Heinrich Heine
University
Düsseldorf

Exploring Arctic Marine Ecosystems: Analysing Multi-Year Data of Arctic Microbial Marine Communities

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Ellen Oldenburg
geboren in
Kleve, Deutschland

Düsseldorf, Juni 2024

Quantitative und Theoretische Biologie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Oliver Ebenhöh
2. Prof. Dr. Björn Usadel
3. Prof. Dr. Ulf Karsten

Tag der mündlichen Prüfung: 16.12.2024

"The world cannot live without the Arctic; it affects every living thing on Earth and acts as a virtual thermostat, reflecting sunlight and cooling the planet. "

Philippe Cousteau, Jr.

Abstract

The Arctic Ocean is undergoing rapid and significant changes due to atmospheric and oceanic warming. The reduction in sea ice cover has raised questions about the ecological impact on biodiversity, primary productivity, and the biological carbon pump. Furthermore, it has led to environmental changes such as atlantification and a rapidly changing marginal ice zone. The diversity and composition of the phytoplankton community, particularly in the sea ice and water column of the Central Arctic Ocean, will be impacted. As the primary food source for several trophic levels, phytoplankton plays a crucial role in the overall productivity and functioning of the Arctic ecosystem. Any changes in the community structure could have significant implications for other trophic levels and carbon sequestration. Arctic Ocean communities play a significant role in the biogeochemical cycle, due carbon sequestration and nutrient cycling. Yet they have received little academic attention. It is clear that their diversity and function are important areas for further research. The objective of this research is to identify which microorganisms are capable of adapting to the rapidly changing Arctic environment and the forthcoming Atlantification. By investigating the diversity and interactions of microbial communities throughout the year, we aim to elucidate how shifting environmental conditions are altering these ecosystems. Utilizing cutting-edge bioinformatics and mathematical tools, this thesis advances our understanding of marine biology. Considering all these points together, a novel framework was developed for the analysis of temporal patterns and interactions of organisms, with a focus on identifying those groups that exhibit stability under various conditions. In addition, a novel tool for precise and efficient analysis of zooplankton images was developed, which will be crucial for further community analyses in our framework. This holistic approach promises to significantly improve our understanding of the vibrant ecosystem of the Arctic Ocean and its response to ongoing environmental change.

Zusammenfassung

Der Arktische Ozean ist aufgrund der Erwärmung der Atmosphäre und der Ozeane raschen und bedeutenden Veränderungen unterworfen. Der Rückgang der Meereisbedeckung hat Fragen zu den ökologischen Auswirkungen auf die biologische Vielfalt, die Primärproduktion und die biologische Kohlenstoffpumpe aufgeworfen. Außerdem hat er zu Umweltveränderungen wie der Verlagerung einer sich rasch verändernden Randeiszone geführt. Die Vielfalt und Zusammensetzung der Phytoplankton-Gemeinschaft, insbesondere im Meereis und in der Wassersäule des zentralen Arktischen Ozeans, werden davon betroffen sein. Als primäre Nahrungsquelle für mehrere trophische Ebenen spielt das Phytoplankton eine entscheidende Rolle für die Gesamtproduktivität und das Funktionieren des arktischen Ökosystems. Jegliche Veränderungen in der Struktur der Gemeinschaft könnten erhebliche Auswirkungen auf andere trophische Ebenen und die Kohlenstoffspeicherung haben. Die Lebensgemeinschaften im Arktischen Ozean spielen eine wichtige Rolle im biogeochemischen Kreislauf, da sie für die Kohlenstoffbindung und den Nährstoffkreislauf verantwortlich sind. Dennoch haben sie bisher nur wenig wissenschaftliche Aufmerksamkeit erhalten. Es ist unbestritten, dass die Vielfalt und Funktion der Mikroorganismen wichtige Forschungsbereiche darstellen. Die vorliegende Forschung zielt darauf ab, die Fähigkeit von Mikroorganismen zu untersuchen, sich an die sich rasch verändernde arktische Umwelt und die bevorstehende Atlantisierung anzupassen. Durch die Untersuchung der Diversität und der Interaktionen mikrobieller Gemeinschaften über das gesamte Jahr hinweg soll eruiert werden, wie sich diese Ökosysteme durch die veränderten Umweltbedingungen verändern. Die Anwendung modernster Bioinformatik und mathematischer Methoden trägt zu einer umfassenderen Betrachtung der Meeresbiologie bei. Unter Berücksichtigung der genannten Punkte wurde ein neuartiger Rahmen für die Analyse zeitlicher Muster und Interaktionen von Organismen entwickelt. Dabei liegt der Schwerpunkt auf der Identifizierung derjenigen Gruppen, die unter verschiedenen Bedingungen Stabilität zeigen. Des Weiteren wurde ein neuartiges Werkzeug zur präzisen und effizienten Analyse von Zooplanktonbildern entwickelt, das für weitere Gemeinschaftsanalysen in unserem Rahmenwerk von Bedeutung sein wird. Der ganzheitliche Ansatz verspricht, unser Verständnis des lebendigen Ökosystems des Arktischen Ozeans und seiner Reaktion auf die laufenden Umweltveränderungen zu verbessern.

Acknowledgements

First of all, I would like to thank my supervisor, Prof. Oliver Ebenhöh, who has been an amazing mentor and boss to me over the last few years. Thank you for giving me the opportunity to take part in this incredible project and encouraging me to try my hand in many areas. Thank you for listening to all my ideas and sparking my interest in marine biology. It was such a pleasure to work with you! I'd also like to thank Dr. Katja Metfies for giving me the chance to gain so much valuable experience during my research expedition. It was such a great opportunity for me to gain a deeper understanding of Arctic marine ecosystems.

I would like to thank the German Research Foundation, the Cluster of Excellence for Plant Sciences, the CEPLAS Graduate School, the Graduate School POLMAR and iGrad for funding my project and for allowing me to broaden my scientific experience at international conferences, research stays and workshops.

I would like to thank you, Barbara, Matthias, Wilken and Taylor, my colleagues from Alfred Wegener Institut for the many fruitful discussions during the meetings and the numerous insights and perspectives into the world of ecology.

Many, many thanks to all current and former members of the QTB team. You are undoubtedly a unique group and the atmosphere that prevails in this institute is incomparable. Thank you Thomas, you were the best office colleague imaginable. Hettie, Mo, Yvan and Chilperic, for the funny conversations. And thank you, Adélaïde, for training me in the right way to present and for all your delicious desserts. I would also like to thank Anna for her support and advice over the years. Thank you Tim for all our great chats, that helped me to see what's really important. Janina and Dan, thank you for fixing my computer problems and always supporting me with your technical knowledge for the modules. Ovidiu, thank you for all your help, for being a partner in crime and for sharing the most amazing experiences with me during the expedition. I would also like to thank Marvin for his help with all the big and small questions and the lovely breakfast chats in the morning. And a big thank you to Nima, who always welcomed me into his office and made my life so much easier with his lovely manner. Mara, you've been such a help with all the admin and the bureaucracy. Thank you for your open ear and your support.

Many thanks to Sieglinde for always having an open ear for me and for your advice. Raphael, thank you for your support, expertise, for listening to all my ideas and for still getting involved in the world of ecology.

And finally, I would like to thank my family and friends who have supported and encouraged me throughout my studies. Without you, I wouldn't be the person I am today.

Contents

Chapter 1

Introduction and Motivation

1.1 The landscape of the Arctic Ocean

The Arctic Ocean, the smallest and northernmost of the world's five oceans, covers only 4% of the total ocean area. It is characterized by two major deep basins, namely the Eurasian and the Amerasian Basin, separated by the Lomonosov Ridge, and surrounded by extensive continental shelves (Jakobsson, 2002; Slagstad et al., 2015). With an average water depth of around 1,200 m, the Arctic Ocean is the shallowest of all oceans, largely due to its shelves, collectively comprising over 50% of its total area (Jakobsson et al., 2003, 2004, 2014).

Its distinctive geographical location and landlocked nature gives rise to several unique characteristics that profoundly influence its physicochemical environment and biological dynamics. The Arctic Ocean receives a significant amount of freshwater and organic matter from some of the planet's largest river systems, such as the Lena River, which accounts for approximately 10% of the global river discharge (Dittmar and Kattner, 2003). This freshwater inflow results in strong stratification within the Arctic Ocean, with a less dense fresh upper ocean layer overlaying denser, more saline deeper water. River discharge plays a pivotal role in regulating mixing patterns, nutrient distribution, heat exchange, and sea-ice formation within the region (Carmack et al., 2016; Osadchiev et al., 2024).

The pronounced stratification of the Arctic Ocean serves as a driving force for its interaction with the Pacific and Atlantic Oceans (Rudels et al., 2013; Rudels and Friedrich, 2000). Water from the Pacific enters the Arctic via the Bering Strait (Woodgate et al., 2012), but the most significant water mass exchange occurs with the North Atlantic through Fram Strait and the Barents Sea (Figure 1.1) (Rudels et al., 2013; Rudels et al., 2004; Wefing et al., 2021). The Fram Strait is a passage, that acts as a gateway between the Arctic Ocean and the Atlantic Ocean (Pawlowicz et al., 1995). Long-term time-series studies in the Arctic marine ecosystem provide a baseline for understanding ongoing changes in the Fram Strait. These studies significantly contribute to global efforts to comprehend variations in ecosystem structure and functioning, aiding in future predictions under different climate scenarios (Soltwedel et al., 2016). The Long-Term Ecological Research (LTER) site, HAUSGARTEN (Soltwedel et al., 2005), encompasses several stations spanning the West Spitzbergen Current (WSC) and East Greenland Current (EGC). The WSC is primarily characterized by Atlantic conditions with

minimal polar influences, whereas the EGC experiences typical polar conditions and regular sea-ice cover. The transition zone, represented by station HGIV, situated between the WSC and EGC, is subject to influences from both Arctic and Atlantic conditions, frequently experiencing intrusion of Polar Water (PW) and ice coverage throughout the annual cycle (Soltwedel et al., 2013). Notably, the EGC area demonstrates a more pronounced polar influence.

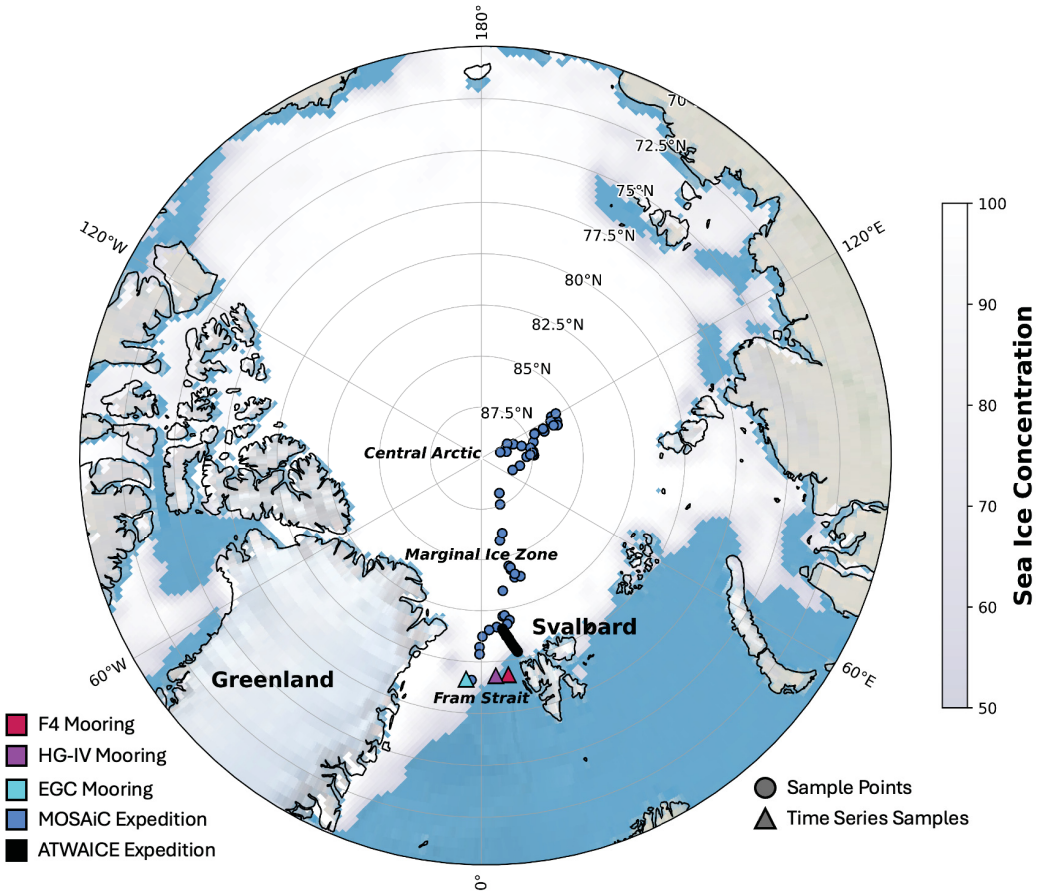

Figure 1.1: **Expedition tracks: Sampling locations in the Arctic Ocean** The map shows the Arctic Ocean. Blue dots are representing samples from the MOSAiC Expedition (Mock et al., 2022) and the black ones from the ATWAICE Expedition (Kanzow, 2021). The red, purple, and cyan dots are representing the mooring stations F4, HG-IV, and EGC. These stations are equipped with autonomous samplers, which collect biological samples and measure environmental data over several years. Sea ice concentration shown for illustrative purposes only (Based on data from PolarWatch ERDDAP).

