

Beyond blooms: the winter ecosystem reset determines microeukaryotic community dynamics in the Fram Strait

Ellen Oldenburg, Raphael M. Kronberg, Katja Metfies, Matthias Wietz, Wilken-Jon von Appen, Christina Bienhold, Ovidiu Popa & Oliver Ebenhöh

Article - Version of Record



Suggested Citation:

Oldenburg, E., Kronberg, R. M., Metfies, K., Wietz, M., von Appen, W.-J., Bienhold, C., Popa, O., & Ebenhöh, O. (2024). Beyond blooms: the winter ecosystem reset determines microeukaryotic community dynamics in the Fram Strait. *Communications Earth & Environment*, 5, Article 643.
<https://doi.org/10.1038/s43247-024-01782-0>

Wissen, wo das Wissen ist.



UNIVERSITÄTS- UND
LANDESBIBLIOTHEK
DÜSSELDORF

This version is available at:

URN: <https://nbn-resolving.org/urn:nbn:de:hbz:061-20241209-105559-9>

Terms of Use:

This work is licensed under the Creative Commons Attribution 4.0 International License.

For more information see: <https://creativecommons.org/licenses/by/4.0>

<https://doi.org/10.1038/s43247-024-01782-0>

Beyond blooms: the winter ecosystem reset determines microeukaryotic community dynamics in the Fram Strait

Check for updates

Ellen Oldenburg ^{1,2,3,10} ✉, Raphael M. Kronberg ^{3,4,10}, Katja Metfies ^{5,6}, Matthias Wietz ^{3,7,8}, Wilken-Jon von Appen ⁹, Christina Bienhold ^{3,7}, Ovidiu Popa ^{1,11} & Oliver Ebenhöh ^{1,2,11}

The Arctic Ocean is undergoing a major transition as a result of global warming, with uncertain consequences for its ecosystems. Our study introduces an integrated analytical approach using co-occurrence networks, convergent cross-mapping, and energy landscape analysis. Applied to four years of amplicon data from Fram Strait, located at the boundary between Arctic and Atlantic waters, our method identifies keystone species in seasonal microbial communities, elucidates causal interactions, and predicts stable configurations across changing environments. We find strong evidence for a “winter reset”, implying that organisms representing the spring bloom are largely determined by prevailing environmental conditions during winter. In addition, our analysis suggests that winter communities may adapt more readily to expected Atlantification than summer communities. These results highlight the utility of innovative time-series analyses in disentangling ecosystem dynamics. This approach provides critical insights into Arctic ecological interactions, dynamics, and resilience and aids in understanding ecosystem responses to environmental change.

The Arctic Ocean is a unique ecosystem undergoing major transitions during climate change. Over the past two decades, temperatures have risen more than twice compared to the global average¹, combined with a reduction in sea ice and snow cover, exacerbating warming trends. In particular, the extent of Arctic sea ice has declined¹. These environmental changes have a wide range of consequences, including profound shifts in biodiversity², and thus have a fundamental impact on ecosystems of the Arctic Ocean. There are first signs that the geographical ranges of temperate species are shifting northwards³, while polar fish and ice-associated species experience a reduction in their habitat due to changing environmental conditions. These ecological changes impact the entire ecosystem stability¹. The complex relationship between biodiversity and ecosystem stability remains poorly understood, particularly in the Arctic Ocean. Consequently, the rapid changes in Arctic sea ice and environmental conditions require an improved understanding of the mechanisms governing the resilience and stability of biological processes and ecosystem functions in the Arctic Ocean. Within

marine ecosystems, primary production is a key service supporting all trophic levels^{4,5}, with implications for biodiversity, the abundance and community structure at higher trophic levels, and carbon sequestration. This distinct ecosystem feature is supported by a highly productive microalgal community that thrives in sea ice, accompanied by a remarkably diverse heterotrophic community ranging from bacteria to metazoans⁶.

Recent decades have seen an increase in pelagic phytoplankton and primary production in the Arctic Ocean, a direct consequence of global warming^{7–9}. In the Central Arctic Ocean (CAO), sea-ice algae rather than phytoplankton account for much of the primary production^{10,11} as they have the potential to initiate pelagic blooms beneath the ice¹². Typically, phytoplankton growth starts mainly within the marginal ice zone in spring, co-occurring with increased solar radiation and meltwater-induced stratification¹³. Over the past three decades, increasing evidence has documented the occurrence of under-ice blooms in the Arctic Ocean^{14–17}, while phytoplankton in the water column below the ice shows significant

¹Institute of Quantitative and Theoretical Biology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ²Cluster of Excellence on Plant Sciences, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ³Deep-Sea Ecology and Technology, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany. ⁴Mathematical Modelling of Biological Systems, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ⁵Polar Biological Oceanography, Alfred-Wegener-Institut Helmholtz Center for Polar and Marine Research, Bremerhaven, Germany. ⁶Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg, Oldenburg, Germany. ⁷Max Planck Institute for Marine Microbiology, Bremen, Germany. ⁸Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Oldenburg, Germany. ⁹Physical Oceanography of the Polar Seas, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany. ¹⁰These authors contributed equally: Ellen Oldenburg, Raphael M. Kronberg. ¹¹These authors jointly supervised this work: Ovidiu Popa, Oliver Ebenhöh. ✉e-mail: Ellen.Oldenburg@hhu.de

differences from the microalgal communities in the sea ice¹⁸. However, changes in the diversity of key species related to the increase in Arctic pelagic primary production and its impact on the marine ecosystem stability are currently unresolved.

Recent findings indicate that high temperatures in natural ecosystems may affect ecological stability, whereas the consequences of alterations to biodiversity remain variable¹⁹. Nevertheless, the underlying mechanisms remain a subject of debate and limited understanding²⁰. The presence of nearly 2000 phytoplankton taxa and 1000 ice-associated protists in the Arctic²¹ indicates the relevance of identifying keystone species in this wealth of Arctic marine microbial diversity that accounts for ecosystem stability^{22–24}.

Understanding biological and ecological dynamics across seasonal environmental gradients is substantially fostered by novel statistical approaches. In polar ecosystems, these gradients include polar day and night, as well as variations in sea ice cover, stratification, or nutrient concentrations. Techniques are now accessible to assess the impact of ecological variables on ecosystem stability. For instance, co-occurrence networks (CON) determine and visualize how species coexist within communities or ecosystems^{25,26}. However, in natural ecosystems, species interactions are subject to variation as a result of changes in environmental conditions, which can cause a transition from one stable state of co-occurrence to another²⁷. Cross-convergence mapping (CCM) helps to identify the causality of co-occurrence in complex ecosystems, i.e., which organisms might share mutual or other direct relationships. Energy Landscape Analysis (ELA) aids in building ecological models that simulate and predict how ecosystems respond to disturbances or changes of environmental parameters^{28–30}.

In this study, we establish a mathematical methodology to reveal seasonal patterns, suggest causal ecological relationships, and identify microbial keystone species in Western Fram Strait.

This major gateway between the Arctic and Atlantic Oceans has been studied for over 20 years within the framework of the long-term HAUSGARTEN and FRAM observatories³¹. Our study contributes an extended mathematical perspective on microbial inventories in Fram Strait, showing seasonal patterns and the influence of sea-ice on microbial dynamics and the biological carbon pump^{32–36}. Based on a 4-year metabarcoding dataset of microeukaryotic taxa in context of rich oceanographic data, sampled year-round in approximately biweekly intervals, we develop scenarios of their long-term resilience. Additionally, we predict taxa that play a crucial role in maintaining stable communities among Arctic eukaryotic planktonic food web. Furthermore, we seek to define keystone species that can serve as indicators for monitoring the consequences of environmental change for Arctic marine ecosystem stability. Using an unprecedented combination of network analysis techniques like co-occurrence networks and cross convergence mapping, along with energy landscape analysis, our objective is to elucidate which factors might determine the stability of Arctic marine ecosystems. This approach will greatly improve our understanding of the effects of climate change on this ecosystem.

Results

We examined a dataset of 1019 eukaryotic ASVs and eight environmental parameters compiled over 4 years at mooring site F4 in the West Spitsbergen Current (WSC) in Fram Strait. The aim was to characterize species communities, analyze causal relationships between ASVs, and identify keystone and resilient taxa with respect to the impact of various environmental conditions.

To accomplish this, we established a novel computational pipeline, coupling co-occurrence analysis with convergence cross mapping and energy landscape analysis. This allowed us to identify causal interactions among taxa in a co-occurring community and to identify stable community states across different environmental conditions.

Co-occurrence network reveals seasonal dynamics

The co-occurrence network (CON) comprised eight connected components, with a major component accounting for 98% (935) of all nodes, which

are connected by 8610 edges. In the following, we focus on this major connected component. The resulting undirected graph notably displays a clear seasonal cyclic pattern (Fig. 1A).

The network was partitioned using the Louvain community detection algorithm³⁷, revealing ten discrete community clusters (Fig. 1B) labeled by the season in which the majority of cluster members had their maximum abundance (Table 1). To further group the clusters, we submerge each 3 month period into one season. Two clusters were assigned to the transition autumn period (01TA and 02TA), three clusters were associated with the low light winter period (03LW, 04LW, and 05LW), and three clusters with the transition spring period (06TS, 07TS, and 08TS). Finally, clusters 09 and 10 were allocated to the high light summer period (09HS, 10HS).

Community composition

We explored the taxonomic composition per cluster to explore the seasonal associations of each specific taxonomic group. Alpha biodiversity, measured by Shannon entropy, decreases from summer through autumn and winter, gradually decreasing towards spring (Supplementary Information Fig. S1) except for cluster LW05. The beta biodiversity, measured by Bray-Curtis distance (Supplementary Information Fig. S2), between the winter and spring clusters (03LW, 04LW, 05LW, and 06TS) is notably lower than between most other cluster pairs, except that between 01TA and 10HS. Cluster 02TA shows, on average, a higher beta diversity compared to all other clusters, which can be explained by the fact that 02TA is the smallest cluster in terms of the number of ASVs (Supplementary Information Fig. S2).

We found distinct taxonomic compositions within various clusters. Photosynthetic organisms like Ochrophyta and Haptophyta dominate the light phases³⁸. In late spring (cluster 08TS), phototrophs make up more than 75% of ASVs, while during summer (clusters 08TS and 09HS) and early autumn (10HS), they still comprise over 25% of all ASVs. Mixotrophs are highly abundant in most clusters, while they clearly dominate during the late autumn transition (cluster 02TA). Through the complete dark period (clusters 03LW, 04LW, and 05LW) as well as in early spring (06TS and 07TS), heterotrophs, particularly Syndiniales, are dominant with a clear peak of abundance (more than 90% of ASVs) in mid-winter (cluster 04LW). During early spring (cluster 06TS) when sunlight appears again, mixotrophs increase in their abundance, highlighting the nuanced trophic dynamics during the annual cycle. This comprehensive analysis at taxa level provides insights into the composition of these clusters, shedding light on the prevalence and distribution of specific classes within distinct seasonal communities (Fig. 1C).

Convergent cross mapping identifies community interactions

Convergent cross mapping (CCM) was applied to predict causal relationships within and between seasonal clusters based on the underlying ASV dynamics. We project the CCM-derived weights onto the co-occurrence network, resulting in a directed graph consisting of 17,220 directed edges and 935 nodes. Here, a directed edge indicates that knowledge of the dynamics of the source node allows predicting the dynamics of the target node.

A comparative analysis of edge weights within the CCM network was conducted. The connectivity derived from the co-occurrence network was compared with theoretical edge weights and randomly permuted connections. To perform this comparison, a two-sided Kolmogorov-Smirnov test was used. The theoretical edge weights were derived from all possible connections between pairs of nodes in the co-occurrence network, excluding the existing true links (Section “Convergent cross mapping identifies community interactions”). The findings clearly show that there is a stronger causal influence (higher NMI values) between co-occurring ASVs compared to random or to unconnected nodes (Supplementary Information Figs. S8 and S9 and Tables S1 and S2). The significance level was set at a nonparametric *p* value of less than 0.05 calculated similar to³⁶. Trimming edges with non-significant NMI values (Supplementary Information

Fig. S10), produced a network graph consisting of 4597 edges and 719 nodes, divided into 18 disconnected components, with the largest module encompassing 706 nodes with 4572 edges (Fig. 2A, B).

