice algal aggregates in the central Arctic Ocean. *PLoS ONE.* 2014;9:e107452.
89. Piontek J, Galgani L, Nöthig E-M, Peeken I, Engel A. Organic matter composition and heterotrophic bacterial activity at declining summer sea ice in the central Arctic Ocean. *Limnol Oceanogr.* 2021;66:5343–5362.
90. Colatriano D, Tran PQ, Guéguen C, Williams WJ, Lovejoy C, Walsh DA. Genomic evidence for the degradation of terrestrial organic matter by pelagic Arctic Ocean Chloroflexi bacteria. *Commun Biol.* 2018;1:90.
91. Grevesse T, Guéguen C, Onana VE, Walsh DA. Degradation pathways for organic matter of terrestrial origin are widespread and expressed in Arctic Ocean microbiomes. *Microbiome.* 2022;10:237.
92. Jones BM, Arp CD, Jorgenson MT, Hinkel KM, Schmutz JA, Flint PL. Increase in the rate and uniformity of coastline erosion in Arctic Alaska. *Geophys Res Lett.* 2009;36:L03503.
93. Gondeev VV, Martin JM, Sidorov IS, Sidorova MV. A reassessment of the Eurasian river input of water, sediment, major elements, and nutrients to the Arctic Ocean. *Am J Sci.* 1996;296:664–91.
94. Opsahl S, Benner R, Amon RMW. Major flux of terrigenous dissolved organic matter through the Arctic Ocean. *Limnol Oceanogr.* 1999;44:2017–23.
95. Kiene RP. Production of methanethiol from dimethylsulfoniopropionate in marine surface waters. *Mar Chem.* 1996;54:69–83.
96. Uhlig C, Damm E, Peeken I, Krumpen T, Rabe B, Korhonen M, et al. Sea ice and water mass influence dimethylsulfide concentrations in the central Arctic Ocean. *Front Earth Sci.* 2019;7:179.
97. Park K-T, Lee K, Yoon Y-J, Lee H-W, Kim H-C, Lee B-Y, et al. Linking atmospheric dimethyl sulfide and the Arctic Ocean spring bloom. *Geophys Res Lett.* 2013;40:155–60.
98. Conte L, Szopa S, Séférian R, Bopp L. The oceanic cycle of carbon monoxide and its emissions to the atmosphere. *Biogeosciences.* 2019;16:881–902.
99. Luo H, Moran MA. Evolutionary ecology of the marine *Roseobacter* clade. *Microbiol Mol Biol Rev.* 2014;78:573–87.
100. Ekwurzel B, Schlosser P, Mortlock RA, Fairbanks RG, Swift JH. River runoff, sea ice meltwater, and Pacific water distribution and mean residence times in the Arctic Ocean. *J Geophys Res Oceans.* 2001;106:9075–92.

## ACKNOWLEDGEMENTS

We thank Jana Bäger, Theresa Hargesheimer, Rafael Stiens and Lili Hufnagel for RAS operation; Daniel Scholz for RAS and sensor operations and programming; Normen Lochthofen, Janine Ludszuweit, Lennard Frommhold and Jonas Hagemann for mooring operation; Jakob Barz, Swantje Ziemann and Anja Batzke for DNA extraction and library preparation, and Bruno Huettel, Christian Woehle and the technicians at the Max Planck Genome Centre in Cologne for metagenome sequencing. The captain, crew and scientists of RV Polarstern cruises PS99.2, PS107, PS114, PS121 and PS126 are gratefully acknowledged. We thank Oliver Ebenhöf and Eva-Maria Nöthig for helpful discussions. Ian Salter made essential contributions during the early phase of FRAM.

## AUTHOR CONTRIBUTIONS

TP performed ASV and metagenomics analysis. MW processed amplicon raw data into ASVs and coordinated the data analysis. TP and MW wrote the paper. WJvA contributed quality-controlled oceanographic data, and coordinated the mooring operations. EO and OP performed network analyses. STV provided quality-controlled chlorophyll data. CB, KM and AB co-designed and coordinated the autonomous sampling and mooring strategy, and contributed to interpretation of the results. TM and WB provided access and background information on MOSAiC data, and contributed to interpretation of the results. BMF and RA contributed to interpretation of the results and development of the story. All authors contributed to the final manuscript.

## FUNDING

Open Access funding enabled and organized by Projekt DEAL. This project has received funding from the European Research Council (ERC) under the European Union's Seventh Framework Program (FP7/2007-2013) research project ABYSS (Grant Agreement no. 294757) to AB. Additional funding came from the Helmholtz Association, specifically for the FRAM infrastructure, and from the Max Planck Society.

## 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/s41396-023-01461-6>.

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

**Reprints and permission 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023