The Arctic Ocean, reigning supreme north of the Arctic Circle (66°34' N), is characterized by its extensive sea ice cover. This icy blanket acts as a critical regulator for the entire Arctic ecosystem. It controls how much heat escapes to the atmosphere by reflecting sunlight

(albedo) and influences the amount of light reaching the water below (Lindsey, 2009; Perovich and Polashenski, 2012). As sea ice freezes and melts, it dictates the temperature and salinity of the surface ocean, which in turn has a domino effect on ocean layering (stratification), mixing patterns, and nutrient availability for marine life (Korhonen et al., 2013; Zhang et al., 2023). Beyond its physical influence, sea ice provides a vital habitat for a diverse array of organisms, offering shelter for those living on the ice surface and those thriving in the hidden world beneath (Arrigo, 2014, 2017; Bluhm and Gradinger, 2008).

The Arctic system experiences a pronounced seasonality, characterized by an extended winter period with several months of polar night, followed by spring, summer, and fall months with continuous daylight supporting autotrophic life in the sea ice and upper ocean (Wassmann, 2011). Arctic sea ice reaches its maximum seasonal extent in March after growing throughout the winter, and undergoes melting and breakup in spring and summer, reaching its minimum extent in September (Perovich et al., 2020; Serreze and Barry, 2011; Stroeve and Notz, 2018). As a result, freshwater input to the system is lowest in winter, when river discharge is limited by freezing and stratification is weakened, allowing for deep mixing and nutrient replenishment (Korhonen et al., 2013; Meredith et al., 2019).

In the context of an anticipated transition to an ice-free Arctic, projections indicate that the earliest instance of completely ice-free conditions during September could occur by the 2020s to 2030s across all emission scenarios, with a high likelihood by 2050. On average, daily September ice-free conditions are expected to occur about four years earlier, possibly surpassing monthly metrics by up to a decade (Jahn et al., 2024; Kinnard et al., 2011). Consistently ice-free conditions during September are projected to become regular by mid-century, ranging from 2035 to 2067, depending on emission trajectories. The IPCC's assessment report suggests that under intermediate and high greenhouse gas emissions scenarios, the Arctic may be nearly ice-free in September by mid-century (Masson-Delmotte et al., 2021b). A study employing an attribution analysis approach highlights the significant influence of greenhouse gas emissions on Arctic sea ice area, with forecasts indicating ice-free Arctic conditions by mid-century across all scenarios. This underscores the urgent need to prepare for, and adapt to a seasonally ice-free Arctic (Kim et al., 2023; Stroeve et al., 2007). The diminishing extent of Arctic sea ice has profound environmental and economic implications for global climate change trajectories. Statistical forecasts based on extensive satellite data suggest an accelerating decline in sea ice coverage, diverging from Coupled Model Intercomparison Project (CMIP5) model projections (Driscoll et al., 2012). Long-term statistical projections indicate a nearly 60% likelihood of an essentially ice-free Arctic Ocean by the 2030s, earlier than average global climate model projections (Diebold and Rudebusch, 2022). Efforts within the Paris Agreement framework aim to stabilize global temperatures, but current emission reduction policies may limit warming to around 3.0°C by 2100. The study suggests that transitioning from 2.0°C to 1.5°C warming could significantly reduce ice-free conditions, while at 3.0°C warming, permanent summer ice-free conditions are likely, emphasizing the urgency of stronger commitments to the Paris Agreement (Bamber, 2022; Diebold and Rudebusch, 2022).

1.2 Ocean warming: A changing landscape

The Earth's natural atmospheric greenhouse effect has undergone changes. Since 1750, the atmospheric concentrations of greenhouse gases such as carbon dioxide, methane, and nitrous oxide have increased, leading to the absorption of outgoing heat and subsequent warming of the Earth's surface (Crowley, 2000; Hartmann et al., 2013; Katelaris and Beggs, 2018). The global mean surface temperature has increased by approximately (Khilyuk, 2003) 0.85°C over the late 19th century. The period from 1983 to 2012 was the warmest in the past 800 years (Masson-Delmotte et al., 2013). This rise in surface temperatures is also evident in the world's oceans, with notable climate-related changes occurring in the upper 700 m of the Northern Hemisphere between 1971 and 2010 (Figure Section 1.8 (a))(Howes et al., 2015; Rhein et al., 2013).

Furthermore, the heat flux of Atlantic Water (AW) through the Fram Strait into the Central Arctic Ocean (CAO) has increased significantly since 1997 due to elevated surface water temperatures (0.06°C per year, mean temperature from 1997 to 2010) and intensified water flow (Beszczynska-Möller et al., 2012; Karam et al., 2024; Long et al., 2024; Schauer et al., 2008; Schauer et al., 2004). Poleward ocean heat transport plays a critical role in the Earth's system. Over the past five decades, the Arctic Ocean, encompassing the Nordic and Barents Seas, has experienced significant warming, aligning with amplified ocean heat transport and sea ice reduction. This warming trend has notably contributed to the recession of marine-terminating glaciers in Greenland. In the Nordic Seas, where approximately 60% of total heat loss to the atmosphere occurs, variability is primarily influenced by the frequency of Cold Air Outbreaks and cyclones, yet no long-term statistically significant trend has been observed. Conversely, heat loss from the Barents Sea (approximately 30%) and the more northern Arctic seas (approximately 10%) exhibits considerable positive trends. Since 1900, there has been a notable increase in Atlantic Water (AW) inflow, total heat loss to the atmosphere, and dense outflow, with these parameters consistently linked through theoretical scaling. Additionally, the increase in AW inflow is influenced by wind dynamics. Moreover, Arctic Ocean Carbon Dioxide (CO_2) uptake has surged by approximately 30% over the past century, aligning with the loss of Arctic sea ice, which enhances air-sea interaction. This uptick represents roughly 8% of the global CO_2 uptake, indicating the Arctic's increasing significance in the global carbon cycle (Smedsrud et al., 2022; Tesi et al., 2021). The Arctic Ocean is a dynamic and complex ecosystem, in which the biological pump plays a crucial role in carbon cycling (Figure Figure 1.3). Primary producers within the sea ice and the upper water column utilise solar energy to convert carbon dioxide (CO_2) into biomass, initiating a cascade of ecological processes. This conversion results in the generation of a diverse range of dissolved and particulate organic matter, a high proportion of which is respired by heterotrophic organisms that inhabit the sea ice and surface waters. Heterotrophic microbes, predominantly bacteria, decompose and remineralise the organic matter, thereby recycling essential inorganic nutrients that support ongoing primary production. In addition, the viral shunt mechanism releases organic matter into the surrounding seawater or sea ice through viral-induced cell lysis. The microbial loop, a process whereby microbial activity mediates the cycling and transformation of carbon, pre-

dominantly occurs in the surface ocean, where the majority of organic material is consumed and respired. As sinking material descends, microbial degradation and grazing further reduce the export flux and modify the organic matter. Zooplankton consume and respire a portion of this organic matter, transferring it to higher trophic levels. Only a small proportion of surface production reaches the deep-sea floor, supporting benthic heterotrophic organisms. Their metabolic activity results in the remineralisation of the majority of organic material, with only a negligible proportion (0.1%) undergoing burial and contributing to long-term carbon sequestration from the atmosphere (Jahnke and Jackson, 1992). The heightened heat flux, combined with atmospheric warming, has had a negative impact on sea ice conditions in the CAO (Cao et al., 2018; Comiso, 2003; Comiso et al., 2008; Stroeve et al., 2008; Stroeve et al., 2012b). It is projected that there may be a nearly ice-free summer period between 2020 and 2050 (Kim et al., 2023; Kirtman et al., 2013), which would have significant ecological consequences for species associated with and adapted to sea ice (see Pagano and Williams, 2021; Post et al., 2013 for a comprehensive review).

1.3 Arctic sea ice: Sea ice decline and its repercussions in the Arctic Ocean

In the last decade, the Central Arctic Ocean has experienced significant sea ice loss due to atmospheric and oceanic warming (Yadav et al., 2020; Zhao et al., 2019; Zhou et al., 2024). Data shows that the average sea ice extent has decreased by approximately 10% per decade (Comiso et al., 2008; Stroeve et al., 2012b). At the same time, the ice has become thinner and fresher, and there has been an increase in the prevalence of melt ponds. The Arctic Ocean has experienced a decrease in the thickness of multi-year ice (MYI, 1.5 - 3 m), with thin first-year ice (FYI, 0.5 - 1.5 m) becoming the dominant ice type (Fetterer and Untersteiner, 1998; Kwok and Rothrock, 2009; Laxon et al., 2013; Ye et al., 2016).

Sea ice conditions in the Arctic Ocean show considerable seasonal and annual variability. The polar day, which starts towards the end of February or early March, marks the beginning of the productive season in the pelagic realm due to increased light intensity (Berge et al., 2015b; Cohen et al., 2020; Ellertsen, 1993; Perrette et al., 2011). The summer minimum sea ice extent is usually reached in August or September, resulting in extensive open water areas (Arrigo et al., 2008). Significant losses of sea ice occurred in September of 2007 and 2012. The latter experienced an extreme and sudden reduction in ice cover in the Beaufort, Chukchi, and East Siberian Seas (NSIDC; `http://nsidc.org`).

The thinning of sea ice in the Central Arctic Ocean, combined with strong southward geostrophic winds, has resulted in an increased export of sea ice through the Fram Strait out of the CAO (Halvorsen et al., 2015; Krumpen et al., 2015; Smedsrud et al., 2011). During autumn, new ice forms as ocean surface temperatures drop below the freezing point of seawater (-1.8°C and a salinity of approximately 34 psu) (Arrigo, 2014; Golden et al., 1998; Krembs et al., 2011; Le Hir et al., 2014; Meredith et al., 2019; Yang et al., 2023). Sea ice extent and concentration peak

during the dark winter period, from late October to late February. Climate-related changes are becoming increasingly evident during the Arctic winter. On March 24, 2016, the Arctic experienced its first recorded seasonal minimum sea ice extent since the winter of 1979. This led to large ice-free areas in the Arctic shelf regions (Perovich et al., 2020; Serreze and Barry, 2011).

The sea ice in the CAO has unique physico-chemical conditions and serves as a vital habitat for unicellular eukaryotes, fish, crustaceans, and higher trophic levels such as seals and polar bears. The anticipated decline in sea ice and subsequent loss of habitat is to have a significant impact on the biodiversity and ecology of highly adapted Arctic species. The sea ice habitats, such as the water column, under-ice water, sea ice, and melt ponds, are home to unicellular eukaryotes that serve as the basis for the marine and benthic food webs.

1.4 The Arctic water column: A haven for diverse microbial life

The Arctic Ocean boasts a unique water column structure, shaped by a constant influx of freshwater from rivers and melting ice (Aagaard et al., 1985; Carmack et al., 2016; Talley et al., 2011). Evidence suggests distinct microbial communities thrive within each of these stratified layers (Hamdan et al., 2013).

The polar surface water, the uppermost layer, extends to roughly 200 meters deep (Jones, 2001; Talley et al., 2011). Nestled beneath it lies the Polar Mixed Layer (PML), with expansion depths of 25-50 m. Characterized by low salinity (27-34 psu), and near-freezing temperatures with significant seasonal variations, the PML plays a crucial role for the carbon cycle (Appen et al., 2021; Talley et al., 2011). A sharp halocline separates the surface layer from the warmer, saltier intermediate Atlantic layer below, restricting deep-to-surface ocean exchange (Aagaard et al., 1985; Rudels, 2012; Talley et al., 2011). Regionally, water column properties vary due to its journey through the Arctic, leading to distinct vertical profiles in the Amerasian and Eurasian basins, as well as extensive shelf regions (Jakobsson et al., 2004; Talley et al., 2011).