Of the total 12,648 edges eliminated during trimming, 18.16% represent connections between taxa reaching their peak abundance in March. This selective removal has a profound impact on the network structure (Fig. 2A). In contrast to the co-occurrence network, the causal interaction network “breaks” during spring, as demonstrated in Fig. 2A (months March

and April). The winter cluster 05LW and the spring cluster 06TS collapse (see Fig. 2B), meaning that nodes disintegrate and large parts of the clusters are no longer connected to the rest of the network. This suggests that the corresponding connections in the co-occurrence network are not a result of causal interactions but rather result from other factors, possibly caused by the prevailing environmental conditions. Analysis of the betweenness centrality reveals that ASVs of Picozoa, Leegaardiella, Acantharea, Dinophyceae, MAST-1, and Syndiniales serve as essential hub nodes throughout

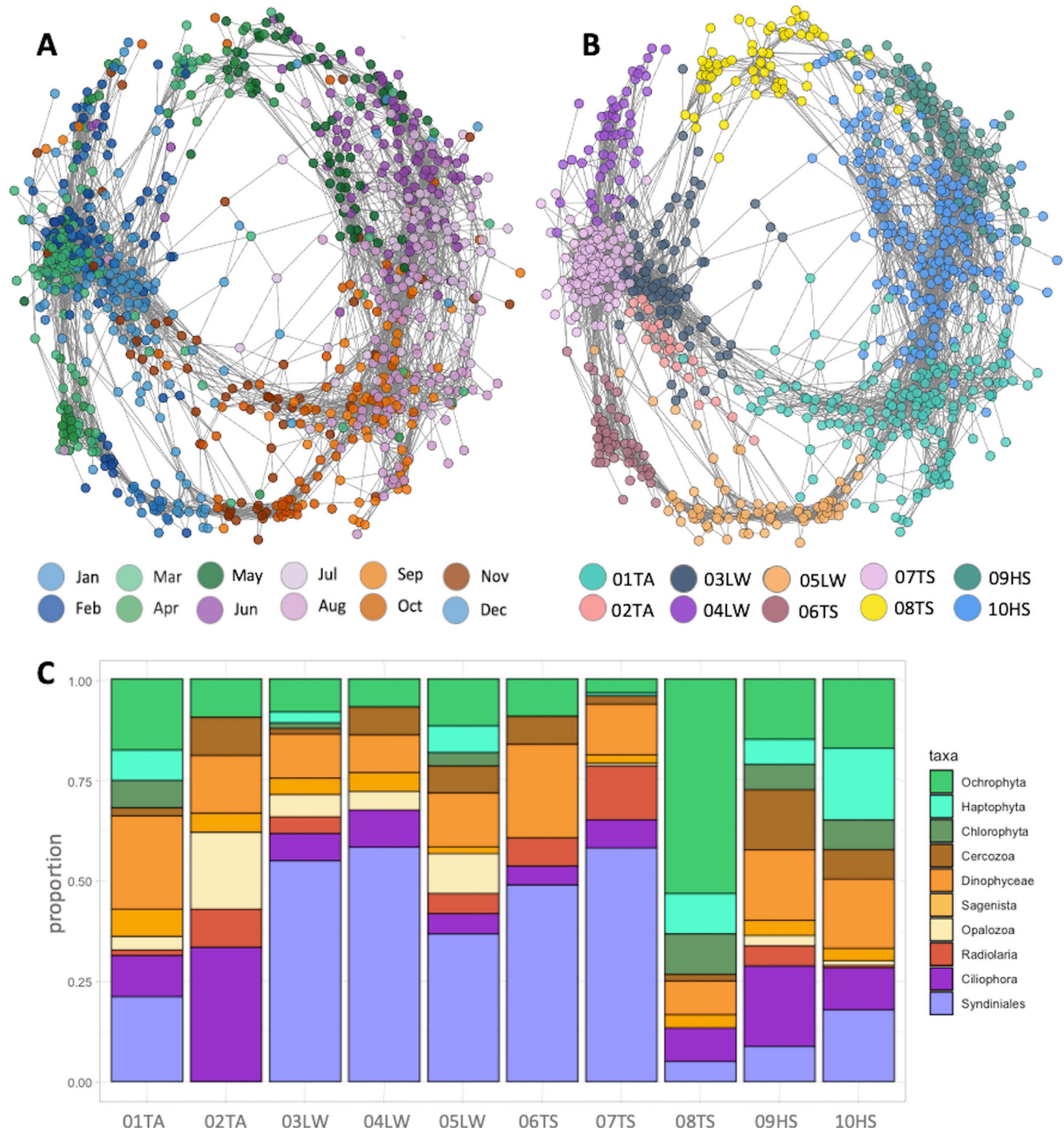


Fig. 1 | Co-occurrence networks and microeukaryotic composition at mooring F4 from 2016-08-01 to 2020-09-17. Each network node represents an ASV, and each edge represents a similar temporal pattern of two ASVs. The edge weights correspond to the Pearson correlation coefficients determined from the comparison of the individual ASV temporal profiles. ASVs are connected if the coefficient is $r > 0.7$, $p < 0.05$. Thus, edges in this figure indicate if the temporal profiles of two nodes display a significant correlation. **A** Node color reflect the month in which the

ASV exhibits maximal abundance, calculated from the maximum abundance mode for each year ranging from January to December. **B** In this representation, nodes are colored based on the community membership that was determined by the Louvain community detection algorithm. **C** The proportion of the top 10 taxonomic classes by cluster (“HS” high light summer, “LW” low light winter, “TS” transition spring and “TA” transition autumn). Color shades illustrate the assignment to auto- (green), mixo- (orange) or heterotroph (purple).

Table 1 | Co-occurrence network clusters

Cluster	MaxMonth	Season	Number of ASVs	Number of edges
01TA	Aug*	Autumn*	162	997
02TA	Nov	Autumn	24	86
03LW	Dec	Winter	78	522
04LW	Feb	Winter	52	212
05LW	Dec	Winter	74	467
06TS	Mar	Spring	51	388
07TS	Mar	Spring	153	3040
08TS	Apr	Spring	63	262
09HS	Jun	Summer	88	509
10HS	Jul	Summer	190	1189

Ten labeled clusters with their assigned season based on month in which the ASV exhibit maximal abundance, calculated from the maximum abundance mode (majority vote) for each year ranging from January to December, the number of ASVs per cluster, and the number of significant edges in network graph. *The stats denote that the MaxMonth is in August, but several nodes are also in September and October, therefore we mapped it to autumn, to model a transition autumn cluster.

the seasonal cycle in the network, highlighting their crucial function in maintaining network stability.

Community interaction

Analyzing cluster interactions revealed distinct patterns. We measure distance of clusters by “network distance”, a metric designed to evaluate the separation between clusters. This measure is computed by assessing the distance between the centroids of clusters within the dimension-reduced UMAP embedding space (Supplementary Information Fig. S4). Distances between clusters thus determined range between one and seven. Proximate clusters (a network distance of two to four) exhibited notably higher connectivity compared to clusters situated further apart (a distance of five to seven). Figure 2C provides a visual and quantitative representation of these interactions. Subsequent analysis revealed a prevalence of connections at a network distance of two (92% of total connections), followed by distances of three (6% of total connections), and four (2% of total connections) (Supplementary Information Figs. S4–S6).

For each seasonal cluster, we investigate in detail the mutual influence (NMI) of four taxonomic groups selected from the top ten classifications (Fig. 2D): Bacillariophyta, Syndiniales, Dinophyceae, and MAST (all MAST-X variants were classified under MAST). Bacillariophyta primarily comprises photo-autotrophic species³⁹, while Syndiniales include parasitic species, most of them characterized by their heterotrophic lifestyle⁴⁰. Dinophyceae are known for their diverse array of species and ecological roles, from symbionts to planktonic autotrophs⁴¹. MASTs are heterotrophic pico-eukaryotic protists and contribute substantially to protist abundances in the ocean. They play a crucial role in marine ecosystems, being among the dominant eukaryotes in the Arctic Ocean^{41,42}.

These four taxonomic groups are primarily distinguished by their unique lifestyles and ecological roles as primary producers, consumers, parasites, or endosymbiotic interactors. These distinctions form the basis for our analysis of their contributions to the ecosystem. Hence, by summarizing the members of each group into single nodes, we analyzed their cross-interactions using the information obtained from the CCM network.

The strength and direction of interactions between these taxonomic groups varied over the annual cycle (see Fig. 2D). During the spring-summer and summer-autumn transition, clusters 02TA and 06TS displayed fewer and weaker connections compared to other clusters (see Fig. 2B). This suggests dynamic changes in community structure during these transition phases, with ecological interactions between individual species either yet to be established or no longer present. At the beginning of the polar night (cluster 05LW), we detected the most substantial influence from the

dinoflagellates (shown by the thickest arrow in Fig. 2D) to the pico-eukaryotic heterotrophic groups Syndiniales and MAST, suggesting an ecological role of dinoflagellates for the establishment of the winter community. Overall, the strength of the links between the taxonomic groups decreases as the polar night progresses and has its minimum at the peak of the polar night in December (Cluster 04LW and 03LW). The lack of strong connections between taxonomic groups during the deepest polar night indicates that there are only a few ecologically significant interactions within the microeukaryotic community.

During polar day, the connections between taxonomic groups become stronger and reach their peak during the time span around summer solstice (Cluster 10HS). Notably, Bacillariophyta (i.e., diatoms) showed the most robust connections during the polar day, owing to their role as predominant phototrophic biomass producers at the foundation of the marine food web. However, towards the end of the growth period, the impact of MAST on Dinoflagellates becomes more pronounced (as shown by the thicker arrow), indicating an essential involvement of this pico-eukaryotic heterotroph in the ecosystem during the late polar day; a signature of the transition from primary production to recycling.

Community and environment interactions

For a more detailed understanding of which environmental conditions align with seasonal community clusters, we conducted a correlation analysis (Fig. 3). For our calculation, we define the time series of clusters to be the weighted sum of the abundance values of the ASVs assigned to this cluster. Cluster 10HS displayed a significant positive correlation with Photosynthetically Active Radiation (PAR) (0.64) and temperature (0.52) but a significant negative correlation with Mixed Layer Depth (MLD) (−0.64). This cluster thrives in environments with high light and temperature levels but stronger stratification. Cluster 03LW exhibits the opposite behavior, showing a moderately positive correlation with MLD (0.25) and polar water fraction (PW_frac) (0.21) while displaying an inverse relationship with PAR and temperature (−0.35 and −0.37, respectively).

Energy landscape analysis determines stability of microbial communities

We employed energy landscape analysis (ELA) to assess the stability of communities under the prevailing environmental conditions. We focus on four clusters representing the four seasons (01TA for autumn, 03LW for winter, 08TS for spring, and 10HS for summer). For each of these clusters, we determined the energy landscape, which is a highly complex function that depends on the abundances of all ASVs and the environmental parameters (see Fig. 4). To approximately visualize this landscape, we plot an interpolated smooth surface as a function of the two most significant NMDS dimensions. In addition, for each time point, we evaluate the energy landscape function and represent each energy value by a point in the three-dimensional diagram, where the z-axis represents the energy value.

For the landscape reconstructed for cluster 01TA (Fig. 4A), the autumn communities display lower energy values than the communities found in other seasons. This demonstrates that the autumn communities exhibit high stability. For the winter cluster 03LW (Fig. 4B), the picture is less clear. Whereas the interpolated energy landscape has a more pronounced minimum, the energy values of the observed communities are not clearly separated. As a tendency, the summer communities have a high energy value, demonstrating that summer communities are unstable in winter conditions. However, spring and autumn communities exhibit comparable energy values as winter communities, which indicates that stable community structures in winter conditions are not clearly defined. This trend is even more pronounced for the spring cluster 08TS (Fig. 4C). Here, the interpolated energy landscape shows a broad and shallow minimum, and the energy values of all observed communities, regardless of the season in which they are found, are very similar. This suggests that under spring conditions, community structures are not very stable and that community compositions show a high plasticity. As a consequence, many different communities may exist under spring conditions.

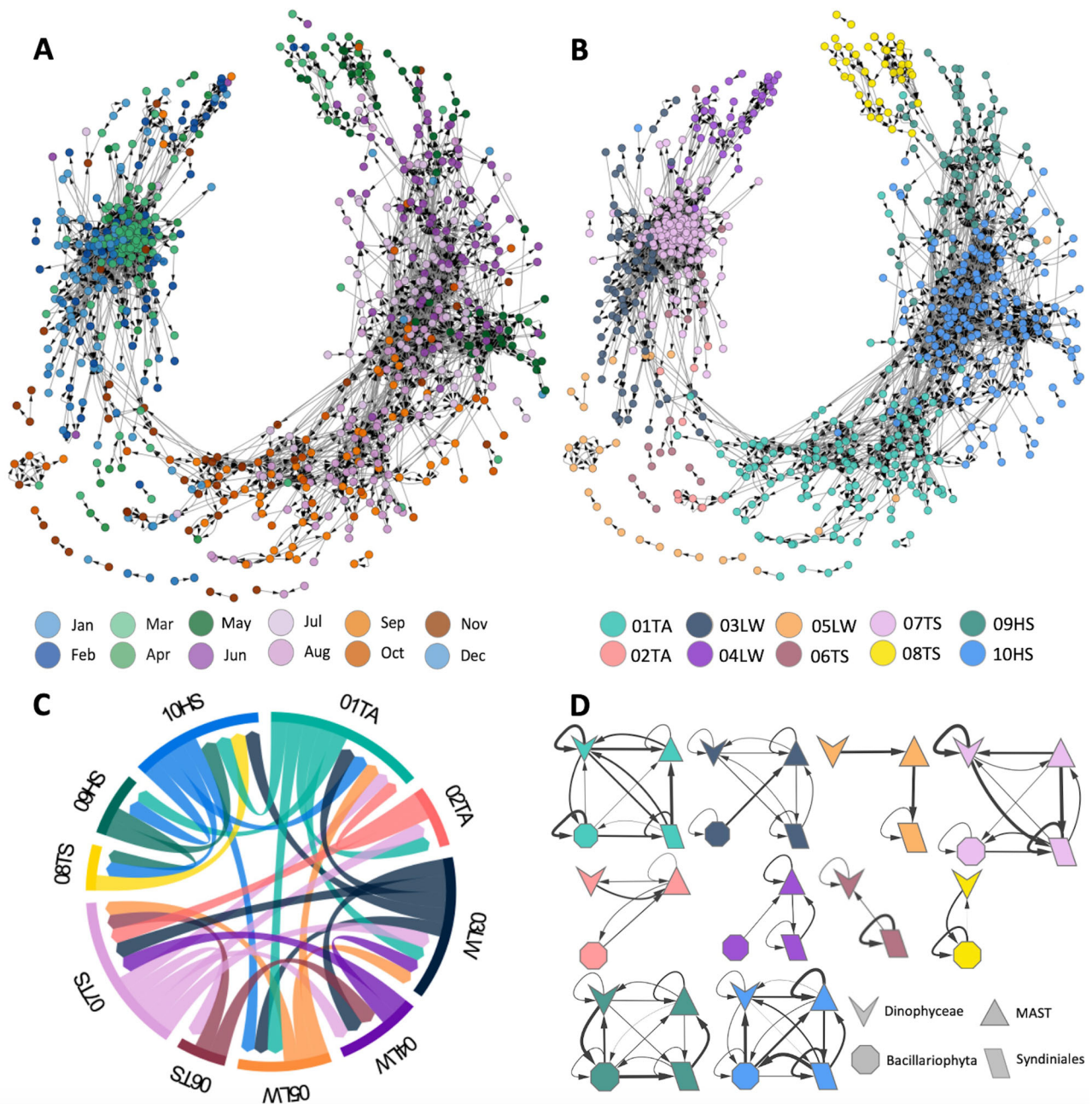


Fig. 2 | Convergence cross mapping networks of microeukaryotes at mooring F4 from 2016-08-01 to 2020-09-17. Each node in the CCM network represents an ASV, and each edge represents the causal influences. The edge weight corresponds to the Normalized Mutual Information determined from the comparison of the individual ASV and their predicted representation in the shadow manifold. ASVs are connected if the smoothed p value of the weight is $p < 0.05$ (Supplementary Information Fig. S3). **A** Node color reflects the month in which the ASV exhibits maximal abundance, calculated from the maximum abundance mode for each year ranging from January to December. **B** In this representation, nodes are colored based on the community membership that was determined by the Louvain community detection

algorithm. “HS” labels denote high light summer, “LW” represents low light winter, “TS” corresponds to transition spring, and “TA” indicates transition autumn. **C** The Normalized Mutual Information aggregated across the edges between the clusters, visually represented by thickness of the arrows corresponding to their respective values. Colors visually represent the clusters. **D** Interaction analysis between taxonomic clusters. For each of the ten clusters, the interactions between ASV groups are examined at class level, considering “Syndiniales”, “Dinophyceae”, “Bacillariophyta” and “MAST”. The cluster assignments are marked by different colors. The thickness of the arrows denotes the strength of the interaction, while the shapes represent the various taxa groups at the class level.

These findings demonstrate that knowledge of the composition of winter communities does not allow for the prediction of the composition of spring communities, which is in conjunction with the observation of the CCM analysis and the gap between winter and spring clusters (Fig. 2B). Finally, the energy landscape in summer (Fig. 4D) shows a pronounced minimum, in which the observed summer communities are also found. This indicates that summer conditions support well-defined communities with a high degree of stability.

Figure 5 displays how the energy values of the spring, summer, autumn, and winter clusters vary over the 4 years. These values demonstrate clear seasonal patterns over the 4-year period from August 2016 to September 2020. In general, the Autumn Cluster (01TA) displays minimal energy values during the autumn months (September, October, and November), indicating that autumn communities are typically more stable during this season (Fig. 5). In contrast, the seasonal minima for the Winter Cluster (03LW) are less clear, showing that the stability of this cluster is rather

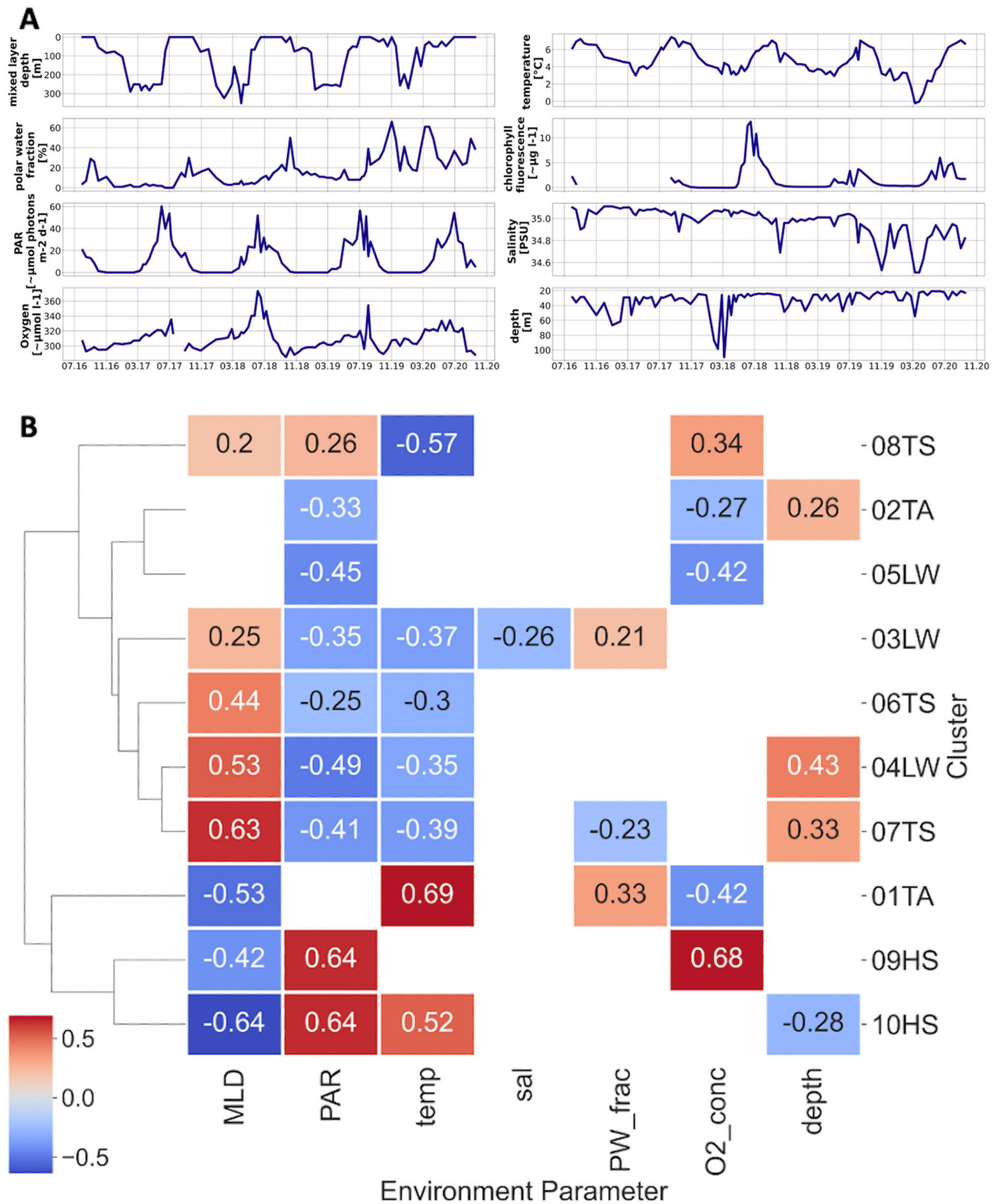


Fig. 3 | Environmental data and their correlation with Louvain clusters.

A Environmental data for F4 from 2016-08-01 to 2020-09-17. The x-axis represents the time period, while the y-axis indicates the following parameters: Mixed Layer Depth (MLD) [m], Temperature [°C], Chlorophyll Fluorescence [$\mu\text{g l}^{-1}$], Polar Water Fraction [%], Photosynthetically Active Radiation (PAR) [$\mu\text{mol photons m}^{-2} \text{d}^{-1}$], Salinity [Practical Salinity Units (PSU)], Oxygen Concentration [$\mu\text{mol l}^{-1}$], Depth of measurement [m]. **B** Correlations between environmental parameters and seasonal Louvain clusters. The displayed chart shows the environmental parameters from panel **A** in relation to seasonal clusters. These clusters are

characterized by the cumulative relative abundance of ASVs. The time series for the clusters are calculated as the weighted sum of the abundance values of the ASVs assigned to them. “TA” denotes transition autumn, “LW” low light winter, “TS” transition spring and “HS” high light summer. The color gradient used in the heatmap illustrates the strength of the correlation visually, with blue shades indicating negative correlations and red shades the positive correlations. It is worth noting that a significance mask has been applied to show only correlations that are statistically significant.

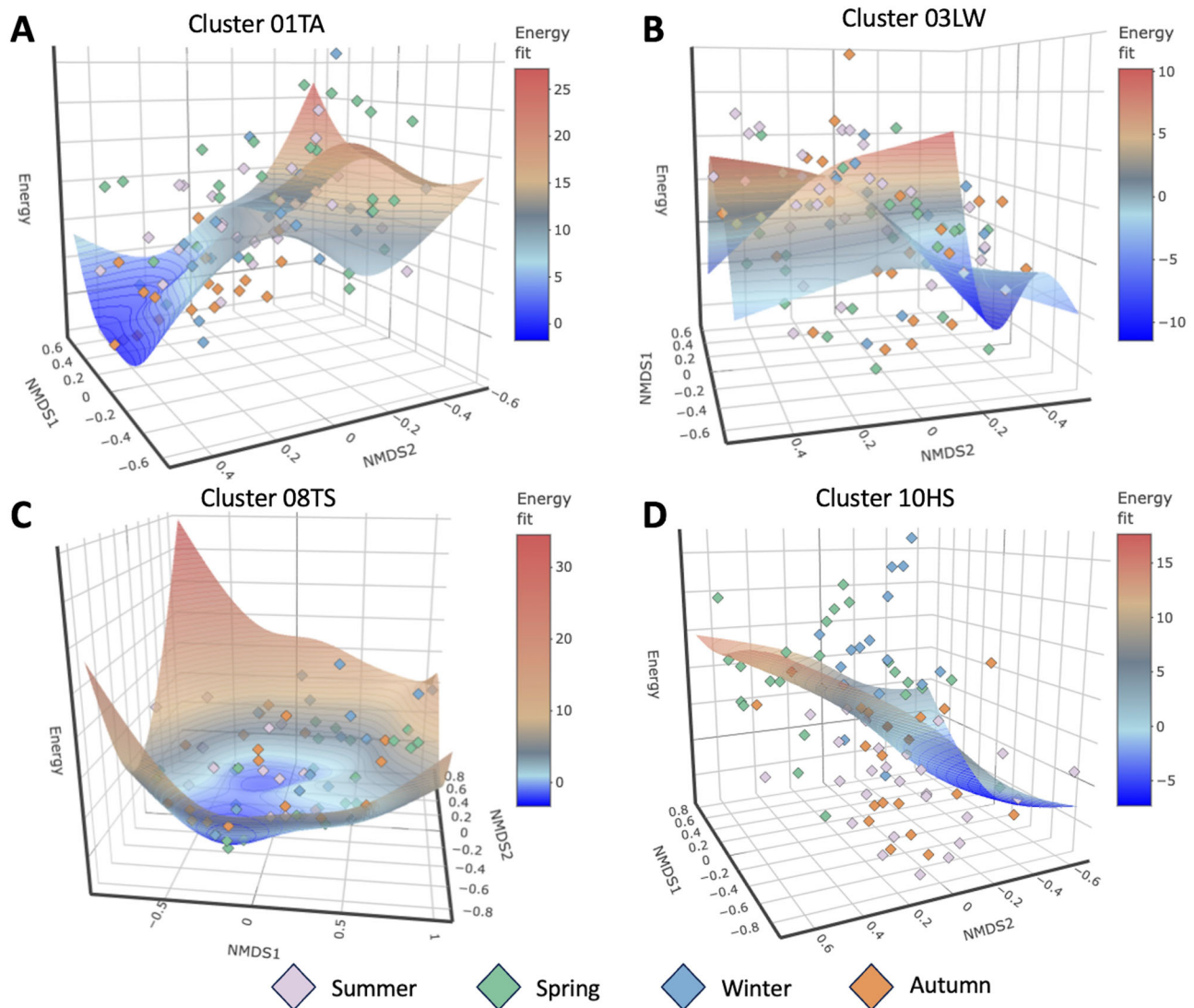


Fig. 4 | Energy landscapes depicting community structure dynamics. The plots display the reconstructed energy landscape on the NMDS surface for a cluster of each season. Environmental landscapes over the NMDS surface are reconstructed for each of the four example clusters. The z-axis displays the energy, while the x- and y-axes display the first and second NMDS dimensions. The landscape contours were