Biological production thrives primarily in the surface ocean, further divided into three distinct zones. The first zone boasts a deep euphotic zone (40-50 m), vigorous vertical mixing, and high productivity (Sigman and Hain, 2012). The second zone experiences seasonal ice cover, resulting in stronger stratification, reduced mixing, and a shallower euphotic zone, leading to lower primary productivity. However, occasional phytoplankton blooms can still occur. Areas with permanent ice cover have a very shallow euphotic zone, resulting in minimal productivity and biomass (Laney et al., 2014; Wassmann, 2011; Wassmann et al., 2011). Ice-free or seasonally ice-covered shelf regions typically exhibit higher productivity compared to the perpetually ice-covered deep central basins (Carmack et al., 2016; Holt et al., 2016; Tremblay et al., 2015).

Regardless of location, studies consistently show oligotrophic Alphaproteobacteria, primarily identified as Pelagibacter or members of the SAR11 clade, to be dominant pelagic bacteria in Arctic surface waters (Boeuf et al., 2014; Bowman et al., 2012; Salcher et al., 2011; Wilson et

al., 2017). Algal bloom events result in an increase in the abundance of taxa that thrive in high biomass and nutrient-rich conditions. Members of the Bacteroidetes, Gammaproteobacteria, and specific Alphaproteobacteria are among these taxa (Teeling et al., 2012; Wells et al., 2015; Zhou et al., 2018). Arctic pelagic communities exhibit successional patterns similar to bacterial communities in warmer oceans (Boeuf et al., 2014; Bunse and Pinhassi, 2017; Teeling et al., 2016).

The deeper water column, characterized by light deprivation, higher concentrations of inorganic nutrients, and reduced seasonal variability, offers a more stable habitat (Aristegui et al., 2009; Orcutt et al., 2011).

Here, chemolithotrophic Thaumarchaeota and representatives of the Chloroflexi-type SAR202 clade appear to dominate year-round (Wilson et al., 2017). While some studies suggest minimal seasonal changes in community structure (Kirchman et al., 2010), others report significant differences, including the presence of deep-water groups near the surface during winter (Alonso-Sáez et al., 2008; Wilson et al., 2017).

A question surrounds the exchange of bacteria and microeukaryotes between the Arctic Ocean's water column and sea ice. We don't fully understand how surface seawater communities influence the microbial denizens within the sea ice during winter freeze-up. Similarly, the impact of melting sea ice and the release of its trapped biomass on the composition of summer surface water communities remains unclear. Furthermore, research is needed to identify the specific heterotrophic groups that associate with algal biomass in both the sea ice and the water column. These interactions likely play a crucial role in nutrient recycling within the upper ocean, but their exact nature awaits further exploration. By unraveling these intricate connections, we can gain a deeper understanding of the delicate balance within the Arctic marine ecosystem.

1.5 Sea-ice ecosystem

The Arctic Ocean contains regions with both perennial and seasonal ice coverage. Perennial sea ice, which persists year-round, is mainly found in the deep basins. In contrast, sizable portions of the shelves and neighboring seas experience seasonal ice cover, which is present only during specific times of the year (Polyak et al., 2010). Ice formation primarily occurs over open waters on the shelves during autumn and winter. It is then transported towards the central basins by wind and currents (Polyak et al., 2010). The ice can remain adrift for several years before undergoing significant melting, and the remaining ice eventually exits the Arctic mainly via the Fram Strait (Kwok and Rothrock, 2009; Rudels et al., 2013). FYI and MYI can differ significantly in thickness, albedo, salinity, and brine inclusion (Weeks and Ackley, 1986). During the process of sea ice formation, salt is expelled and accumulates as brine liquids within a network of pores and channels that permeate the entire structure of the ice, from its surface to its base (Petrich and Eicken, 2010; Thomas and Dieckmann, 2002). The concentration of salt within these networks is high enough to prevent the water from freezing even in subzero temperatures. This creates a viable habitat within the ice structure (Petrich and Eicken, 2010;

Thomas and Dieckmann, 2002). Variations in ice morphology, along with pronounced gradients in temperature, light, nutrients, salinity, volume, and pore space contribute to the remarkable spatial heterogeneity of this environment (Arrigo et al., 2008; Deming et al., 2007; Dieckmann and Hellmer, n.d.; Petrich and Eicken, 2010).

The sea-ice matrix supports a diverse range of organisms, including viruses, bacteria, algae, protists, various meiofauna, and small crustaceans (Bluhm et al., 2018; Legendre et al., 1992; Steiner et al., 2021), which form their own ice-associated food web. Photoautotrophic algae serve as the foundation, and diatoms are the primary contributors to biomass. They accumulate in the lower part of the ice, reaching concentrations that can give the ice a brownish tint or form extended filamentous strands that attach to the underside of the ice (Arrigo, 2014; Katlein et al., 2015; Melnikov, 1987; Mock and Gradinger, 1999). As the ice melts, significant amounts of sea ice biomass can be released into the water column and descend to the seafloor (Appen et al., 2021; Boetius et al., 2013; Swoboda et al., 2024; Tamelander et al., 2009). This biomass not only sustains benthic life, but also facilitates vertical dispersal of associated microorganisms, thereby promoting microbial community connectivity.

Sea ice organisms are thought to originate primarily from the surface ocean, where they are recruited during the freezing process, adhere to, or become trapped between newly formed ice crystals, and are subsequently incorporated into the consolidating ice matrix (Ewert and Deming, 2013). However, they can also originate from entrained sediments (Nürnberg et al., 1994; Pfirman et al., 1997; Wegner et al., 2017) or atmospheric deposition (Price et al., 2009). In their new environment, organisms must adapt to harsh abiotic conditions, including changes in available space, light levels, salinity, nutrient concentrations, and extremely low temperatures (Gradinger and Ikävalko, 1998). Many organisms inhabiting sea ice exhibit specific physiological or biochemical adaptations, such as specialized membrane compositions that maintain fluidity at low temperatures (Bayer-Giraldi et al., 2011; Feng et al., 2014; Ramasamy et al., 2023), psychro- and halophilic enzymes (Pomeroy and Wiebe, 2001), or the ability to encyst to endure specific time intervals (Stoecker et al., 1998). These adaptations provide sea ice microorganisms with the necessary mechanisms to survive sudden melting and relocation events, such as transitions from sea ice to the water column.

Bacterial populations within sea ice can reach densities of up to 10^7 cells per milliliter (Bowman, 2013; Gosink et al., 1993), and even higher when associated with high ice algal biomass (Assmy et al., 2013; Deal et al., 2011; Fernández-Méndez et al., 2014).

Within the confined brine channel system, they thrive due to reduced grazing pressure from large metazoan predators (Krembs et al., 2000), close spatial association with ice algae (Deming et al., 2007; Krembs et al., 2000), and elevated concentrations of dissolved organic matter resulting from cell death, lysis, or exudation (Collins and Deming, 2011; Paul et al., 2012; Thomas et al., 2001; Thornton, 2014). Both ice algae and bacteria are prolific producers of extracellular polymeric substances (EPS) (Costa et al., 2018; Decho and Gutierrez, 2017; Krembs et al., 2002), which consist primarily of polysaccharides and glycoproteins (Verdugo et al., 2004). These substances enhance cell surface adhesion and can facilitate the selective incorporation of cells into the ice during freezing (Cai et al., 2020; Gradinger and Ikävalko,

1998), as well as provide potential attachment sites (Junge et al., 2004). EPS also serve as carbon-rich substrates for bacteria in sea ice (Meiners et al., 2004) and are thought to act as cryoprotectants, buffering against abrupt changes in pH or salinity (Collins et al., 2008; Decho and Gutierrez, 2017). Through EPS production, bacteria modify the microstructure of the ice, improving its habitability and permeability (Krembs et al., 2011) and influencing nutrient regeneration processes (Riedel et al., 2007). The majority of existing knowledge about the structure of Arctic sea ice communities is derived from research conducted during the spring and summer seasons. During these seasons, heterotrophic taxa, primarily Flavobacteria and Gammaproteobacteria, flourish due to the copious organic matter produced by algae (Bowman, 2015; Bowman et al., 2012; Eronen-Rasimus et al., 2016). Notably, these spring and summer ice communities differ from source communities in the underlying seawater (Collins et al., 2010; Han et al., 2014; Hatam et al., 2014), prompting questions about selection processes or seasonal succession within sea ice and connectivity between ice and water column communities. Initial findings on winter ice community structure and the early stages of sea ice community formation revealed that the dominant members in winter ice, oligotrophic Alpha- and Gammaproteobacteria, mirrored those in the underlying waters, arguing against selective incorporation of specific bacterial groups (Collins et al., 2010; Comeau et al., 2011). With temperatures in the ice potentially dropping to -35°C (Deming and Eicken, 2007; Deming et al., 2007), gradual selection for psychrotrophic types with the capacity and metabolic traits to survive at these temperatures is conceivable (Feng et al., 2014; Heinz et al., 2018; Junge et al., 2011, 2004). In addition, winter conditions appear to favor species with the ability to adhere to surfaces or particles, such as members of the bacteroidetes (Larose et al., 2013). The sea ice environment exhibits pronounced physical and chemical vertical gradients that evolve over time as the ice matures (Meier et al., 2014; Yamanouchi and Takata, 2020). Consequently, distinct bacterial communities have been observed at different depth levels (Eronen-Rasimus et al., 2016; Hatam et al., 2014), as well as in multi-year ice versus first-year ice (Hatam et al., 2016). Elevated levels of presumed brackish or freshwater groups, such as members of the Actinobacteria and Betaproteobacteria, in the surface layer of ice have been associated with the presence of melt ponds on sea ice (Aparıécio, 2023; Hatam et al., 2014). At the same time, the prevalence of copiotrophic members of Flavobacteria, Alpha-, Gammaproteobacteria, and Verrucomicrobia has often been associated with high algal biomass (Bowman et al., 2012; Collins et al., 2010; Eronen-Rasimus et al., 2016). The availability of algal-derived substrates has been identified as a primary factor shaping sea ice bacterial communities during the productive season (Cowie et al., 2014; Eronen-Rasimus et al., 2015; Hatam et al., 2016), and potentially even during the dark winter months (Eronen-Rasimus et al., 2017; Junge et al., 2004). Heterotrophic bacteria are essential components of the sea ice ecosystem (Figure 1.3), as their metabolic activity releases remineralized inorganic nutrients to support the growth of sea ice algae (Kottmeier and Sullivan, 1990; Roukaerts et al., 2021). Their efficiency in metabolizing organic matter within the ice matrix also influences the amount and composition of materials released into the water column during ice melt (Deming, 2010). Despite their critical role, our understanding of Arctic sea ice bacteria and microeukaryotes remains limited, based primarily on sporadic observations focused primarily on the ice shelves or the Amerasian Basin. Furthermore, detailed

investigations of which types of heterotrophic bacteria are associated with sea ice algal biomass and whether these associations are specific are lacking and therefore largely unknown (Bowman, 2015; Ramanan et al., 2016).

1.6 Microeukaryotic biodiversity

Microeukaryotes are an integral part of the Arctic ecosystem and play an important role in the diversity and productivity of various water and ice environments. Autotrophic species, such as phytoplankton and ice algae (e.g. diatoms, flagellates), are major contributors to primary production and thus influence energy fluxes and carbon sequestration in the Arctic Ocean (Boetius et al., 2013; Fernández-Méndez et al., 2015; Gosselin et al., 1997; Gradinger, 2009; Kudryavtseva et al., 2023).

Phytoplankton and ice algae serve as food sources for heterotrophic zooplankton (e.g. flagellates, dinoflagellates, ciliates). In addition, carbon incorporated into algal aggregates and zooplankton fecal pellets is transported to the Arctic deep sea, where it provides food for benthic communities and is preserved in sediments (Bathmann et al., 1990; Bauerfeind et al., 1994; Kohlbach, 2017).

Diatoms, belonging to the supergroup Stramenopiles, are among the most diverse planktonic organisms in the Arctic and play a crucial role in the ecosystem (Figure 1.2). Diatoms make up a significant portion of both the water column and sea ice ecosystems (Arrigo, 2014). Within these environments, they serve as a vital food source for heterotrophic flagellates or crustaceans (zooplankton). Diatoms vary in size, ranging from 5 to 500 micrometers, with some species exceeding 1 millimeter in length. The silica-based frustule (cell wall) of diatoms provides protection against grazers (Behrenfeld et al., 2021; Passow, 1991).

Under favorable physicochemical conditions in spring or summer, certain diatom species (e.g. *Chaetoceros spp.*, Figure 1.2 , *Thalassiosira spp.*, and *Fragilariopsis spp.*) can initiate blooms in open water, at ice margins, or under sea ice (Ardyna et al., 2020; Arrigo et al., 2012; Ellertsen, 1993; Von Quillfeldt, 2000). Factors that drive bloom formation include increased water stratification, nutrient availability, and light exposure. Following blooms, aggregates composed of chain-forming diatoms (e.g. *Fragilariopsis*, *Melosira*, *Thalassiosira*, Figure 1.2) or highly silicified diatoms (e.g. *Fragilariopsis*, *Coscinodiscus*, *Melosira*) exhibit high sinking rates and contribute significantly to carbon export to the deep seafloor in the Fram Strait and the Arctic Ocean (Bauerfeind et al., 2009; Boetius et al., 2013).

The Arctic Ocean boasts a vibrant tapestry of eukaryotic microorganisms, each with a critical function within the ecosystem. This work zooms in on some of these essential inhabitants. The supergroup Haptophyta includes a collection of autotrophic flagellates (Figure 1.2) typically found in the pico- and nanoplankton size range (3-20 micrometers).

Another member of the Haptophyta supergroup is *Phaeocystis*. Existing as either single-celled organisms or gelatinous colonies, *Phaeocystis* boasts a wide distribution, particularly in the

warm North Atlantic waters of Fram Strait and the Arctic Ocean (Eikrem et al., 2016; Metfies et al., 2016; Reigstad and Wassmann, 2007). Unlike Coccolithophores, *Phaeocystis* plays a less significant role in vertical export of organic carbon. However, their abundance makes them a critical food source for mesozooplankton in the upper ocean (Baumas and Bizic, 2024; Gasparini et al., 2000).

Green algae, or Chlorophyta, encompass a diverse group, including both multicellular and unicellular taxa, found commonly in marine and freshwater phytoplankton communities (Caron et al., 2012). The smallest marine eukaryotes, measuring less than 3 micrometers, belong to the chlorophyte class Mamiellophyceae. Ostreococcus exemplifies this group. Micromonas, another member of Mamiellophyceae, enjoys a wide distribution and forms a significant portion ($\geq 1\%$ abundance) of the Arctic Ocean biosphere (Lovejoy et al., 2007; Yung et al., 2022). Recent observations have identified *Micromonas pusilla* in the Arctic halocline, particularly associated with Atlantic Water (Metfies et al., 2016). Interestingly, *M. pusilla* exhibits high adaptability to various environments through its diverse genetic lineages (Foulon et al., 2008; Worden et al., 2009).

The supergroup Haptophyta encompasses a group of autotrophic flagellates (Figure 1.2) primarily found in the pico- and nanoplankton size range (3-20 micrometers). Certain members, like *Emiliania huxleyi* and *Coccolithus pelagicus*, are known as Coccolithophores. These fascinating organisms produce intricate calcified scales called Coccoliths. During massive blooms, Coccolithophores are even visible from space. Their contribution is significant, as they act as major regional contributors of carbonate and a vital source of calcite carbon (Bates et al., 2013; Brown and Yoder, 1994; Winter and Siesser, 2006).

Dinoflagellates, belonging to the Alveolata supergroup, are arguably the most prominent protozooplankton in the Arctic pelagic realm, ranging in size from 3 to 200 micrometers. Heterotrophic dinoflagellates, like *Akashiwo*, *Prorocentrum*, and *Protoperidinium* (Figure 1.2), boast relatively large cells and play a vital role as grazers of phytoplankton and smaller flagellates. Mixotrophic dinoflagellates, including *Gymnodinium* (Figure 1.2) and *Karlodinium*, possess the unique ability to both photosynthesize and ingest prey within a single cell (Lee et al., 2014; Li et al., 2022; Mitra et al., 2016). Depending on environmental factors, these adaptable organisms can switch between feeding modes through gene regulation (Matantseva and Skarlato, 2013).

Another subgroup within the Alveolata supergroup is the Ciliophora (Figure 1.2). This group primarily consists of heterotrophic grazers of small flagellates and diatoms, although some mixotrophic species also exist (Posch et al., 2015). Ciliate growth rates are closely linked to the abundance and productivity of phytoplankton communities (Jensen and Hansen, 2000; Verity, 1985). For example, mixotrophic ciliates like *Mesodinium* consume *Cryptophytes* (Gustafson Jr et al., 2000). Tintinnids (e.g. *Acanthostomella*, *Parafavella*, *Tintinnopsis*) feed on organisms like *Phaeocystis pouchetii* (a haptophyte, Figure 1.2), especially during spring blooms observed in the Dutch Wadden Sea and coastal North Sea (Admiraal and Venekamp, 1986).

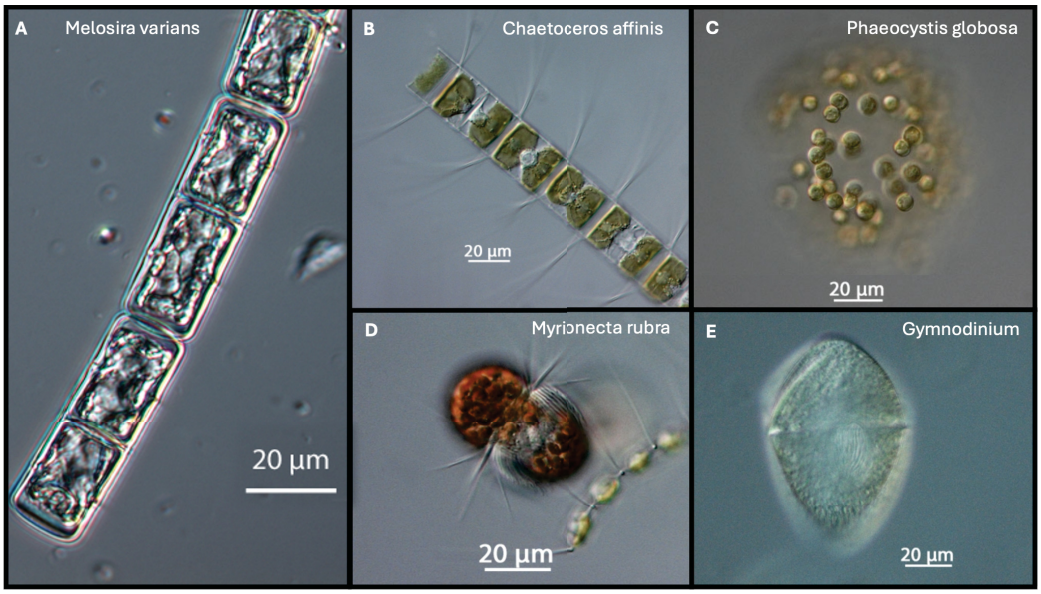

Figure 1.2: **Representatives of microeukaryotes found in the Fram Strait. They can be characterized as follows:** Diatoms (*Melosira varians*, *Chaetoceros affinis*) and haptophytes (*Phaeocystis globosa*) are autotrophic, while dinoflagellates (*Gymnodinium*) are mainly mixotrophic. Ciliates (*Myrionecta rubra*) exhibit a mixotrophic lifestyle (*Mesodinium sp.*) or are heterotrophic. The images were selected and modified from various publications for illustrative purposes.The scales for A to D are 20 µm respectively. All images were sourced from Plankton Net, a provider of biodiversity data. All images are licensed under the Creative Commons Attribution 3.0 (planktonnet.awi.de, 2024).

1.7 The essential role of Zooplankton in the Arctic food web

Zooplankton are a diverse array of small organisms that drift in the water column of oceans, seas, and freshwater systems (Stockdale et al., 2015). It is important to acknowledge the significance of zooplankton, as they include a variety of animal species, such as copepods, krill, jellyfish larvae, and other invertebrates. They are integral components of the Arctic food web and play a crucial role in marine ecosystems (Spisla, 2021).

In the Arctic ecosystem, zooplankton play a crucial role as a link between primary producers, such as phytoplankton, and higher trophic levels, including fish, seabirds, and marine mammals (Spisla, 2021). Zooplankton consume phytoplankton and other organic particles, transferring energy from lower to higher trophic levels through predation and consumption. They are the primary food source for many Arctic organisms, forming the base of the marine food web. This highlights the importance of zooplankton in maintaining the Arctic ecosystem's delicate balance.

Furthermore, zooplankton are of vital importance in the cycling of nutrients and the sequestration of carbon in Arctic waters (Botterell et al., 2023; Lebrato et al., 2019). They release nutrients through excretion and fecal pellets while consuming phytoplankton and other organic matter, which in turn stimulates primary production and supports the growth of marine plants (Kyewalyanga, 2016). When zooplankton die or are consumed by predators, their organic matter sinks to the seafloor. The ecological significance of zooplankton populations in the Arctic highlights the need to comprehend and preserve them, especially considering the current environmental changes and human impacts (Lomartire et al., 2021). It is crucial to acknowledge the importance of these populations. This organic matter can become buried and stored as sedimentary carbon, contributing to the long-term storage of carbon in Arctic ecosystems. Zooplankton play a pivotal role in regulating carbon dioxide (CO_2) levels in both the atmosphere and the ocean. Their feeding activities and respiration significantly influence the balance of CO_2 between the surface ocean and the atmosphere, which can have a profound impact on global climate dynamics (Cavan et al., 2017; Lean, 2017). The abundance, distribution, or composition of zooplankton populations in the Arctic can have far-reaching consequences for ecosystem structure and function, as well as for climate regulation on a broader scale (Yamuza-Magdaleno et al., 2024).

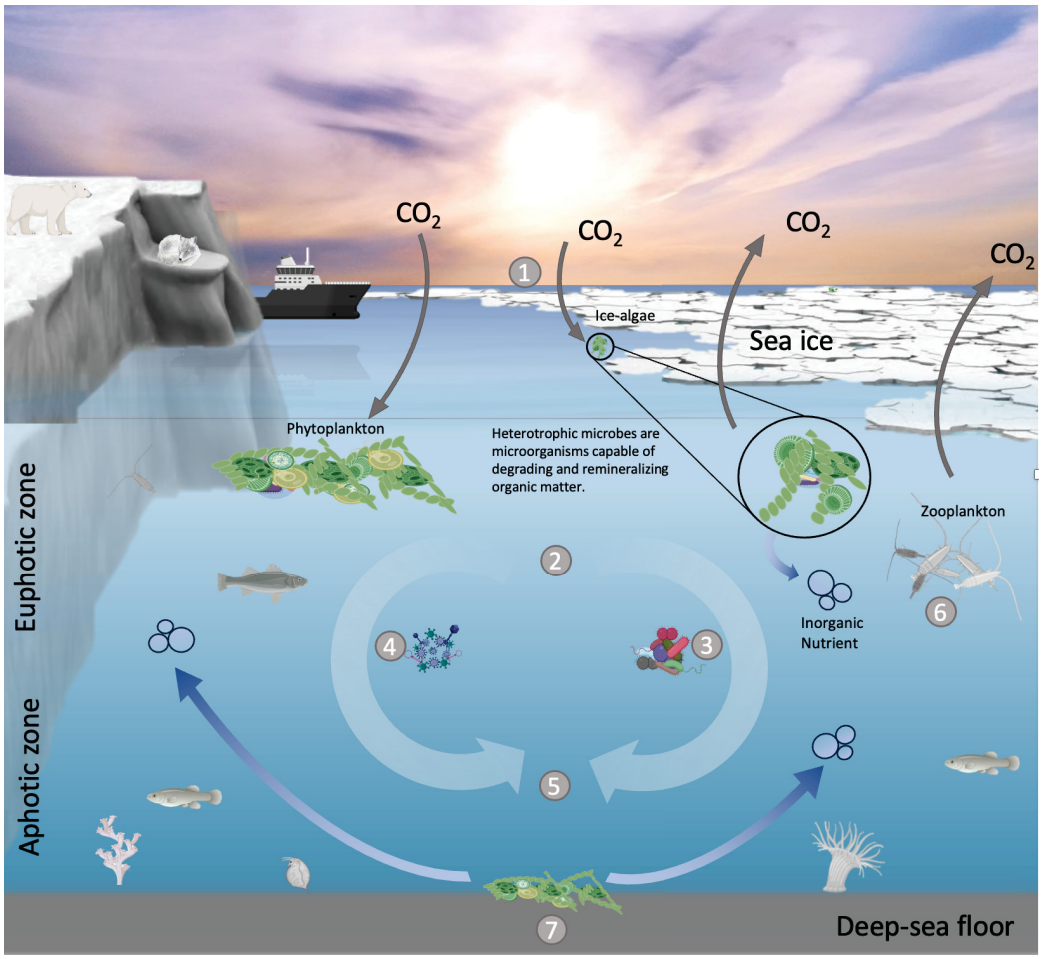

Figure 1.3: **Fundamental processes of the biological carbon pump operating in the Arctic Ocean.** (1) Photosynthetic primary producers convert CO_2 into biomass. (2) Organic matter is respired by heterotrophs. (3) Bacteria decompose and recycle nutrients. (4) Viral lysis releases organic matter, contributing to the microbial loop. (5) Most organic material is consumed in the surface ocean. (6) Zooplankton and higher trophic levels transfer organic matter. (7) A small fraction reaches the deep-sea floor, with about 0.1% buried for long-term sequestration.