estimated using a smoothing spline approach with optimized penalty parameters. Community states, which are defined by ASV compositions and occupy lower-energy regions, indicate higher stability within the energy landscapes. **A** The transition autumn cluster 01TA. **B** The low light winter cluster 03LW. **C** The transition spring cluster 08TS. **D** The high light summer cluster 10HS.

independent of the season. The Spring Cluster (08TS) exhibits minima predominantly during the spring months (March, April, and May), suggesting a tendency towards stability during this season (Fig. 5). Likewise, the Summer Cluster (10HS) shows local minima during the summer months (June, July, and August), indicating a tendency towards stability in energy levels during this season (Fig. 5). Overall, each cluster demonstrates a proclivity for higher stability during its respective season, with the notable exception of the winter cluster. This underscores the influence of seasonal dynamics on the stability of microbial communities in these ecological communities in the Fram Strait but also indicates a potentially high plasticity in winter conditions.

Predicting keystone microeukaryotes in the Fram Strait

According to our definition (Section “Keystone species definition”), a keystone species is highly connected in the co-occurrence network, has a high influence on other species, and appears in a stable community. By contextualizing the evidence from CON, CCM, and ELA, we predict 38 keystone species across the annual cycle within the measured environmental profile (Table 2). 14 of these keystone species are associated with summer

clusters (three and eleven are found in clusters 09HS and 10HS, respectively), 13 with winter (eleven and two in the winter clusters 03LW and 04LW, respectively), eight are associated with autumn (cluster 01TA) and three with spring (cluster 07TS). The 14 keystone species from the summer clusters belong to the taxonomic groups Ochrophyta (6), Dinophyceae (4), Ciliophora (3), and Cryptophyta (1). These groups include *Fragilariopsis*, *Pseudo-nitzschia*, and *Thalassiosira*, major diatom taxa during Arctic blooms⁴³ that also serve as prey for microzooplankton^{44,45}. Notably, *Fragilariopsis* and *Thalassiosira* exhibited the highest abundance within this cluster. The keystone species in the winter clusters comprise Syndiniales (10), Radiolaria (1), Ochrophyta (1), and Dinophyceae (1); autumn cluster keystone species are Syndiniales (3), Ochrophyta (2), Chlorophyta (1), Dinophyceae (1) and unclassified Eukaryota (1). The spring keystone species belongs to Syndiniales (1) and Radiolaria (2), reflecting the major ecological strategies, including primary production, heterotrophy, and parasitism. The finding of only a few spring keystone species aligns with the greatest variability, as shown by ELA (Section “Energy landscape analysis determines stability of microbial communities”). The emergence of Chlorophyta during early autumn suggests a shift in primary production

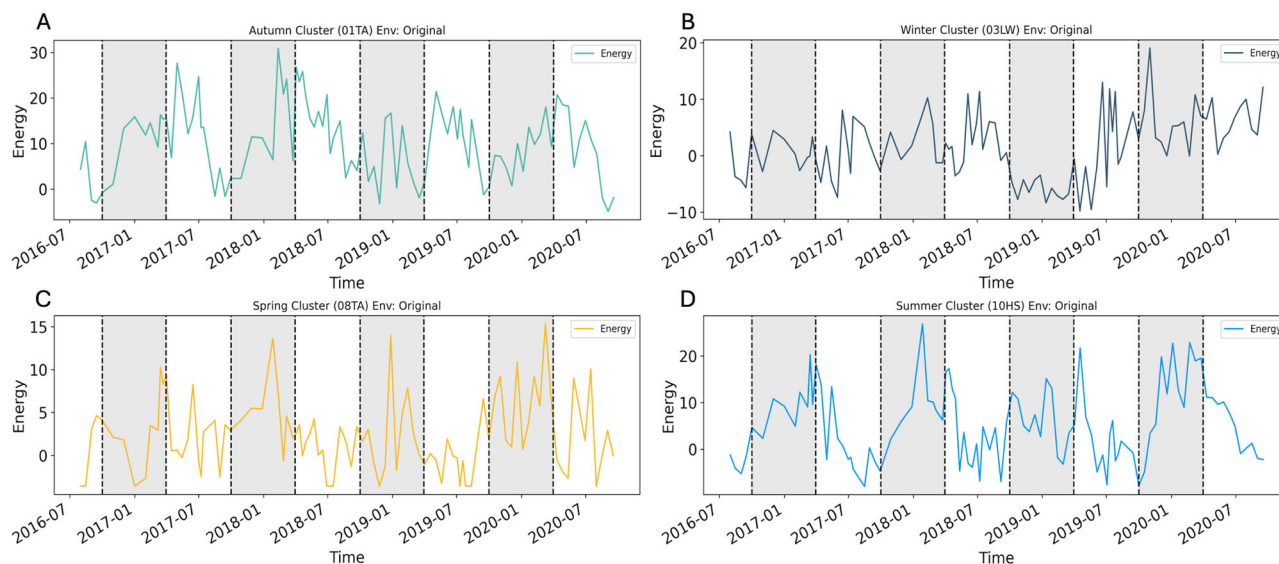


Fig. 5 | Energy landscape analysis illustrating seasonal dynamics of community energy levels over a 4-year span (08.2016–09.2020), with each subplot depicting a distinct season (autumn, winter, spring, summer). The energy curves, colored to represent specific clusters, are based on abundance and environmental data, as indicated. The x-axis denotes time, encompassing the specified period, while

environmental data integration underscores the complexity of ecological interactions shaping energy distributions. Lower energy levels are typically more stable, in contrast to higher energy levels. **A** The transition autumn cluster 01TA. **B** The low light winter cluster 03LW. **C** The transition spring cluster 08TS. **D** The high light summer cluster 10HS.

from Ochrophyta to Chlorophyta, including taxa that may prefer colder temperatures (ref. 46 and are better adapted to nutrient limitation⁴⁷.

Discussion

In this study, we present a new strategy for investigating ecological time series data based on 18S metabarcoding by combining three data analysis methods: Co-Occurrence Networks (CON)⁴⁸, Convergent Cross Mapping (CCM)²⁸, and Energy Landscape Analysis (ELA)³⁰. Integrating these three methodological approaches, we aimed to predict and characterize the abundance of keystone microeukaryotes in the West Spitsbergen Current across different seasons, environmental parameters, and in relation to other organisms. We also investigated how different taxa groups affect others and how their effects vary with seasonal shifts and environmental factors.

Sequence-based surveys of the 18S rRNA gene are effective for identifying and comparing variations in microbial eukaryote communities in marine environments. This approach offers a valuable tool for understanding the diversity and distribution of different taxonomic groups in the ocean⁴⁹. However, several aspects must be kept in mind when analyzing the data.

The taxonomic level of an amplicon sequence variant can vary in resolution due to missing references in the database or lacking diversity in the marker loci between strains. The presence of intracellular polymorphism, variation in rDNA copy number, and the existence of pseudogenes, in conjunction with the high sequencing depth of next-generation technologies, can lead to an inflation of diversity metrics. This is due to the fact that they capture a multitude of copies, pseudogenes, and other 18S rRNA variants within each organism, which consequently increases the number of predicted OTUs^{50,51}. For example, when determining dinoflagellate dominance using molecular methods, caution is required due to the influence of their disproportionately large genome size. Indeed, in the Fram Strait and extensive areas of the central Arctic Ocean, a significant proportion of sequence reads were attributed to Phaeocystaceae, *Micromonas* sp., Dinophyceae, and Syndiniales⁵².

A more detailed discussion about advantages and disadvantages of the metabarcoding approach is found in ref. 50. On the other hand, some studies showed that quality-controlled amplicon methods have high quantitative potential for determining the diversity and composition of marine protist communities and the relative abundances of specific ASVs and lineages^{53,54}.

Our co-occurrence network based on Fourier decomposition differs from previous methods that rely directly on the raw time series signals^{26,55,56}. The resulting network accurately captured seasonal states and transitions, revealing community clusters that reflect the prevailing community structure⁵⁷: in spring (cluster 08TS), primary producers such as Bacillariophyta appear and remain throughout the summer (09HS, 10HS), while mixotrophs increase in autumn (01TA, 02TA, 03LW) until almost exclusively heterotrophic and parasitic taxa dominate in winter (04LW, 05LW, 07TS). The considerable difference of spring clusters to other seasonal clusters (Supplementary Information Fig. S2) can be explained by the rapid environmental changes during this period (i.e., change from darkness to constant daylight within 20 days). The predominance of dinoflagellates in the intermediate phases of spring and autumn indicates that these mixotrophic organisms play a crucial role during transition phases^{58–60}.

Our CCM analysis revealed that by far, not all co-occurring ASVs actually influence each other (Fig. 2). A striking example is between clusters 06TS and 05LW, which were closely connected in the CON but not in the CCM network (Fig. 2). This co-occurrence without apparent causal connections could be explained by unique environmental conditions shaping both of these clusters, such as polar water influx. Even more pronounced is the separation of cluster 03LW, mainly heterotrophs, and 08TS, mainly phototrophs, which are tightly connected by co-occurrence but show not a single causal link in the CCM network. The organisms in these two clusters are primarily influenced by environmental parameters, particularly light. Additionally, these photosynthetic and heterotrophic organisms are sometimes preyed upon by the same predators⁶¹ such as Syndiniales. This explains the simultaneous occurrence and similar seasonality of these taxa but indicates that they do not have a direct influence on each other. The lack of causal influence during the transition from polar night to day is clearly visible in the CCM network (see Fig. 2). We interpret this gap between winter and spring clusters as a “winter reset” (Supplementary Information Fig. S11). This phase is characterized by the predominance of Syndiniales and Dinophyceae. With the emergence of light, a new period of primary production begins, shaped by the prevailing environmental conditions. The ambient environmental conditions then determine which species will subsequently prevail. By reflecting causal interactions between species, the CCM network even stronger reflects the cyclic microbiome structure than

Table 2 | ASV identified as potential keystone species for clusters 10HS, 06TS, 03LW and 01TA