1.8 The impact of climate change on the Arctic Ocean

Although geographically isolated, the Arctic Ocean has emerged as a sentinel for global warming, exhibiting temperature increases two to three times faster than the global average (Belan et al., 2022; Rantanen et al., 2022) (Figure **??**(a)). Although the extent and thickness of sea ice in the Arctic Ocean has varied significantly throughout its geological history, including episodes of ice-free conditions during particularly warm periods (Polyak et al., 2010), the observed rate of ice loss in recent decades is unprecedented and without comparison to any historical record (Polyak et al., 2010; Schweiger et al., 2019). Due to atmospheric warming, rising temperatures in the Arctic Ocean are leading to a reduction in sea ice extent (Cavalieri and Parkinson, 2012; Serreze et al., 2007; Stroeve et al., 2007; Stroeve et al., 2012b) and a decrease in mean ice thickness (Laxon et al., 2013; Rothrock et al., 2008).

This loss of ice is accompanied by changes in its physical properties (Figure **??** (b), (d), (e)), with the average age of the ice decreasing and shifting to a thinner and younger ice cover (Kwok, 2007; Maslanik et al., 2007; Perovich et al., 2017; Tschudi et al., 2016). During the 1980s, the spring ice cover was predominantly composed of thick multi-year ice, but this has now changed dramatically, with first-year ice largely taking its place (Stroeve et al., 2012a; Tucker III et al., 2001). This shift renders the ice cover more susceptible to summer melt (Perovich et al., 2017) and results in a lengthening of the summer melt season (Figure **??** (b)), characterized by an earlier onset and later freeze-up (highlighted by Stroeve et al., 2014). In addition, there has been an increase in drift velocity and ice export through Fram Strait due to thinning of Arctic sea ice (Smedsrud et al., 2017; Spreen et al., 2020), which may also promote more frequent lead and ridge formation (Assmy et al., 2017). The observed loss of ice in the Arctic is directly related to anthropogenic CO_2 emissions (Notz and Stroeve, 2016), with various model projections suggesting that the remaining summer sea ice in the Arctic will disappear before mid-century (Notz and Stroeve, 2016; Snape and Forster, 2014; Zege et al., 2015) unless there is a significant reduction in emissions (Mahlstein and Knutti, 2012; Notz and Stroeve, 2016; Overland et al., 2014). However, most models currently underestimate the rate of decline, with observed losses exceeding predictions (Notz and Stroeve, 2016; Overland and Wang, 2013; Winton, 2011). In addition, to the drastic reduction in sea ice cover (Figure **??** (b)), the Arctic Ocean is undergoing a process of freshening (Carmack et al., 2016; Jean-Michel et al., 2021; Lellouche et al., 2018) and warming (McLaughlin et al., 2009; Polyakov et al., 2012; Steele and Dickinson, 2016). Remarkably, these observed changes are intricately linked and to some extent contribute to each other, forming a complex positive feedback loop known as Arctic amplification (Alexeev and Jackson, 2013; Box et al., 2019; Dai et al., 2019; Taylor et al., 2013).

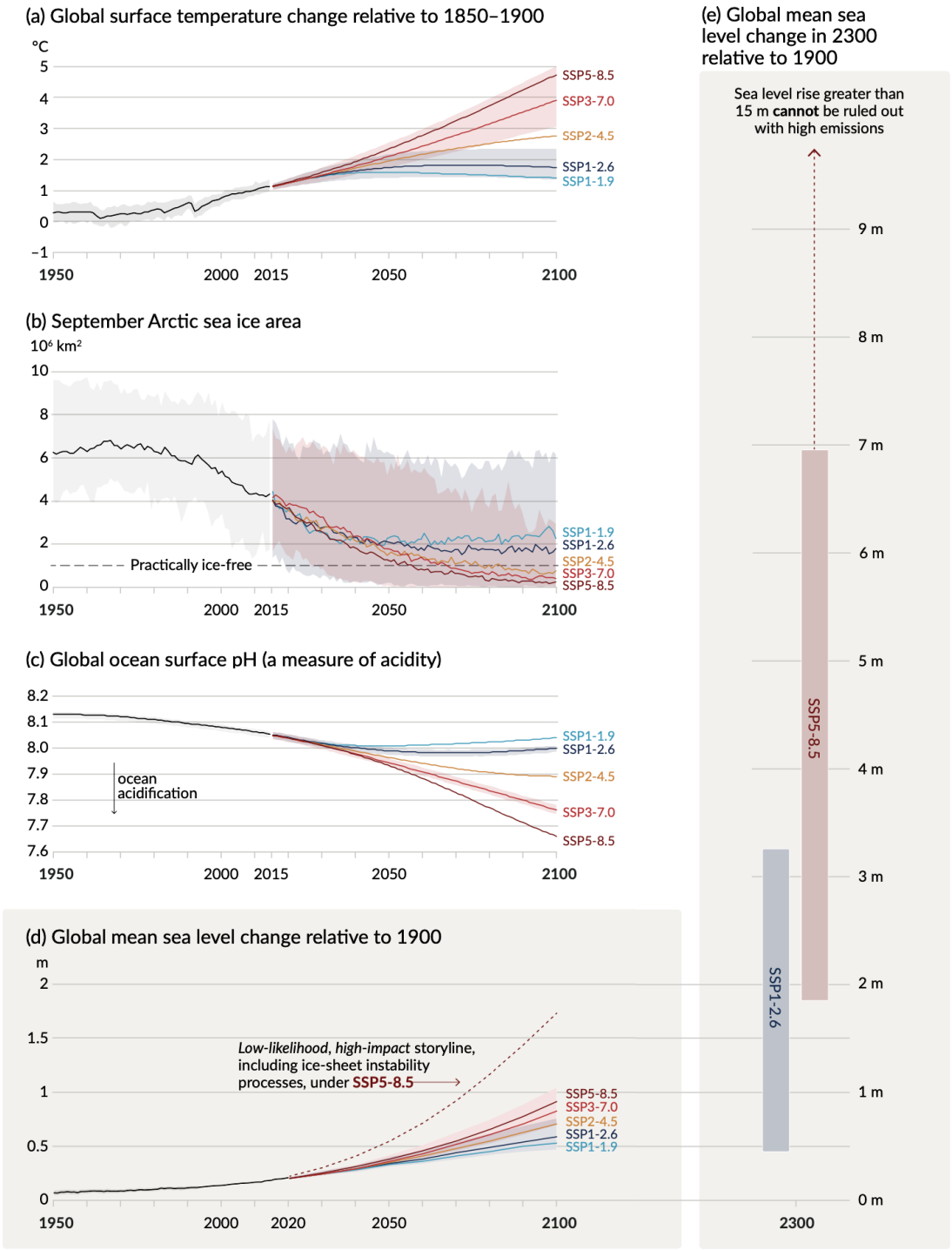

Figure SPM.8 | Selected indicators of global climate change under the five illustrative scenarios used in this Report.

Figure 1.4: The projections for each of the five scenarios are shown in colour. Shades represent uncertainty ranges – more detail is provided for each panel below. The black curves represent the historical simulations (panels a, b, c) or the observations (panel d). Historical values are included in all graphs to provide context for the projected future changes. **Panel (a) Global surface temperature changes** in °C relative to 1850–1900. These changes were obtained by combining Coupled Model Intercomparison Project Phase 6 (CMIP6) model simulations with observational constraints based on past simulated warming, as well as an updated assessment of equilibrium climate sensitivity (see Box SPM.1). Changes relative to 1850–1900 based on 20-year averaging periods are calculated by adding 0.85°C (the observed global surface temperature increase from 1850–1900 to 1995–2014) to simulated changes relative to 1995–2014. *Very likely* ranges are shown for SSP1-2.6 and SSP3-7.0. **Panel (b) September Arctic sea ice area** in 10^6 km^2 based on CMIP6 model simulations. *Very likely* ranges are shown for SSP1-2.6 and SSP3-7.0. The Arctic is projected to be practically ice-free near mid-century under intermediate and high GHG emissions scenarios. **Panel (c) Global ocean surface pH** (a measure of acidity) based on CMIP6 model simulations. *Very likely* ranges are shown for SSP1-2.6 and SSP3-7.0. **Panel (d) Global mean sea level change** in metres, relative to 1900. The historical changes are observed (from tide gauges before 1992 and altimeters afterwards), and the future changes are assessed consistently with observational constraints based on emulation of CMIP, ice-sheet, and glacier models. *Likely* ranges are shown for SSP1-2.6 and SSP3-7.0. Only *likely* ranges are assessed for sea level changes due to difficulties in estimating the distribution of deeply uncertain processes. The dashed curve indicates the potential impact of these deeply uncertain processes. It shows the 83rd percentile of SSP5-8.5 projections that include low-likelihood, high-impact ice-sheet processes that cannot be ruled out; because of *low confidence* in projections of these processes, this curve does not constitute part of a *likely* range. Changes relative to 1900 are calculated by adding 0.158 m (observed global mean sea level rise from 1900 to 1995–2014) to simulated and observed changes relative to 1995–2014. **Panel (e) Global mean sea level change at 2300** in metres relative to 1900. Only SSP1-2.6 and SSP5-8.5 are projected at 2300, as simulations that extend beyond 2100 for the other scenarios are too few for robust results. The 17th–83rd percentile ranges are shaded. The dashed arrow illustrates the 83rd percentile of SSP5-8.5 projections that include low-likelihood, high-impact ice-sheet processes that cannot be ruled out. Panels (b) and (c) are based on single simulations from each model, and so include a component of internal variability. Panels (a), (d) and (e) are based on long-term averages, and hence the contributions from internal variability are small. {4.3; Figures 4.2, 4.8, and 4.11; 9.6; Figure 9.27; Figures TS.8 and TS.11; Box TS.4, Figure 1} This figure (Figure SPM.8) has been reproduced with the approval of the IPCC (Intergovernmental Panel on Climate Change), 2021: Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (Masson-Delmotte et al., 2021a).

1.9 Melting ice is reshaping microbial communities in the Arctic Ocean

Rapid climate change is already affecting Arctic ecosystems, most notably in larger marine organisms from zooplankton to polar bears (Wassmann, 2011), with implications for food web structure and the link between pelagic and benthic environments (highlighted by (Grebmeier et al., 2006)). While microbial responses to these ecosystem changes are less well understood, reports of such responses have emerged from different regions of the Arctic Ocean (Barber et al., 2015; Vincent, 2010; Wassmann, 2011).

Concomitant with alterations in the physicochemical properties of the sea ice cover in the Amerasian Basin from the 1970s to the late 1990s, there have been notable changes in the composition of sea ice algal communities. These changes have led to a reduction in overall diversity and abundance, likely attributable to ice melt and freshwater input into the upper ocean (Dupont, 2012; Kohlbach, 2017).

Changes in the duration of the melt season (Kahru et al., 2010; Rapp, 2018; Stroeve et al., 2014), as well as changes in the extent and thickness of the ice cover (Arrigo and Dijken, 2015), have extended the growing season for phytoplankton and increased light availability, thereby increasing primary productivity in Arctic coastal regions (Arrigo and Dijken, 2015; Arrigo et al., 2012; Negrete-Garcıéa et al., 2024; Pabi et al., 2008; Tremblay et al., 2011). The rise in water temperature and the increased inflow of Atlantic and Pacific water masses (Korhonen et al., 2013; Polyakov et al., 2012; Rudels et al., 2013; Spielhagen et al., 2011) have facilitated the invasion of phytoplankton species typical of the North Atlantic and Pacific oceans into the Arctic, such as the Coccolithophore *Emiliania huxleyi*, the picocyanobacterium Synechococcus, and the diatom *Neodenticula seminae* (Nöthig et al., 2015; Paulsen et al., 2016). This is the first time in over 800,000 years that Pacific phytoplankton has been able to migrate through the Arctic to the Atlantic Ocean (Miettinen et al., 2013), indicating the potential for increased trans-Arctic exchange if warming trends continue (Zhang et al., 2019). The intensification of upper ocean stratification caused by increased freshwater input has been particularly noticeable in the oligotrophic deep basins, potentially reducing nutrient delivery to the surface (Chen et al., 2021; McLaughlin and Carmack, 2010; Slagstad et al., 2015) and thereby limiting primary production within the water column despite increased illumination (Tremblay and Gagnon, 2009). A change in phytoplankton composition has been observed, characterized by an increase in small algae and bacteria, with a concomitant decrease in larger algae, which may be less efficient at nutrient acquisition (Franz, 2012; Li et al., 2009; Paulsen et al., 2016; Pinhassi et al., 2004). Ocean freshening following significant ice melt in 2007 also triggered changes in the structure of the pelagic bacterial community in the Canadian Arctic, resulting in a less diverse community after 2007 and a notable decline in Bacteroidetes (Comeau et al., 2011). These documented shifts in microbial diversity and community composition are likely to have implications for the biological pump and carbon cycling in the Arctic (Appen et al., 2021; Rapp, 2018). Certainly, the first indications of an impact on the benthic-pelagic coupling came from extensive monitoring in the Fram Strait. This monitoring revealed a consistent rise in water temperature and salinity in the North Atlantic inflow (Walczowski et al., 2017), which

coincided with changes in phytoplankton community composition and export flux from 1998 to 2011. Even in the depths of the deep seafloor, deviations in organic matter supply observed during the 2005-2007 warm phase were reflected in changes in benthic bacterial community structure and reduced diversity (Jacob, 2014; Soltwedel et al., 2016). In 2012, extensive ice melt in the central Arctic resulted in the release of substantial sub-ice filaments of the diatom *Melosira arctica* that sank to the seafloor, leading to widespread deposition of fresh ice-algae material at depths of 4400 m (Boetius et al., 2013). Increased rates of oxygen consumption in sediments with algal deposits suggest bacterial remineralization in response to increased carbon flux rates (Boetius et al., 2013). Thus, it seems clear that the environmental changes observed in the Arctic are affecting microbial communities from shelf to deep basin and from sea ice to the seafloor.