Nodes	Phylum	Class	Genus	Species ^a	Cluster	Abundance proportion ^b	Closeness centrality
euk_asv19	Ochrophyta	Bacillariophyta	Thalassiosira	<i>Thalassiosira uc</i>	10HS	0.010	0.463
euk_asv12	Ochrophyta	Bacillariophyta	Fragilariopsis	<i>Fragilariopsis uc</i>	10HS	0.009	0.521
euk_asv29	Ciliophora	Spirotrichea	Strombidiidae_M uc	<i>Strombidiidae_M uc_sp.</i>	10HS	0.007	0.499
euk_asv35	Ochrophyta	Bacillariophyta	Fragilariopsis	<i>Fragilariopsis_sublineata</i>	10HS	0.007	0.533
euk_asv28	Ochrophyta	Bacillariophyta	Thalassiosira	<i>Thalassiosira uc</i>	10HS	0.006	0.473
euk_asv54	Dinoflagellata	Dinophyceae	Gyrodinium	<i>Gyrodinium_fusiforme</i>	10HS	0.006	0.469
euk_asv24	Ochrophyta	Bacillariophyta	Pseudo-nitzschia	<i>Pseudo-nitzschia_sp.</i>	10HS	0.006	0.446
euk_asv60	Dinoflagellata	Dinophyceae	Woloszynskia	<i>Woloszynskia_sp.</i>	10HS	0.005	0.489
euk_asv73	Cryptophyta	Cryptophyceae	Plagioselmis	<i>Plagioselmis_prolonga</i>	10HS	0.005	0.513
euk_asv52	Dinoflagellata	Dinophyceae	Peridinales uc	<i>Peridinales uc</i>	10HS	0.004	0.486
euk_asv125	Ciliophora	Spirotrichea	Dadayiella	<i>Dadayiella_ganymedes</i>	10HS	0.003	0.504
euk_asv15	Dinoflagellata	Dinophyceae	Gyrodinium	<i>Gyrodinium_fusiforme</i>	09HS	0.011	0.399
euk_asv115	Ochrophyta	Bacillariophyta	Mediophyceae uc	<i>Mediophyceae uc</i>	09HS	0.004	0.404
euk_asv186	Ciliophora	Spirotrichea	Strombidiidae_H uc	<i>Strombidiidae_H uc_sp.</i>	09HS	0.002	0.446
euk_asv5	Radiolaria	RAD-C	RAD-C uc	<i>RAD-C uc_sp.</i>	07TS	0.014	0.444
euk_asv23	Dinoflagellata	Syndiniales	Dino-I-1 uc	<i>Dino-I-1 uc_sp.</i>	07TS	0.005	0.404
euk_asv42	Radiolaria	RAD-C	RAD-C uc	<i>RAD-C uc_sp.</i>	07TS	0.005	0.419
euk_asv21	Dinoflagellata	Syndiniales	Dino-II-10-and-11 uc	<i>Dino-II-10-and-11 uc_sp.</i>	04LW	0.008	0.420
euk_asv607	Dinoflagellata	Syndiniales	Dino-II-10-and-11 uc	<i>Dino-II-10-and-11 uc_sp.</i>	04LW	0.000	0.411
euk_asv87	Radiolaria	RAD-B	RAD-B-Group-IV uc	<i>RAD-B-Group-IV uc_sp.</i>	03LW	0.002	0.489
euk_asv79	Ochrophyta	Chrysophyceae	Chrysophyceae H uc	<i>Chrysophyceae H uc_sp.</i>	03LW	0.002	0.409
euk_asv236	Dinoflagellata	Syndiniales	Dino-II-9 uc	<i>Dino-II-9 uc_sp.</i>	03LW	0.001	0.438
euk_asv213	Dinoflagellata	Dinophyceae	Gymnodinium	<i>Gymnodinium_sp.</i>	03LW	0.001	0.434
euk_asv198	Dinoflagellata	Syndiniales	Dino-II-6 uc	<i>Dino-II-6 uc_sp.</i>	03LW	0.001	0.433
euk_asv411	Dinoflagellata	Syndiniales	Dino-II-20 uc	<i>Dino-II-20 uc_sp.</i>	03LW	0.001	0.428
euk_asv615	Dinoflagellata	Syndiniales	Dino-II-20 uc	<i>Dino-II-20 uc_sp.</i>	03LW	0.001	0.508
euk_asv511	Dinoflagellata	Syndiniales	Dino-II-10-and-11 uc	<i>Dino-II-10-and-11 uc_sp.</i>	03LW	0.001	0.400
euk_asv553	Dinoflagellata	Syndiniales	Dino-II-21 uc	<i>Dino-II-21 uc_sp.</i>	03LW	0.001	0.430
euk_asv780	Dinoflagellata	Syndiniales	Dino-I-5 uc	<i>Dino-I-5 uc_sp.</i>	03LW	0.000	0.410
euk_asv1293	Dinoflagellata	Syndiniales	Dino-II-10-and-11 uc	<i>Dino-II-10-and-11 uc_sp.</i>	03LW	0.000	0.407
euk_asv9	Ochrophyta	Bacillariophyta	Pseudo-nitzschia	<i>Pseudo-nitzschia_sp.</i>	01TA	0.007	0.468
euk_asv44	Dinoflagellata	Syndiniales	Dino-II-23 uc	<i>Dino-II-23 uc_sp.</i>	01TA	0.006	0.526
euk_asv36	Chlorophyta	Chloropicophyceae	Chloroparvula	<i>Chloroparvula_pacifica</i>	01TA	0.005	0.501
euk_asv51	Ochrophyta	Bacillariophyta	Rhizosolenia	<i>Rhizosolenia_imbricata_var_shrubsolei</i>	01TA	0.003	0.404
euk_asv80	Eukaryota uc	Eukaryota uc	Eukaryota uc	<i>Eukaryota uc</i>	01TA	0.003	0.435
euk_asv101	Dinoflagellata	Syndiniales	Dino-I-3 uc	<i>Dino-I-3 uc_sp.</i>	01TA	0.003	0.445
euk_asv163	Dinoflagellata	Syndiniales	Dino-I-1 uc	<i>Dino-I-1 uc_sp.</i>	01TA	0.003	0.514
euk_asv177	Dinoflagellata	Dinophyceae	Gymnodiniaceae uc	<i>Gymnodiniaceae uc</i>	01TA	0.002	0.458

The taxonomic classifications, clusters, proportion of cluster relative abundance and Closeness centrality values are presented in summary form over the 4-year observation period. The ASV (Nodes) in this table has at least one significant CCM connection measured in Normalized Mutual Information.

^aSpecies are formatted in italic.

^bHellinger transformed raw abundance was summed over the 4 year period and divided by the total abundance of all ASVs.

the co-occurrence network. The cycle begins with photoautotrophs (cluster 08TS) in early spring and ends with the hetero- and mixotrophs (cluster 01TA) in late autumn. As light intensity decreases, mixotrophs become more prevalent than photoautotrophs, leading to a shift towards a heterotrophic lifestyle and a transition from carbon fixation to consumption. This transition into a low light period is characterized by parasitic species, suggesting an “eat and be eaten” scenario. The causal links from autumn to winter are much fewer than between other seasons (except winter to spring; Fig. 2B).

Energy landscape analysis was used to assess the stability of observed communities in their respective environments. The reconstructed landscape for autumn cluster 01TA shows that autumn communities are highly stable. However, for the winter cluster 03LW, the energy values of observed communities lack clear separation, making the situation less straightforward. Communities of the winter clusters which are still present in summer tend to display high energy values, indicating instability in winter conditions. The spring cluster 08TS shows an even more notable trend, indicating that community structures lack stability and exhibit high plasticity under

spring conditions (Fig. 5). This suggested plasticity during spring strongly supports our hypothesis of a “winter reset”, which was first proposed to explain the weak causal interactions between winter and spring clusters in the CCM network.

The “winter reset” phenomenon marks an ecological transition characterized by a major reorganization of microbial communities. This reset shifts the ecosystem from winter heterotroph dominance (e.g., Cluster 03LW) to spring photoautotroph prominence (e.g., Cluster 08TS). The lack of direct causal connections between these clusters in the Convergent Cross Mapping (CCM) network, despite their co-occurrence in the Co-Occurrence Network (CON), suggests that these transitions are driven by external environmental factors rather than direct species interactions. A similar observation was made with the 16S data, indicating that seasonal variations in environmental variables play a more significant role than trophic interactions⁶².

The increasing light availability in early spring drives the growth of photoautotrophic organisms. However, how well the individual species thrive and which of these species will dominate the population, critically depends on various other factors. The ambient temperature will favor some species more than others, as will the nutrient availability and salinity. These factors, in turn, are heavily influenced by water stratification, which creates its own unique spatial pattern of environmental conditions. Together with ocean currents, which carry material and organisms, this creates a situation in which low abundant resident populations of phototrophic microbes are in competition with species introduced by currents from other locations^{34,38}. Since every species possesses its own specific preferences for external factors such as temperature, salinity or nutrients, the outcome of this competition is likely to be highly sensitive on the prevailing environmental parameters, and much less so on the composition of the communities in the previous winter months, providing a possible explanation for the winter reset.

According to traditional ecological theory, keystone species are often defined as those with the most biomass^{63,64}. Here, we followed a more elaborate approach and defined keystone species as those found in stable communities (as determined by ELA), which are also highly connected in the CON and CCM networks. This way, keystone species have a strong causal interaction with many other species and form part of stable communities. Besides identifying highly abundant keystone species (for example *Fragilariopsis* or *Pseudo-nitzschia* diatoms), we also identified several lowly abundant ASVs as keystone, suggesting that both common and rare members contribute to ecosystem stability. The keystone species can be linked to central ecological functions carried out within the community network, including primary production, consumption, and parasitic interactions. During the beginning of autumn, Chlorophyta emerges as a keystone species, indicating a shift in primary production from Ochrophyta to Chlorophyta. This shift may be explained by the preference of Chlorophyta for colder temperatures and a better adaptation to nutrient limitation.

Our approach represents a strong advancement in the field of ecological network analysis, combining the techniques of co-occurrence network construction and convergent cross mapping to elucidate species interactions and community structures. In comparison to traditional methodologies that frequently rely on simplistic correlation analyses, our approach offers a number of distinct advantages. Traditional ecological network studies frequently employ simple correlation coefficients to infer species interactions, which may result in the overlooking of non-linear relationships⁶⁵. Clustering based on Fourier coefficients enables the capture of temporal profiles with greater fidelity to effectively summarize seasonal and long-term dynamics of species abundance⁶⁶ and additionally reduces the compositional bias³⁸. We choose the Hellinger transformation for our network analysis to normalize the abundance of data. The research demonstrated that transformation based on relative abundance is a crucial aspect of conducting appropriate analyses of eDNA metabarcoding data. Furthermore, the findings indicated that the Hellinger transformation exhibited a slight advantage over other methods⁶⁷. The Hellinger distance is recommended for clustering or ordination of species abundance data^{68,69}. In conducting the Energy Landscape analysis, the normalization method recommended by the author was

employed. Consequently, the abundance data was subsequently binarised^{29,30}.

A key limitation of correlation analyses is that they do not offer insights into causal relations, and therefore, it is difficult to infer species interactions from co-occurrence networks only⁷⁰. Integrating CCM addresses this limitation and extends beyond conventional correlation-based techniques by explicitly evaluating causal relationships between species pairs²⁸. Specifically, we compared our approach with two previous studies^{71,72}. While our approach focuses on the specific interactions within and between clusters, Ushio et al. provide a more general framework for predicting community diversity based on interaction capacity, temperature and abundance. The emphasis on mechanistic explanations for observed ecological patterns distinguishes the two approaches. Our methodology provides a comprehensive understanding of keystone species in a specific context, while Ushio's study provides broader insights into the factors influencing community diversity in different ecosystems. Both studies use similar techniques such as correlation and CCM⁷¹. Fujita's study used controlled experiments with six isolated community replicates, subjected to diverse treatments over 110 days. Regarding Takens' Theorem and Convergence Cross Mapping, Fujita et al. used Simplex projection to forecast population size⁷², while our study utilized pairwise CCM on ASV time series signals within clusters to predict keystone species.

The results presented in this study not only have practical implications for ecosystem management by improving our understanding and ability to predict change in complex ecological systems but also provide systematic insights into the mechanisms responsible for shaping and maintaining spatiotemporal heterogeneity in ecosystem composition²⁹.

Methods

Sampling and data

Samples were collected with Remote Access Samplers (RAS; McLane) deployed in conjunction with oceanographic sensors over four annual cycles (01.08.2016 to 16.09.2020 (96 Samples)) at the F4 mooring (79.0118N 6.9648E) of LTER HAUSGARTEN and FRAM in the Fram Strait^{38,73}. Each RAS contains 48 sterile bags, each collecting water samples of 500 mL at programmed sampling intervals.

The samples collected from 2016 to 2018 reflect the pool of up to two samples collected 1 h apart in two individual bags. Since 2018, we pooled samples taken 7 to 8 days apart from two consecutive weeks⁷⁴. The samples were preserved by adding 700 µl of mercuric chloride (7.5% w/v) to the bags prior to sampling. Pooled samples were filtered onto Sterivex filter cartridges with a pore size of 0.22 µm (Millipore, USA).