1.10 Modern tools for exploring microbial diversity

The "great plate count anomaly", as described by Staley and Konopka in 1985 (Staley and Konopka, 1985), elucidates that only a small proportion of viable marine microorganisms can flourish under traditional laboratory conditions. Consequently, the characterization of microbial communities has historically been constrained by cultivation methods (Berg et al., 2020; Eilers et al., 2000). Over the past few decades, the advent of innovative molecular tools for evaluating microbial diversity has transformed the investigation of microbial communities (Amann et al., 1995; Berg et al., 2020; DeLong and Pace, 2001; Eme and Tamarit, 2024; Lane et al., 1985). In particular, the use of the small subunit ribosomal RNA gene as a molecular marker (Pace et al., 2012; Woese and Fox, 1977; Woese et al., 1990), along with advances in low-cost sequencing technologies and computational methods for data processing, has greatly expanded our ability to understand the taxonomic composition of natural communities. Following this, endeavors in high-throughput sequencing have shed light on the extensive microbial diversity present in the environment, unveiling the genomic makeup and functional capacities of complete microbial populations (Cao et al., 2020; Morey et al., 2013; Pace, 1997; Reuter et al., 2015).

1.11 Amplicon sequencing as a lens into arctic microbial diversity

In contemporary microbiology, amplicon sequencing has emerged as the predominant method for examining microbial diversity in environmental samples. This technique involves targeted sequencing of specific genomic regions of interest, with a primary focus on segments of the 16S or 18S RNA genes to evaluate bacterial and archaeal or eukaryotic diversity, respectively. These gene regions exhibit key characteristics that render them valuable as taxonomic and phylogenetic markers, such as widespread occurrence, significant sequence conservation, and the presence of hypervariable regions (Hugenholtz and Pace, 1996; Lane et al., 1985; Pace, 2009; Woese and Fox, 1977). After DNA is extracted from all cells within a given sample, specific primer sequences are used to target the genes or gene regions of interest for amplification. The

amplified products are then sequenced, and subsequent analysis and comparison to reference databases allows inference of community composition and structure within the original sample. Amplicon sequencing has emerged as a valuable tool for exploring microbial biodiversity in diverse environments, demonstrating its power to reveal the richness and diversity of microbial taxa that are often overlooked or underestimated. It has provided insights into co-occurrence networks among organisms and revealed spatial and temporal distribution patterns of microbial populations (Masenya et al., 2024; Trego et al., 2022). While amplicon sequencing offers numerous benefits, it also presents notable limitations and drawbacks stemming from technical biases inherent in the amplification and sequencing procedures, as well as the choice of gene region. These factors can exert a substantial impact on the detected biodiversity, a concern underscored by various research findings (Ibarbalz et al., 2014; Poretsky et al., 2014; Schloss et al., 2011; Wylie et al., 2012). Furthermore, the applicability of the method is limited to organisms with known dedicated marker genes, potentially missing highly divergent or novel community members beyond the scope of current primer sequences. Differences in the number of RNA gene copies across microbial genomes (Huse et al., 2008; Milanese et al., 2019), along with gene transfer between taxa, can add complexity to determining the true composition of a microbial community in a sample. However, despite these challenges, utilizing amplicon sequencing of the RNA gene remains a cost-effective and time-efficient method for investigating diversity, community structure, and the identity of individual members within extensive microbial datasets.

1.12 Bioinformatics: Unraveling marine ecosystems

In marine biology, understanding the intricate interactions between different organisms, both eukaryotic and bacterial, is critical to understanding ecosystem dynamics and functioning. Bioinformatics, particularly techniques such as network analysis, play a key role in unraveling this complexity and shedding light on the interconnectedness within marine communities.

1.12.1 Data processing and analysis

Marine ecosystems are intricate and data-rich, demanding robust tools for analysis. Bioinformatics has become essential in this field, enabling researchers to handle large-scale genomic, transcriptomic, and proteomic datasets generated by high-throughput sequencing technologies (Lightbody et al., 2019; Manzoni et al., 2018; Ward et al., 2013). These tools facilitate data analysis, integration of multi-omics data, and development of predictive models. Bioinformatics empowers researchers to uncover biological features, elucidate complex processes, and identify potential applications in marine biotechnology (Jain and Tailor, 2020; Rotter et al., 2020). This thesis examines the application of specific bioinformatics and mathematical tools to enhance our comprehension of marine biology.

1.12.2 Time series analysis

Time series analysis is a method of examining a sequence of data points collected over time intervals. In contrast to intermittent or random data recording, time series analysis records data points consistently over a defined period. This allows for the identification of how variables change over time, thereby providing insights into temporal dependencies and adjustments within the data (Montgomery et al., 2015). Access to time series data is crucial for the effective utilisation of statistical and mathematical tools. Such data consists of observations recorded at regular intervals, providing timestamps for each observation (Stock and Watson, 2020). Examples of time series data include stock prices, weather measurements, economic indicators, and physiological parameters such as heart rate. A large dataset is typically required for time series analysis to ensure reliability, consistency, and noise reduction. Moreover, time series data enables forecasting based on historical trends, allowing for predictions of future outcomes (Box et al., 2015; Chatfield and Xing, 2019; Jose, 2022; Tableau, 2024).
The analysis of time series data reveals temporal patterns, trends, and behaviors, providing valuable insights into temporal relationships and facilitating forecasting efforts (Ducklow et al., 2009; Krueger and Von Storch, 2011; Mudelsee, 2010; Steele, 1985). This is essential for comprehending how data evolve over time, regardless of whether the data are collected at regular or irregular intervals, contingent upon the specific phenomenon being investigated (Popa et al., 2020a).
Time series analysis is a powerful tool for capturing and understanding temporal dependencies and relationships within data. This analytical approach yields valuable insights into the dynamic nature of data, facilitating a deeper understanding of complex systems and processes (Lawton, 1988; Petchey et al., 1997; Pimm and Redfearn, 1988). In the context of microbial communities in the Arctic, time series analysis offers several advantages. Time series analysis provides a comprehensive understanding of microbial community dynamics, allowing scientists to draw more confident conclusions about their behavior (Kraemer et al., 2024). By tracking changes in microbial populations across different seasons and environmental conditions, researchers can observe how microorganisms respond to variations in temperature, light availability, nutrient levels, and other factors inherent in the Arctic environment (Comeau et al., 2011; Deslippe et al., 2012). Time series analysis enables the identification of seasonal patterns and trends in microbial abundance, diversity, and activity. This information is of critical importance for comprehending the ecological dynamics of microbial communities in the Arctic and predicting their responses to ongoing environmental changes, such as climate warming and sea ice melting (Brussaard et al., 2013; Macdonald et al., 2005; Mudelsee, 2010). Moreover, time series analysis facilitates the detection of short-term fluctuations and long-term trends in microbial populations (Nöthig et al., 2015). By examining microbial dynamics over extended periods, researchers can distinguish between transient fluctuations and persistent shifts in community composition or structure (Callaghan et al., 2004; Crump et al., 2003). This insight is essential for assessing the resilience of Arctic microbial ecosystems to environmental perturbations and predicting their future trajectories.
In conclusion, time series analysis provides a robust framework for exploring the dynamics of microorganisms in the Arctic and understanding their response to environmental variability and

change (Doherty et al., 2020). By leveraging this analytical approach, scientists can uncover the main factors driving microbial community dynamics, identify crucial thresholds or tipping points, and develop effective strategies for climate change in Arctic ecosystems (Buttigieg et al., 2018; Mock et al., 2016).

1.12.3 Network analysis: Mapping ecosystem interactions

In marine biology, networks are a powerful tool for analyzing the complex interactions and dynamics of marine organisms in ecosystems. They visually display the relationships between entities, such as species, genes, or ecological variables, enabling researchers to confidently identify patterns, key players, and underlying ecological processes. Co-occurrence networks and convergent cross mapping (CCM) are two common types of networks used in marine biology (Barberán et al., 2012; Berry and Widder, 2014; Javier et al., 2022; Sugihara et al., 2012).

Co-occurrence Networks Co-occurrence networks clearly illustrate the co-occurrence patterns of organisms or ecological variables across samples or environmental conditions. The nodes in the network represent individual entities (e.g. species or environmental factors), and the edges (connections) between nodes indicate significant co-occurrence relationships between them. In a microbial co-occurrence network, nodes represent microbial taxa, and edges represent co-occurrence patterns between taxa across different marine samples. Such networks reveal ecological associations, interactions, and potential dependencies between organisms or environmental factors, providing valuable insights into community structure, species interactions, and ecosystem dynamics (Barberán et al., 2012; Berry and Widder, 2014).

Convergent Cross Mapping CCM Convergent cross mapping CCM is a powerful statistical technique that infers causal relationships between variables based on time series data. By assessing the predictability of one variable using time series data from another variable, CCM provides strong evidence for a potential causal relationship between them. This technique is particularly useful for identifying causal connections between ecological variables or environmental factors, and has been widely adopted by researchers in these fields (Javier et al., 2022; Sugihara et al., 2012). CCM is a powerful tool in marine biology that can be used to understand the behavior of marine organisms.

1.13 Mathematical Modeling: Quantifying ecosystem stability

The application of mathematical models in the study of natural phenomena has a long history, with notable contributions from the physical sciences and, more recently, the biological sciences. Over the past approximately 40 years, mathematical models have been used to gain insights into a wide range of natural phenomena. The fundamental principle behind constructing a mathematical representation of a natural phenomenon is to simplify it to its essential components, discarding extraneous details that are irrelevant to the research question. Two primary approaches to model building exist: the top-down approach, which begins with data and proceeds to deduce the metabolic network structure, and the bottom-up approach, which begins

with knowledge about the phenomenon and proceeds to construct a qualitative mathematical model that reproduces experimental observations. Following the selection of the model-building approach, researchers must choose a modeling technique from various available mathematical methods. With advancements in computational power, numerous software programs are now capable of efficiently solving these models within a reasonable time-frame. Mathematical modeling employs mathematical equations and computational algorithms to accurately simulate real-world processes, phenomena, or systems (Bellouquid and Delitala, 2006; Ji et al., 2017). In the context of Arctic marine communities, mathematical modeling offers numerous benefits for comprehending ecosystem dynamics, species interactions, and responses to environmental changes (Fulford et al., 2020). Mathematical models have a clear advantage in exploring complex ecological relationships and predicting ecosystem responses under different scenarios. Through the use of mathematical models, we can gain insights that may not be achievable through observational or experimental studies alone (Mazur, 2006). This is due to their ability to integrate data from various sources, including field observations, laboratory experiments, and remote sensing, to generate comprehensive representations of marine ecosystems and their dynamics. Changes in temperature, sea ice extent, nutrient availability, and other environmental factors can be predicted through models, allowing for a better understanding of how they may affect species distributions, population dynamics, and ecosystem structure (Chen et al., 2023). Mathematical models inform conservation efforts, resource management decisions, and policy interventions to mitigate climate change effects on Arctic marine ecosystems through simulating various climate scenarios (Parrott et al., 2012).
Energy landscape analysis (ELA) quantifies the stability and resilience of Arctic ecological communities by mapping their energy landscapes. The energy landscape represents the potential energy of a system as a function of its configuration or state, with lower energy states indicating more stable configurations. Through analysis of the topology and properties of the energy landscape, researchers can confidently identify stable community configurations, critical thresholds, and potential regime shifts in marine ecosystems (Fujita et al., 2023; Suzuki et al., 2021). The Energy Landscape Analysis provides valuable insights into the mechanisms driving ecosystem stability, the impacts of environmental disturbances, and the resilience of marine communities in the Arctic. Incorporating ELA into mathematical models of Arctic marine ecosystems may allow researchers to improve predictions of ecosystem responses to climate change and enhance our understanding of the complex interactions that shape Arctic marine communities.

1.14 Artificial intelligence empowers image recognition to improve data analysis

Artificial intelligence (AI) is the broad field of creating machines capable of intelligent behavior. Machine learning (ML) is a subset of AI that involves training algorithms to learn from and make predictions based on data. Deep learning (DL) is a further subset of machine learning

that uses neural networks with many layers to analyze complex patterns in large datasets (Goodfellow et al., 2016).

1.14.1 Improving Accuracy / Precision

Advanced computational algorithms employ machine learning techniques to train on various datasets of annotated images, allowing the algorithms to learn patterns and features indicative of different microbial species or cellular structures. These algorithms can differentiate between closely related strains or morphologically similar organisms. Once trained, DL models can efficiently analyze diverse microbial images from various sources, including laboratory cultures, environmental samples, and clinical specimens. This ability is particularly valuable in fields such as environmental microbiology, where assessing the diversity and complexity of microbial communities using traditional methods can be challenging (Eerola et al., 2024; Luo et al., 2018).

1.14.2 Scalability

Deep Learning in image recognition offers significant benefits, particularly in terms of scalability and adaptability (Burns et al., 2023; Kumar et al., 2023). These programs offer are cost-effectiveness and saving valuable time, allowing researchers to process large datasets efficiently without the need for extensive manual labor or specialized expertise (Cheng et al., 2019; O'Mahony et al., 2020). This is particularly advantageous in the Arctic, where limited resources and logistical challenges often hinder research efforts. Automating the identification process based on deep learning maximizes the utility of available data and allows researchers to focus their efforts on data analysis, interpretation, and scientific discovery (Oldenburg et al., 2023b).
In conclusion, DL applied on marine organisms in the Arctic, providing improved efficiency, accuracy, scalability, and adaptability to changing environmental conditions (Oldenburg et al., 2023b). AI technologies are advancing rapidly, providing great potential for enhancing our understanding of Arctic marine ecosystems in the face of ongoing environmental changes (Andersson et al., 2021; Buškus et al., 2021).

1.15 Synergy for Arctic Ocean research

The integration of bioinformatics, mathematical modeling, and artificial intelligence offers unique advantages for the study of marine biology in the Arctic Ocean. In this dynamic and environmentally sensitive region, where samples are limited, environmental variability is substantial, and future predictions are critical, the synergy of these three approaches is invaluable. By leveraging comprehensive data integration, advanced modeling techniques, and AI-based analysis, researchers can gain deeper insights into the structure, function, and resilience of Arctic marine ecosystems (Mueter et al., 2021). This holistic understanding is essential for informing conservation efforts, guiding sustainable management practices, and mitigating the impacts of climate change on Arctic marine ecosystems and the broader Earth system.

1.16 Research Questions

The classic paradigm that marine life comes to a standstill during Arctic winter has recently been challenged by data showing that even during the polar night, zooplankton species were active and feeding, and some key phytoplankton species were found to be widely distributed in Svalbard waters (Berge et al., 2015a). The underpinning mechanisms and consequences for Arctic marine biogeochemical cycles, ecosystem functionality and functions are still unknown. The overall aim of this research project is to gain deeper insights into the key components that drive the dynamics of Arctic microbial ecosystems. Special attention will be paid to the transition period from the end of summer to the beginning of winter, which will be investigated by analyzing time series data. This is crucial for understanding ecosystem dynamics, predicting future changes, and informing strategies for conservation. This doctoral thesis will investigate several critical research questions related to the influence of environmental factors on phytoplankton and microbial communities in the Arctic Ocean.

RQ1: How do environmental conditions, including sea ice melt, influence the prevalence of specific phytoplankton species in the Arctic, and how does the tolerance of different phytoplankton taxa to varying oceanographic conditions affect their ability to persist and spread in Arctic waters?

Relevance: Recognizing the significance of environmental changes, particularly the melting of sea ice, on phytoplankton dynamics in the Arctic is of utmost importance for Arctic ecology and climate research. Phytoplankton serve as the cornerstone of Arctic marine ecosystems, acting as primary producers that sustain higher trophic levels. Changes in phytoplankton composition and distribution can have significant impacts throughout the food web, affecting the abundance and distribution of marine organisms, from zooplankton to marine mammals. Additionally, phytoplankton play a critical role in carbon cycling and the regulation of global climate. Insights gained from studying the response of phytoplankton communities to environmental changes provide valuable information for predicting and mitigating the ecological and climatic consequences of ongoing Arctic warming and sea ice decline.

Our Approach: Our approach aims to comprehensively understand the environmental preferences of Arctic and Atlantic microbial communities and their spatial-temporal variations. We have observed distinct microbial preferences for different water regimes, such as mixed-layer and meltwater regimes. Additionally, we have analyzed how microbial community composition changes over space and time by comparing locations and years, which highlights the dynamic nature of Arctic microbial ecosystems. We have thoroughly analysed the seasonal succession of microbial communities and have identified clusters that represent synchronized occurrences of microbes (see Chapter 2).

RQ2: How does the increasing influx of Atlantic water and changes in sea-ice cover affect the distribution and ecological roles of prokaryotic communities in the Arctic Ocean, and what are the implications for ecosystem dynamics?

Relevance: Understanding the ecological and dynamic changes in the Arctic Ocean as a result of global warming is crucial for understanding ecosystem structure and function. By analyzing

high-resolution data sets spanning several years and different spatial and temporal scales, the effects of Atlantic water inflow and sea ice loss on Arctic bacterial communities will be revealed. Examining how microbial communities respond to changing environmental conditions provides insight into shifts in bacterial populations, with densely ice-covered polar waters harboring stable resident microbiomes, while the influx of Atlantic water leads to the dominance of seasonally fluctuating populations. This allows the identification of characteristic bacterial populations associated with different environmental conditions and a better understanding of their ecological roles. Importantly, the studies provide evidence for metabolic differences between bacteria adapted to Arctic and Atlantic conditions, with implications for food webs and biogeochemical cycles. Overall, this work provides new insights into Arctic ecology and points to an ongoing biological atlanticization of the warming Arctic Ocean.

Our Approach: Our study aims to unravel the effects of Atlantification and sea-ice decline on the composition, diversity, and functional capabilities of Arctic microbial communities. To achieve this, we analyze high-resolution data spanning four years, including high-resolution amplicon data and metagenomes generated from the East Greenland Current (EGC) and supplemented with existing datasets from TARA Arctic (Gascard et al., 2008) and MOSAiC (Mock et al., 2022). This approach allows us to assess how variations in sea-ice cover and Atlantic water influx influence microbial community dynamics. By identifying microbial signatures associated with specific environmental conditions and exploring their ecological roles, we gain insight into the mechanisms driving changes in Arctic ecosystem dynamics in response to ongoing environmental shifts (see Chapter 3).

RQ3: How do microbial communities assemble and function across seasonal and inter-annual scales in the pelagic Arctic Ocean, and what are the underlying environmental drivers of these dynamics?

Relevance: An understanding of the dynamics of microbial communities in marine ecosystems is essential to comprehend how these ecosystems function and how they adapt to environmental change. This study highlights the seasonal and interannual dynamics of prokaryotic and microeukaryotic communities in pelagic marine ecosystems of the Arctic. Using advanced sampling techniques over a period of four years, the study reveals recurrent fluctuations in the predominant populations and gene content of the community organised in different seasonal modules. The identification of distinct microbial signatures associated with specific environmental conditions, such as the polar night and polar summer, emphasises the complex links between taxonomy, function and oceanographic factors. These modules represent unique ecological states that are influenced by specific microeukaryotic populations and environmental conditions. Furthermore, the study reveals heterogeneous environmental selection processes in these ecological states, providing valuable insights into the structural and functional organisation of microbial communities in rapidly changing polar marine ecosystems. In conclusion, this research contributes significantly to our understanding of how microbiomes respond to pronounced environmental variability in understudied marine regions. Furthermore, this approach is crucial, as microbial communities are not only defined by their taxonomic composition, but also by their functional roles and interactions with the environment. By considering these

multiple dimensions, the study contributes to a better understanding of microbial ecology and ecosystem functioning.

Our Approach: The study utilizes a multi-faceted approach to explore microbial community dynamics in the Arctic Ocean. Moorings with autonomous devices continuously monitor taxonomic, functional, and environmental parameters over four years, offering high-resolution data. Advanced sequencing methods provide detailed inventories of microbial taxa and gene content, integrated with environmental data. Sophisticated analytical techniques like Fourier transformation and co-occurrence network analysis identify recurrent patterns, ecological states, and environmental drivers, revealing insights into microbial community structure and function (see Chapter 4).

RQ4: How do microbial communities in the Arctic Ocean respond to seasonal changes and environmental shifts, and what are the key species driving these dynamics?

Relevance: Given the significant environmental changes currently occurring in the Arctic Ocean, it is critical to gain a deeper understanding of how microbial communities respond in order to accurately predict ecosystem resilience. Conventional analytical methods employed to examine microbial communities often struggle to identify key species that are essential for maintaining ecosystem stability. To address this challenge, we have developed an innovative analytical approach that integrates co-occurrence networks, convergent cross mapping, and energy landscape analysis. These methods enable the identification of seasonal microbial communities and their interactions, providing a more comprehensive understanding of the complex dynamics within these ecosystems. Convergent Cross Mapping Network enables the clear visualisation of seasonal patterns in phytoplankton composition and species interactions. Additionally, a "winter reset" phenomenon is revealed, influencing the dynamics of the community. Energy landscape analysis also demonstrates the differing stability of winter communities compared to summer communities. Identifying key species that determine microbial community dynamics allows for the identification of species most sensitive to environmental change. An understanding of these responses provides insight into potential impacts at higher trophic levels, which is of importance to industries such as commercial fisheries and indigenous communities that rely on Arctic marine resources.

Our Approach: Our approach focuses on understanding Arctic microbial communities' responses to environmental changes by utilizing three advanced analytical methods. By integrating Co-Occurrence Networks, Convergent Cross Mapping, and Energy Landscape Analysis, we identify key species driving community dynamics and assess ecosystem stability. With this comprehensive approach, we can recognise seasonal shifts in microbial community interactions and thus gain a deeper insight into how communities interact in the Arctic Ocean and what effects changes can have (see Chapter 5).

RQ5: How can we quickly and efficiently pre-sort and analyze zooplankton images without internet access and limited computing power on a ship?

Relevance: Zooplankton plays a crucial role in the marine ecosystem, exerting a significant influence on food chains and the nutrient cycle. Nevertheless, the vertical distribution of

these organisms is challenging to quantify. During an Arctic expedition, a device known as the Lightframe On-sight Keyspecies Investigation (LOKI) is employed at various positions to analyse the vertical distribution of zooplankton. The LOKI is a method of concentrating zooplankton by means of a net, which leads to a flow chamber in which a camera records images. The high-resolution images permit the identification of zooplankton taxa, frequently to the genus or species level, and developmental stages. A considerable number of images are generated during each expedition, necessitating manual analysis on board. This process currently requires a significant amount of time and an internet connection. The advent of Big Data and the use of AI in biology has led to an increased importance in the development of data pipelines. This enables rapid, effective and time-saving analysis, allowing scientists to gain an overview of the population at an earlier stage and make time-saving decisions on the ship, for example with regard to the further course of the experiments.

Our Approach: To address the questions, we employed a two-step approach. Initially, we implemented an automated deep transfer learning approach for pre-sorting the images based on the preferred taxonomic level. Subsequently, we developed a graphical user interface and validated our results with the state-of-the-art technology currently in use by the scientists on the ship. This process not only offers a significantly faster alternative to manual classification but also demonstrates greater accuracy compared to the software solution currently in use. The solution can be executed on a standard MacBook, rendering it suitable for deployment in remote locations lacking a stable internet connection, such as the Central Arctic. The algorithm necessitates training with already sorted images, and the superior the training set, the more accurate the prediction. Our model achieved an impressive accuracy of 83.9%, which is twice as accurate as conventional methods such as EcoTaxa. Furthermore, AI can be employed to identify anomalies and enhance data quality. This methodology is not dependent on the system in question and can be employed with different imaging systems, provided that sufficient labelled data is available (see Chapter 6).