Filters were stored at −20 °C until DNA extraction and ribosomal metabarcoding of 18S rRNA reads using primers 528f (GCGTAATTC-CAGCTCCAA) and 926r (ACTTTCGTTCTTGATYRR)⁷⁵. The resulting amplicon sequence variants (ASVs) were classified using the PR2 4.12 database (see Supplementary Information Supplementary Methods: Data Preprocessing). We normalized raw ASV counts for CON and CCM using the Hellinger transformation but did not for the energy landscape analysis; hence a different normalization is introduced for the rELA implementation³⁰.

Temperature, salinity, and oxygen concentration were measured with a CTD-O₂ attached to the RAS. Physical oceanography sensors were manufacturer-calibrated and processed as described under³⁴. Raw and processed mooring data are available at PANGAEA <https://doi.org/10.1594/PANGAEA.904565>, <https://doi.org/10.1594/PANGAEA.940744>, <https://doi.org/10.1594/PANGAEA.941125> and <https://doi.org/10.1594/PANGAEA.946447>. For chemical sensors, the raw sensor readouts are reported. The fraction of Atlantic and Polar Water were computed following⁷⁶ for each sampling event and reported along with distance below the surface (due to mooring blowdown). Sea ice concentration derived from the Advanced Microwave Scanning Radiometer sensor AMSR-2⁷⁷ were downloaded from the Institute of Environmental Physics, University of Bremen (<https://seaice.uni-bremen.de/sea-ice-concentration-amsr-eamsr2>). Sentinel 3A OLCI chlorophyll surface concentrations were

downloaded from <https://earth.esa.int/web/sentinel/sentinel-data-access>. For all satellite-derived data, we considered grid points within a radius of 15 km around the moorings. 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>).

We considered eight environmental variables: mixed layer depth (MLD in m), water temperature (temp °C), polar-water fraction (PW_frac %), chlorophyll concentration from in situ sensor (chl_sens ~ $\mu\text{g l}^{-1}$), PAR ($\mu\text{mol photons m}^{-2}\text{s}^{-1}$), Salinity (PSU), oxygen concentration ($\text{O}_2\text{-conc } \mu\text{mol l}^{-1}$) and sampling depth (depth m)³⁴.

Co-occurrence network

The abundance of taxa over the full observation period was converted into temporal profiles by employing Fourier transformation techniques⁷⁸ to time-series signals. These temporal profiles rely on the 14 Fourier coefficients.

We chose 14 coefficients because they reflect the majority (662 of 1019) of observed ASV abundance peaks within the 4 years, which have peaked exactly once each year but no more than two times a year. To investigate the similarity of temporal profiles between ASV pairs, we performed pairwise correlations between the individual temporal profiles, where pairs with higher Pearson correlation values also show a similar temporal profile. Pairs with at least 0.7 ($p < 0.05$) Pearson correlation were then visualized in an undirected graph, corrected with Benjamini–Hochberg correction for multiple testing.

Only positive correlations were retained in the co-occurrence network to later focus exclusively on co-operative relationships, where both taxa were present. These relationships were then used to prune the causal network, ensuring that only co-operative interactions were considered in the following analysis. To identify strongly connected components that reflect the existing communities of co-occurring taxa, we applied the Louvain community detection algorithm³⁷ on the entire graph. The entire process was implemented using the CCM and networkx packages in Python; visualization was performed using Cytoscape with the Edge-weighted Spring-Embedded Layout⁷⁹. The whole co-occurrence network construction is described in Supplementary Information Supplementary Methods: Co-Occurrence Network.

Distance between clusters

To measure the distance between previously defined Louvain communities (clusters), we applied UMAP⁸⁰ on time-series signals obtained after Fourier decomposition of the abundance data. From this, we generated a three-dimensional embedding space. Centroids for each cluster were calculated within this space (see Supplementary Information Figs. S4–S6). The network distance between clusters was determined as the Euclidean distance between their centroids. Subsequently, a distance matrix was created, and distances were rounded to integers, with only significant connections retained.

Convergent cross mapping

Convergent cross mapping (CCM) identifies potential causal relationships between variables in time series data. It quantifies how knowledge of the time series of one taxon allows predicting the time series of another. We first built a CCM network from all pairwise combinations. From this, we extracted the in- and outgoing edges between nodes that are also connected in the co-occurrence network. We used the implementation of Normalized Mutual Information (NMI) from <https://github.com/polys/ennemi> by Petri Laarne and the Convergent Cross Mapping by Implementation from Prince Javier https://github.com/PrinceJavier/causal_ccm⁸¹ to measure the strength of the causal relationship considering also non-linear relations. We could show that the implementation of Normalized Mutual Information (NMI) yields similar findings as the original implementation based on Pearson correlation^{28,82} (see Supplementary Information Fig. S7).

Using a permutation approach²⁶ on the connectivity of the network, we calculated significance values for the edge weights, quantifying whether the respective NMI values are greater than expected for random edges (see Supplementary Information Fig. S8). The whole CCM network construction and validation are described in Supplementary Information Supplementary Methods: Convergent Cross Mapping and Supplementary Notes 2: Convergent Cross Mapping.

Aggregation on cluster level

We simplify the network of interactions between single ASVs into a network of interactions between clusters. For this, we assign a weight to a directed edge between two clusters by calculating the arithmetic mean of NMI of all (directed) edges connecting ASVs belonging to the respective clusters. This process effectively reduces the number of items in the node cloud, representing clusters through a unified composite node.

Energy landscape analysis

Energy landscape analysis is a method based on statistical physics. From data for many points in time, which contain taxon abundance and environmental variables, an *energy landscape* is reconstructed. This energy landscape is a function that maps ASV abundance and environmental variables to an energy value. In analogy to the potential energy in physics, a (local) minimum of this energy landscape indicates a stable community state. Here, we reconstruct the energy landscape function based on the complete time series of ASV abundance together with the available environmental data. We use the reconstructed function to determine the stability of observed communities, and in particular the seasonal clusters determined by the co-occurrence network, and we predict the most stable community compositions. Details of our analysis, including parameters and thresholds applied, are described in Supplementary Information Supplementary Methods: Energy Landscape Analysis. Understanding the existence and the nature of stable community states and how they change in response to environmental shifts is crucial for comprehending the resilience and adaptability of ecosystems in the face of various ecological challenges.

Definition of state space, pairwise maximum entropy models and energy landscape

Formalizing the stability landscape concept requires a precise definition of the state space within an energy landscape. We represent a community composition as a binary vector of length S , where S signifies the total number of ASVs. Within this framework, there exist 2^S unique community compositions. Specifically, a community composition of the k -th sample is denoted as $\sigma^{(k)} = (\sigma_1^{(k)}, \sigma_2^{(k)}, \dots, \sigma_S^{(k)})$, where $\sigma_i^{(k)} \in \{0, 1\}$ indicates the presence/absence status of the i -th ASV. To establish links between community compositions, we adopt the assumption that transitions occur incrementally. Consequently, two community compositions are linked if and only if they differ in the presence/absence status of precisely one taxon. This leads to the formation of a structured network wherein each node is connected to S neighbors^{29,30}. We attribute energy values to individual community compositions and establish the potential structure within the state space through the introduction of the extended pairwise maximum entropy model. This model governs the likelihood of observing community composition $\sigma^{(k)}$ under an environmental condition, denoted by $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_M)$, where M denotes the number of environmental variables and ϵ_i represents the continuous values denoting these environmental factors, such as resource availability, pH, temperature, salinity, etc. The probability of $\sigma^{(k)}$ occurring in condition ϵ is given by:

$$P(\sigma^{(k)}|\epsilon) = \frac{e^{-E(\sigma^{(k)}, \epsilon)}}{Z}, \quad (1)$$

with the energy defined as:

$$E(\sigma^{(k)}, \epsilon) = -\sum_{i=1}^S h_i \sigma_i^{(k)} - \sum_{i=1}^S \sum_{j=1}^M g_{ij} \epsilon_i \sigma_j^{(k)} - \sum_{j=1}^S \sum_{i=1, i \neq j}^S J_{ij} \sigma_i^{(k)} \sigma_j^{(k)} / 2, \quad (2)$$

where $E(\sigma^{(k)}, \epsilon)$ represents the energy of community composition $\sigma^{(k)}$ and:

$$Z = \sum_{k=0}^{2^S-1} e^{-E(\sigma^{(k)}, \epsilon)}. \quad (3)$$

The function E represents the energy landscape and is defined by the parameters, $\mathbf{h} = (h_1, h_2, \dots, h_S)$, matrix $\mathbf{J} = (J_{ij})$, and matrix $\mathbf{g} = (g_{ij})$. Based on the observed communities and the associated environmental factors, these parameters are estimated by Energy Landscape Analysis through a maximum likelihood approach^{29,30}. These estimates can be acquired through stochastic approximation for the extended pairwise maximum entropy model (Eq. (2)), as described in the Supplements Supplementary Information Supplementary Methods: Energy Landscape Analysis.

Here, h_i signifies the net impact of unobserved environmental factors favoring ($h_i > 0$) or hindering ($h_i < 0$) the presence of taxon i , and g_{ij} represents the influence of the i -th observed environmental factor on the occurrence of taxon j . The model captures pairwise relationships, as each taxon is interconnected with all others through J_{ij} . It is pertinent to note that the term $E(\sigma^{(k)}, \epsilon)$ is labeled as energy due to its analog in statistical physics⁸³, although it serves as an exponent in Eq. (1) and indicates the likelihood of observing a community composition within an ecological context. It does not correspond directly to physical energy as used in ecological studies.

Rather, the energy signifies the directionality of transitions between community compositions. For two adjacent nodes, $\sigma^{(k)}$ and $\sigma^{(k_0)}$, if $E(\sigma^{(k)}) > E(\sigma^{(k_0)})$, then the transition from $\sigma^{(k)}$ to $\sigma^{(k_0)}$ is more likely than the reverse.

In ecological studies, analyzing microbial communities often involves sifting through vast amounts of data, which presents major computational challenges due to the exponential increase in possible states as the number of amplicon sequence variants (ASVs) increases. To address this, limiting the analysis to the Top 100 ASVs provides a practical compromise between computational feasibility and analytical depth (as recommended by the Framework Owner, see Suzuki et al.³⁰). This approach allows for efficient management of the data while still yielding meaningful insights into the community structure. By concentrating on the most abundant ASVs, the key contributors to the microbial ecosystem dynamics can be captured because these variants generally hold greater ecological and biological significance due to their higher prevalence. Furthermore, focusing on the top 100 ASVs helps minimize the impact of less abundant variants, which may introduce noise and potentially distort the results. However, this method is not without its drawbacks; there is a risk of bias toward dominant ASVs, potentially overlooking less prevalent but ecologically important species. Such biases could influence the interpretation of community dynamics and functional roles, highlighting the need for careful consideration of these limitations in ecological analyses. In the Supplementary Information Fig. S12, we validated that the cutoff of $N = 100$ ASVs is a good choice by testing the stricter cutoff values $N = 50$ and $N = 75$. Outliers were excluded solely for the purpose of plotting.

Abundance normalization

For the network approaches (CON and CCM), we apply the Hellinger transformation to the abundance matrix. In contrast to previous studies that employed Pearson correlation for pairwise comparisons of relative abundance to ascertain co-occurrence patterns, we utilized Fourier decomposition to extract distinctive temporal profiles for each ASV. This approach mitigates the potential for bias associated with Pearson correlation in compositional data analysis.

Hellinger transformation

In the matrix $\mathbf{M} = \{m_{ij}\}$, the columns represent observed ASVs, and the rows represent different samples. The entries are the read counts of the ASVs in the given sample. The Hellinger transformation was applied to facilitate data normalization and comparison of ecological data, reducing the impact of differences in the scale of abundance values across samples (see Fig. 6A). The normalization is applied column-wise to the raw abundance data as follows:

- 1) Calculate the square root of each element:

$$s_{ij} = \sqrt{m_{ij}} \quad (4)$$

- 2) Calculate the ℓ^2 -norm (Euclidean norm) of each column:

$$r_i = \sqrt{\sum_j s_{ij}^2} \text{ (column-wise)} \quad (5)$$

- 3) Normalize each column by dividing it by its ℓ^2 -norm:

$$h_{ij} = \frac{s_{ij}}{r_i} \quad (6)$$

The energy landscape analysis uses the following normalization.

Proportional normalization

The proportional normalization process applied to the matrix \mathbf{A} (see Fig. 6B). Given a matrix \mathbf{A} of size $m \times n$ the normalized matrix \mathbf{N} is computed as follows³⁰:

- 1) Compute the row sums: define a column vector \mathbf{s} where each element s_i is the sum of the elements in the i -th row of \mathbf{A} :

$$s_i = \sum_j a_{ij} \quad (7)$$

where $\mathbf{A} = \{a_{ij}\}$.

- 2) Add a small constant:

$$d_i = s_i + 10^{-16} \quad (8)$$

- 3) Normalize the matrix: compute the normalized matrix $\mathbf{N} = n_{ij}$ using matrix division:

$$n_{ij} = \frac{a_{ij}}{d_i} \quad (9)$$

This normalization ensures that each row of the resulting matrix sums to 1, with a small adjustment to maintain numerical stability.

Keystone species definition

After collecting attributes from co-occurrence analysis and distinguishing between potential ecological influence and occurrence just by chance, we calculated the stable states for different clusters using ELA. This information was merged to suggest potential keystone species. We defined a keystone species as an ASV with (1) a significant influence on other ASV in the network (significant NMI value), (2) a high centrality (closeness) value within its co-occurrence community, and (3) presence in at least one stable state as predicted by ELA. A significant high centrality value was determined by comparing each centrality value of a single node to the average centrality values of all nodes from the graph using a one sided, one-sample t -test with Benjamini–Hochberg correction for multiple testing (similar to^{84,85}).

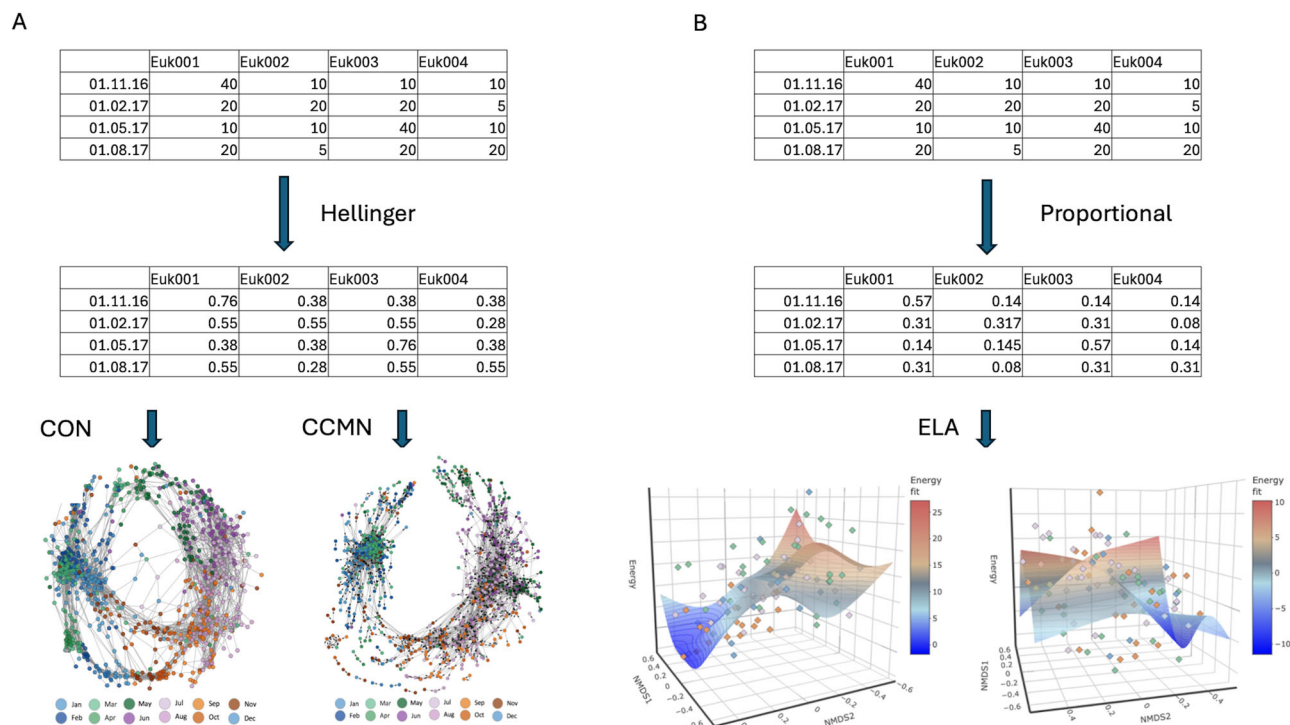


Fig. 6 | Schematic flowchart of our analytical approach. **A** Network analysis. The abundance table is normalized using the Hellinger transformation Section “Hellinger transformation”, followed by co-occurrence network (CON) and convergent

cross mapping network (CCMN) analysis. **B** Energy landscape analysis. The abundance table is normalized using the proportional normalization Section “Proportional Normalization”, followed by energy landscape analysis (ELA).

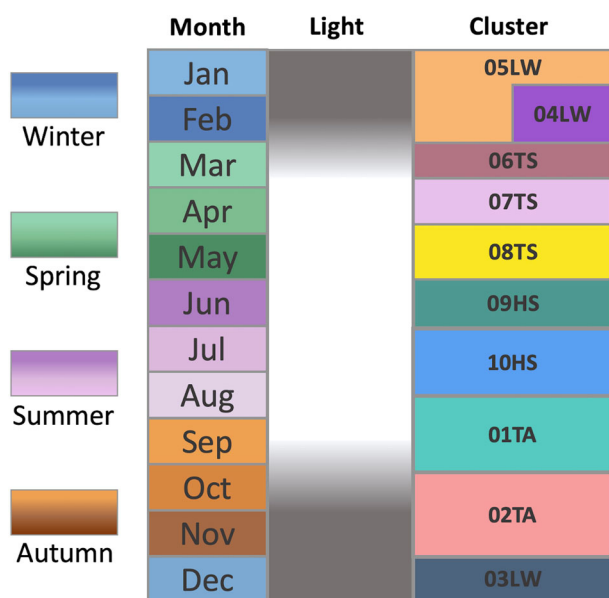


Fig. 7 | Schematic cluster definition. Names are based on light availability (as defined in ref. 38), categorized in transition areas between dark and light (T), high light (H) and low light (L) phases based on PAR parameter (Fig. 3) and the season spring (S), summer(S), autumn (A) and winter (W).

Season definition

For assessing results in context of the entire annual variability over which samples were collected, we defined the seasons as follows, based on month and the availability of light (PAR). In the case of Cluster 01TA, the month of maximal abundance is August. However, several nodes are also present in September and October. Consequently, we mapped this cluster to the

autumn season in order to model a transition from the autumn cluster (see Fig. 7).

Reporting summary

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

Data availability

Raw metabarcoding reads will be accessible under ENA.

Code availability

The code for raw data processing will be available at <https://gitlab.com/qtb-hhu/marine/publications/fastq2abundance> and the code for data analysis at <https://gitlab.com/qtb-hhu/marine/publications/beyondblooms>. In addition we implemented an GUI based framework for community dynamics using network analysis called otter⁸⁶.

Received: 2 May 2024; Accepted: 11 October 2024;

Published online: 29 October 2024

References

1. Meredith, M. et al. Chapter 3 Polar regions. In *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate* (eds Pörtner, H.-O. et al.) 203–320 (Cambridge University Press, Cambridge, UK and New York, NY, USA, 2019).
2. Sala, O. E. et al. Global biodiversity scenarios for the year 2100. *Science* **287**, 1770–1774 (2000).
3. Kraft, A. et al. First evidence of reproductive success in a southern invader indicates possible community shifts among arctic zooplankton. *Mar. Ecol. Prog. Ser.* **493**, 291–296 (2013).
4. Eppey, R. W. & Peterson, B. J. Particulate organic matter flux and planktonic new production in the deep ocean. *Nature* **282**, 677–680 (1979).

5. Lin, I. et al. New evidence for enhanced ocean primary production triggered by tropical cyclone. *Geophys. Res. Lett.* **30**, <https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2003GL017141> (2003).
6. Bluhm, B. A., Swadling, K. M. & Gradinger, R. Sea ice as a habitat for macrograzers. In *Sea Ice*. 394–414 (Wiley Online Books; by John Wiley & Sons, Ltd, 2017).
7. Arrigo, K. R., van Dijken, G. L. & Strong, A. L. Environmental controls of marine productivity hot spots around Antarctica. *J. Geophys. Res. Oceans* **120**, 5545–5565 (2015).
8. Lewis, K., Van Dijken, G. & Arrigo, K. R. Changes in phytoplankton concentration now drive increased arctic ocean primary production. *Science* **369**, 198–202 (2020).
9. Nöthig, E.-M. et al. Summertime chlorophyll a and particulate organic carbon standing stocks in surface waters of the Fram Strait and the arctic ocean (1991–2015). *Front. Mar. Sci.* **7**, 350 (2020).
10. Gosselin, M., Levasseur, M., Wheeler, P. A., Horner, R. A. & Booth, B. C. New measurements of phytoplankton and ice algal production in the arctic ocean. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **44**, 1623–1644 (1997).
11. Fernández-Méndez, M. et al. Photosynthetic production in the central arctic ocean during the record sea-ice minimum in 2012. *Biogeosciences* **12**, 3525–3549 (2015).
12. van Leeuwe, M. A. et al. On the phenology and seeding potential of sea-ice microalgal species. *Elem. Sci. Anth* **10**, 00029 (2022).
13. Clement Kinney, J. et al. Hidden production: on the importance of pelagic phytoplankton blooms beneath arctic sea ice. *J. Geophys. Res. Oceans* **125**, e2020JC016211 (2020).
14. Strass, V. H. & Nöthig, E. M. Seasonal shifts in ice edge phytoplankton blooms in the Barents sea related to the water column stability. *Polar Biol.* **16**, 409–422 (1996).
15. Fortier, M., Fortier, L., Michel, C. & Legendre, L. Climatic and biological forcing of the vertical flux of biogenic particles under seasonal arctic sea ice. *Mar. Ecol. Prog. Ser.* **225**, 1–16 (2002).
16. Leu, E., Søreide, J., Hessen, D., Falk-Petersen, S. & Berge, J. Consequences of changing sea-ice cover for primary and secondary producers in the European arctic shelf seas: timing, quantity, and quality. *Prog. Oceanogr.* **90**, 18–32 (2011).
17. Assmy, P. et al. Leads in arctic pack ice enable early phytoplankton blooms below snow-covered sea ice. *Sci. Rep.* **7**, 40850 (2017).
18. Hardge, K. et al. The importance of sea ice for exchange of habitat-specific protist communities in the central arctic ocean. *J. Mar. Syst.* **165**, 124–138 (2017).
19. Zhao, Q. et al. Relationships of temperature and biodiversity with stability of natural aquatic food webs. *Nat. Commun.* **14**, 3507 (2023).
20. Loreau, M. & De Mazancourt, C. Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecol. Lett.* **16**, 106–115 (2013).
21. Bluhm, B. A., Gradinger, R. & Hopcroft, R. R. Arctic ocean diversity: synthesis. *Mar. Biodivers.* **41**, 1–4 (2011).
22. Frey, K. Arctic ocean primary productivity. In *Arctic Report Card 2017*, NOAA. <https://arctic.noaa.gov/report-card/report-card-2017/arctic-ocean-primary-productivity/> (2017).
23. Barber, D. G. et al. Selected physical, biological and biogeochemical implications of a rapidly changing arctic marginal ice zone. *Prog. Oceanogr.* **139**, 122–150 (2015).
24. Richter-Menge, J. A. & Farrell, S. L. Arctic sea ice conditions in spring 2009–2013 prior to melt. *Geophys. Res. Lett.* **40**, 5888–5893 (2013).
25. 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).
26. Ma, B. et al. Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. *ISME J.* **10**, 1891–1901 (2016).
27. Ives, A. R. & Carpenter, S. R. Stability and diversity of ecosystems. *Science* **317**, 58–62 (2007).
28. Sugihara, G. et al. Detecting causality in complex ecosystems. *Science* **338**, 496–500 (2012).
29. Suzuki, K., Nakaoka, S., Fukuda, S. & Masuya, H. Energy landscape analysis of ecological communities elucidates the phase space of community assembly dynamics. *Ecol. Monogr.* **91**, e01469 (2021).
30. Suzuki, K., Nakaoka, S., Fukuda, S. & Masuya, H. Energy landscape analysis elucidates the multistability of ecological communities across environmental gradients. *Ecol. Monogr.* **91**, e01469 (2021).
31. Soltwedel, T. et al. Natural variability or anthropogenically-induced variation? insights from 15 years of multidisciplinary observations at the arctic marine LTER site HAUSGARTEN. *Ecol. Indic.* **65**, 89–102 (2016).
32. Metfies, K. et al. Protist communities in moored long-term sediment traps (Fram Strait, Arctic)–preservation with mercury chloride allows for PCR-based molecular genetic analyses. *Front. Mar. Sci.* **4**, 301 (2017).
33. 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).
34. 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).
35. Cardozo-Mino, M. G. et al. A decade of microbial community dynamics on sinking particles during high carbon export events in the eastern Fram Strait. *Front. Mar. Sci.* **10**, 1173384 (2023).
36. Wietz, M. et al. The arctic summer microbiome across fram strait: depth, longitude, and substrate concentrations structure microbial diversity in the euphotic zone. *Environ. Microbiol.* **26**, e16568 (2024).
37. 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).
38. 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).
39. Mann, D. G. et al. Bacillariophyta. In *Handbook of the Protists* (eds Archibald, J., Simpson, A. & Slamovits, C.) Vol. 7, 205–266 (Handbook of the Protists. Springer, Cham., 2017).
40. Suter, E. A., Pachiadaki, M., Taylor, G. T. & Edgcomb, V. P. Eukaryotic parasites are integral to a productive microbial food web in oxygen-depleted waters. *Front. Microbiol.* **12**, 764605 (2022).
41. Lin, Y.-C. et al. How communities of marine stramenopiles varied with environmental and biological variables in the subtropical northwestern pacific ocean. *Microb. Ecol.* **83**, 916–928 (2022).
42. Thaler, M. & Lovejoy, C. Environmental selection of marine stramenopile clades in the arctic ocean and coastal waters. *Polar Biol.* **37**, 347–357 (2014).
43. Von Quillfeldt, C. Common diatom species in arctic spring blooms: their distribution and abundance. *Bot. Mar.* **43**, 499–516 (2000).
44. Cleary, A. C., Durbin, E. G., Rynearson, T. A. & Bailey, J. Feeding by pseudocalanus copepods in the Bering sea: trophic linkages and a potential mechanism of niche partitioning. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **134**, 181–189 (2016).
45. Yang, E. J., Ha, H. K. & Kang, S.-H. Microzooplankton community structure and grazing impact on major phytoplankton in the Chukchi sea and the Western Canada basin, Arctic ocean. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **120**, 91–102 (2015).
46. Tragin, M. & Vulot, D. Green microalgae in marine coastal waters: the ocean sampling day (OSD) dataset. *Sci. Rep.* **8**, 14020 (2018).
47. Maat, D. S., Crawford, K. J., Timmermans, K. R. & Brussaard, C. P. Elevated co2 and phosphate limitation favor micromonas pusilla through stimulated growth and reduced viral impact. *Appl. Environ. Microbiol.* **80**, 3119–3127 (2014).
48. Stephens, C. R. et al. Using biotic interaction networks for prediction in biodiversity and emerging diseases. *PLoS ONE* **4**, e5725 (2009).

49. Massana, R., Terrado, R., Forn, I., Lovejoy, C. & Pedrós-Alió, C. Distribution and abundance of uncultured heterotrophic flagellates in the world oceans. *Environ. Microbiol.* **8**, 1515–1522 (2006).
50. de Sousa, M. I. L. *Biogeography of Arctic Eukaryotic Microbiome: A Comparative Approach Between 18S rRNA Gene Metabarcoding and Microscopic Analysis* (Universidade do Porto, 2020).
51. Decelle, J., Romac, S., Sasaki, E., Not, F. & Mahe, F. Intracellular diversity of the v4 and v9 regions of the 18s rna in marine protists (radiolarians) assessed by high-throughput sequencing. *PLoS ONE* **9**, e104297 (2014).
52. Liu, Q., Zhao, Q., McMin, A., Yang, E. J. & Jiang, Y. Planktonic microbial eukaryotes in polar surface waters: recent advances in high-throughput sequencing. *Mar. Life Sci. Technol.* **3**, 94–102 (2021).
53. Wu, S., Xiong, J. & Yu, Y. Taxonomic resolutions based on 18S rRNA genes: a case study of subclass copepoda. *PLoS ONE* **10**, e0131498 (2015).
54. Catlett, D. et al. Evaluation of accuracy and precision in an amplicon sequencing workflow for marine protist communities. *Limnol. Oceanogr. Methods* **18**, 20–40 (2020).
55. Lima-Mendez, G. et al. Determinants of community structure in the global plankton interactome. *Science* **348**, 1262073 (2015).
56. Ma, B. et al. Earth microbial co-occurrence network reveals interconnection pattern across microbiomes. *Microbiome* **8**, 1–12 (2020).
57. Dunne, J. A., Williams, R. J. & Martinez, N. D. Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecol. Lett.* **5**, 558–567 (2002).
58. Jassey, V. E. et al. An unexpected role for mixotrophs in the response of peatland carbon cycling to climate warming. *Sci. Rep.* **5**, 16931 (2015).
59. Bruhn, C. S., Lundholm, N., Hansen, P. J., Wohlrab, S. & John, U. Transition from a mixotrophic/heterotrophic protist community during the dark winter to a photoautotrophic spring community in surface waters of Disko Bay, Greenland. *Frontiers in Microbiology* **15**, 1407888 (2024).
60. Mitra, A. et al. The role of mixotrophic protists in the biological carbon pump. *Biogeosciences* **11**, 995–1005 (2014).
61. Zhao, H. et al. Vertically exported phytoplankton (< 20 µm) and their correlation network with bacterioplankton along a deep-sea seamount. *Front. Mar. Sci.* **9**, 862494 (2022).
62. Gilbert, J. A. et al. Defining seasonal marine microbial community dynamics. *ISME J.* **6**, 298–308 (2012).
63. Kang, S.-H. & Fryxell, G. A. *Fragilariopsis cylindrus* (Grunow) Krieger: the most abundant diatom in water column assemblages of Antarctic marginal ice-edge zones. *Polar Biol.* **12**, 609–627 (1992).
64. Sergeeva, V., Zhitina, L., Mosharov, S., Nedospasov, A. & Polukhin, A. Phytoplankton community structure in the polar front of the eastern barents sea at the end of the growth season. *Oceanology* **58**, 700–709 (2018).
65. Duan, Y. et al. Bacterial community dynamics and co-occurrence network patterns during different stages of biochar-driven composting. *Bioresour. Technol.* **384**, 129358 (2023).
66. Machné, R., Murray, D. B. & Stadler, P. F. Similarity-based segmentation of multi-dimensional signals. *Sci. Rep.* **7**, 12355 (2017).
67. Laporte, M. et al. Proper environmental DNA metabarcoding data transformation reveals temporal stability of fish communities in a dendritic river system. *Environ. DNA* **3**, 1007–1022 (2021).
68. Rao, C. R. *A Review of Canonical Coordinates and an Alternative to Correspondence Analysis Using Hellinger Distance* (Qüestió: quaderns d'estadística i investigació operativa, 1995).
69. Legendre, P. & Gallagher, E. D. Ecologically meaningful transformations for ordination of species data. *Oecologia* **129**, 271–280 (2001).
70. Mikhailov, I. S. et al. Co-occurrence networks among bacteria and microbial eukaryotes of lake Baikal during a spring phytoplankton bloom. *Microb. Ecol.* **77**, 96–109 (2019).
71. Ushio, M. Interaction capacity as a potential driver of community diversity. *Proc. R. Soc. B* **289**, 20212690 (2022).
72. Fujita, H. et al. Alternative stable states, nonlinear behavior, and predictability of microbiome dynamics. *Microbiome* **11**, 63 (2023).
73. Soltwedel, T. et al. Hausgarten: multidisciplinary investigations at a deep-sea, long-term observatory in the arctic ocean. *Oceanography* **18**, 46–61 (2005).
74. Wietz, M. et al. Impact of preservation method and storage period on ribosomal metabarcoding of marine microbes: implications for remote automated samplings. *Front. Microbiol.* **13**, 999925 (2022).
75. Weiß, J. F. et al. Unprecedented insights into extents of biological responses to physical forcing in an arctic sub-mesoscale filament by combining high-resolution measurement approaches. *Sci. Rep.* **14**, 8192 (2024).
76. von Appen, W.-J. et al. Observations of a submesoscale cyclonic filament in the marginal ice zone. *Geophys. Res. Lett.* **45**, 6141–6149 (2018).
77. Spreen, G., Kaleschke, L. & Heygster, G. Sea ice remote sensing using AMSR-E 89-GHz channels. *J. Geophys. Res. Oceans* **113**, <https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2005JC003384> (2008).
78. Cooley, J. W. & Tukey, J. W. An algorithm for the machine calculation of complex Fourier series. *Math. Comput.* **19**, 297–301 (1965).
79. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
80. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software* **3**, 861 (2018).
81. Javier, P. J. E. A., Liponhay, M. P., Dajac, C. V. G. & Monterola, C. P. Causal network inference in a dam system and its implications on feature selection for machine learning forecasting. *Phys. A Stat. Mech. Appl.* **604**, 127893 (2022).
82. Veilleux, B. An analysis of the predatory interaction between paramecium and didinium. *J. Anim. Ecol.* **48**, 787–803 (1979).
83. Azaele, S., Muneepeerakul, R., Rinaldo, A. & Rodriguez-Iturbe, I. Inferring plant ecosystem organization from species occurrences. *J. Theor. Biol.* **262**, 323–329 (2010).
84. Guimera, R. & Nunes Amaral, L. A. Functional cartography of complex metabolic networks. *Nature* **433**, 895–900 (2005).
85. Joyce, K. E., Laurienti, P. J., Burdette, J. H. & Hayasaka, S. A new measure of centrality for brain networks. *PLoS ONE* **5**, e12200 (2010).
86. Kronberg, R. M. & Oldenburg, E. otter—a framework for analysing community dynamics based on time series using co-occurrence, convergent cross mapping networks and Louvain clustering. <https://doi.org/10.5281/zenodo.13840702> (2024).

Acknowledgements

Computational infrastructure and support were provided by the Centre for Information and Media Technology at Heinrich Heine University Düsseldorf. Thanks to Kenta Suzuki for supporting us in adapting rELA to our needs. Thanks to Prince Joseph Erneszer Javier for developing CCM package in python. Thanks to Taylor Priest for providing the PAR data. We thank Theresa Hargesheimer, Jana Bäger, Jakob Barz, Anja Batzke and Daniel Scholz for the RAS support. Moreover we thank the captains and crews of RV Polarstern for excellent support at sea, and the chief scientists for leading the various expeditions conducted for this study. Ship time for RV Polarstern was provided under grants AWI_PS99_00, AWI_PS100_01, AWI_PS107_05, AWI_PS114_01, AWI_PS121_01 of RV Polarstern. The

work was supported by the infrastructure project FRAM (Frontiers in Arctic Marine Monitoring) funded by the Helmholtz Association. Our special thanks go to Stefan Neuhaus for bioinformatic support, Kerstin Korte and Swantje Ziemann for excellent technical support in the laboratory as well as Martina Löbl for coordination of the FRAM-project. We are grateful to three reviewers for their valuable input and time dedicated to improving our manuscript. Their insights have been instrumental in enhancing its quality. This work was supported by The Deutsche Forschungsgemeinschaft (DFG) under grant number EB 418/6-1 (From Dusk till Dawn) and under Germany's Excellence Strategy—EXC-2048/1—project ID 390686111 (CEPLAS) (E.O. and O.E.). Additional funding came from the Helmholtz DataHub Information Infrastructure funds within project iLOVE (E.O., R.M.K., C.B., O.P., M.W.).

Author contributions

W.J.v.A., C.B., M.W. and K.M. are responsible for the sampling design. K.M. contributed 18S metabarcoding data. E.O. and O.P. processed raw metabarcoding reads into ASVs. W.J.v.A. contributed oceanographic data. E.O. devised the project, the main conceptual ideas and study outline. E.O. and R.M.K. designed the model and the computational framework and analyzed the data. E.O., R.M.K., O.P. and K.M. wrote the initial draft. R.M.K. carried out the implementation supported by E.O. E.O., R.M.K. and O.P. interpreted the data and conceptualized the manuscript. O.E. and O.P. advised data evaluation and data interpretation and revised and finalized the manuscript. All authors contributed to the final manuscript, the scientific interpretation and the discussion of results.

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/s43247-024-01782-0>.

Correspondence and requests for materials should be addressed to Ellen Oldenburg.

Peer review information *Communications Earth & Environment* thanks Hideyuki Doi and the other, anonymous, reviewer for their contribution to the peer review of this work. Primary Handling Editor: Alice Drinkwater. 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) 2024