Within the framework of this research, three papers have been accepted for publication: Oldenburg et al., 2023b, 2024b; Priest et al., 2023. In addition, two paper have been submitted for publication: Oldenburg et al., 2024a; Priest et al., 2024.

Furthermore, two software projects based on the papers (DeepLOKI[1]), (otter[2]), have been published under MIT license or will be after acceptance.

1.17 Outline of this thesis

The first chapter provides an overview of the Arctic Ocean and its vulnerability to climate change. The text emphasises the importance of understanding the dynamics of phytoplankton and bacterial communities in response to environmental change, particularly sea ice melt and the influx of Atlantic water and the resulting atlantification. Moreover, the utilisation of diverse methodologies to gain a more profound comprehension of the dynamics of microbial communities and the significance of studying key species in Arctic microbial ecosystems will

[1]Source code: `https://gitlab.com/qtb-hhu/qtb-sda/DeepLOKI`,
[2]Source code: `https://github.com/rakro101/otter`,

be elucidated. Furthermore, the development of an AI-based software framework for higher trophic levels, such as zooplankton analysis, will be presented.
Chapter 2 investigates the impact of sea ice melt on phytoplankton dynamics in the Arctic Ocean. It presents findings from a study conducted over two years at two mooring locations, revealing the dominance of polar phytoplankton during periods of high diatom abundance and the limited prevalence of temperate taxa. Furthermore, it discusses the Atlantification for the northward expansion of temperate species.
Chapter 3 examines the influence of Atlantic water influx on Arctic bacterial communities. It analyses data collected over four years to identify stable microbiomes in densely ice-covered polar waters and fluctuating populations in regions with increased Atlantic water influx. Furthermore, the chapter explores the discovery of bacterial signature populations associated with distinct environmental conditions and metabolic disparities between Arctic and Atlantic-affiliated bacteria.
Chapter 4 presents the investigation of microbial community dynamics in the Arctic Ocean, revealing recurrent seasonal patterns and distinct ecological states over a four-year period. Integrating taxonomic, functional, and environmental data, the study demonstrates heterogeneous environmental selection across seasons, with stronger functional than taxonomic selection shaping microbial communities across pronounced environmental gradients.
Chapter 5 presents an analytical approach for identifying keystone species in Arctic microbial ecosystems. It describes the methodology based on Co-Occurrence Networks, Convergent Cross Mapping, and Energy Landscape Analysis. The chapter discusses the identification of keystone species representative of different trophic modes and explores seasonal patterns in microbial phytoplankton communities and their interactions.
Chapter 6 introduces DeepLOKI, a software framework developed for analysing high-resolution zooplankton images. It details the development process and validation of DeepLOKI using Deep Transfer Learning with a Convolutional Neural Network backbone. Additionally, the chapter demonstrates DeepLOKI's efficiency in achieving high classification accuracy and streamlining the annotation process for zooplankton research.
The concluding last chapter presents a synthesis of the principal findings from each preceding chapter, together with an analysis of their implications for the understanding of Arctic ecosystem functioning. It also outlines prospective avenues for future research, including the application of analytical approaches to other Arctic regions and the integration of multi-omics data for comprehensive ecosystem analysis. Finally, the chapter emphasises the continued necessity for research in order to mitigate the impacts of climate change on Arctic marine ecosystems.

Chapter 2

Eukaryotes: Sea-ice melt determines seasonal phytoplankton dynamics and delimits the habitat of temperate Atlantic taxa as the Arctic Ocean atlantifies

In this chapter, we present our use case concerning temperate eukaryotes at the F4 mooring: Sea-ice melt acts as a barrier. Sea-ice melt determines seasonal phytoplankton dynamics and delimits the habitat of temperate Atlantic taxa as the Arctic Ocean atlantifies.

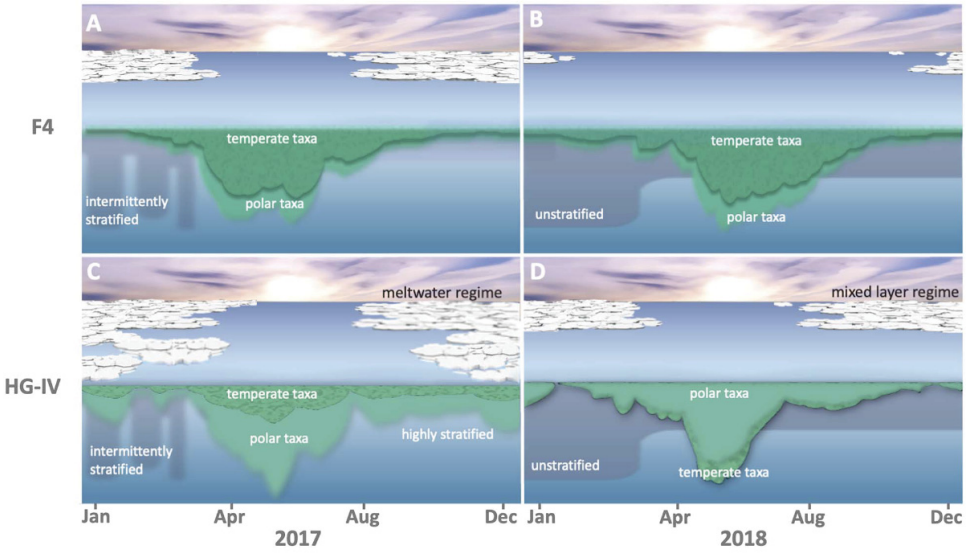

Figure 2.1: **Effects of meltwater and mixed layer conditions on temperate (dark green) and polar (light green) taxa from 2017 to 2018.** A-B: Similar abundances observed in highly stratified meltwater and mixed layer regimes at cluster F-06. C-D: Reduced abundances of temperate and polar taxa in different regimes at cluster H-06 (Oldenburg et al., 2023a).

2.1 Sea-ice melt acts as a barrier establishment of temperate Atlantic taxa in atlantified Arctic Ocean

In this section, we provide an overview of the contributions and impact of our paper (Oldenburg et al., 2023a):

Ellen Oldenburg, Ovidiu Popa, Matthias Wietz, Wilken-Jon von Appen, Sinhue Torres-Valdes, Christina Bienhold, Oliver Ebenhöh, Katja Metfies
"Sea-ice melt acts as a barrier establishment of temperate Atlantic taxa in atlantified Arctic Ocean"
In: *ISME Communications, 2024, 4(1)*

Main Results in Simple Terms

During a two-year period, we conducted a study on phytoplankton in two locations (moorings) in the Fram Strait. While this duration may not be enough to represent long-term changes, it provided us with valuable insights into potential future outcomes. It is possible that climate change could lead to the Arctic Ocean being ice-free during summer but ice-covered during winter. Our research has shown that this melting ice has an impact on the eukaryotes, minuscule plant-like organisms in the water. Specifically, in 2017, we observed a higher concentration of these organisms due to increased ice melt. It was found that certain eukaryote species are more adapted to the conditions similar to the Atlantic Ocean, while others struggle in colder, ice-covered waters. The study suggests that the melting of sea ice in the Arctic could hinder the northward spread of certain phytoplankton species, potentially leading to long-term effects on the ecosystem.

Summary/Abstract

The Arctic Ocean is one of the regions where anthropogenic environmental change is progressing most rapidly and drastically. The impact of rising temperatures and decreasing sea ice on Arctic marine microbial communities is yet not well understood. Microbes form the basis of food webs in the Arctic Ocean, providing energy for larger organisms. Previous studies have shown that Atlantic taxa associated with low light are robust to more polar conditions. We compared to which extent sea ice melt influences light-associated phytoplankton dynamics and biodiversity over two years at two mooring locations in the Fram Strait. One mooring is deployed in pure Atlantic water, and the second in the intermittently ice-covered Marginal Ice Zone. Time-series analysis of amplicon sequence variants abundance over a 2-year period, allowed us to identify communities of co-occurring taxa that exhibit similar patterns throughout the annual cycle. We then examined how alterations in environmental conditions affect the prevalence of species. During high abundance periods of diatoms, polar phytoplankton populations dominated, while temperate taxa were weakly represented. Furthermore, we found that polar pelagic and ice-associated taxa, such as *Fragilariopsis cylindrus* and *Melosira arctica*, were more common in Atlantic conditions, while temperate taxa, such as *Odontella aurita* and *Proboscia alata*, were

less abundant under polar conditions. This suggests that sea ice melt may act as a barrier to the northward expansion of temperate phytoplankton, preventing their dominance in regions still strongly influenced by polar conditions. Our findings highlight the complex interactions between sea ice melt, phytoplankton dynamics, and biodiversity in the Arctic.

Personal Contribution

EO conducted the data analyses. WJvA, CB, MW, STV and KM are responsible for the sampling design. STV contributed nutrient data. WJvA contributed oceanographic data. **EO**, KM and OP interpreted the data, conceptualized and drafted the manuscript. **All authors** contributed to improving the final manuscript, by contributions to the scientific interpretation of the data and the discussion of results.

Importance of the Research and Contribution to this Thesis

Therefore, it answers our first research question: Based on the findings of the paper, environmental conditions such as sea ice melt have a significant influence on the prevalence of specific phytoplankton species in the Arctic. The study observed that certain phytoplankton taxa, particularly those associated with polar pelagic and ice-associated environments, were more common in regions influenced by Atlantic water conditions. Conversely, temperate taxa had limited potential to persist in colder, ice-impacted waters. This suggests that as sea ice melts and Arctic waters experience changes in oceanographic conditions, there could be shifts in the composition and distribution of phytoplankton communities. These changes may have implications for the broader Arctic ecosystem, as the ability of different phytoplankton taxa to tolerate varying environmental conditions can affect their persistence and spread in Arctic waters. Overall, the study provides insights into the complex relationship between environmental changes, phytoplankton dynamics, and ecosystem responses in the Arctic Ocean.

ISME

ISME Communications, 2024, **4(1)**, ycae027
https://doi.org/10.1093/ismeco/ycae027
Advance access publication: 27 February 2024
Original Article

Sea-ice melt determines seasonal phytoplankton dynamics and delimits the habitat of temperate Atlantic taxa as the Arctic Ocean atlantifies

Ellen Oldenburg[1,*], Ovidiu Popa[1], Matthias Wietz[2,3], Wilken-Jon von Appen[4], Sinhue Torres-Valdes[4], Christina Bienhold[2,3], Oliver Ebenhöh[1], Katja Metfies[5]

[1]Institute of Quantitative and Theoretical Biology, Heinrich-Heine-University Düsseldorf, Universitätsstr. 1, D-40225 Düsseldorf, Germany
[2]Max Planck Institute for Marine Microbiology, Celsiusstraße 1 D-28359 Bremen, Germany
[3]Deep-Sea Ecology and Technology, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Am Handelshafen 12 D-27570 Bremerhaven, Germany
[4]Physical Oceanography of the Polar Seas, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Am Handelshafen 12 D-27570 Bremerhaven, Germany
[5]Polar Biological Oceanography, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Am Handelshafen 12 D-27570 Bremerhaven, Germany

*Corresponding author: Ellen Oldenburg, Institute of Quantitative and Theoretical Biology, Heinrich-Heine-University Düsseldorf, Universitätsstr. 1, D-40225 Düsseldorf, Germany. Email: ellen.oldenburg@hhu.de

Abstract

The Arctic Ocean is one of the regions where anthropogenic environmental change is progressing most rapidly and drastically. The impact of rising temperatures and decreasing sea ice on Arctic marine microbial communities is yet not well understood. Microbes form the basis of food webs in the Arctic Ocean, providing energy for larger organisms. Previous studies have shown that Atlantic taxa associated with low light are robust to more polar conditions. We compared to which extent sea ice melt influences light-associated phytoplankton dynamics and biodiversity over two years at two mooring locations in the Fram Strait. One mooring is deployed in pure Atlantic water, and the second in the intermittently ice-covered Marginal Ice Zone. Time-series analysis of amplicon sequence variants abundance over a 2-year period, allowed us to identify communities of co-occurring taxa that exhibit similar patterns throughout the annual cycle. We then examined how alterations in environmental conditions affect the prevalence of species. During high abundance periods of diatoms, polar phytoplankton populations dominated, while temperate taxa were weakly represented. Furthermore, we found that polar pelagic and ice-associated taxa, such as *Fragilariopsis cylindrus* and *Melosira arctica*, were more common in Atlantic conditions, while temperate taxa, such as *Odontella aurita* and *Proboscia alata*, were less abundant under polar conditions. This suggests that sea ice melt may act as a barrier to the northward expansion of temperate phytoplankton, preventing their dominance in regions still strongly influenced by polar conditions. Our findings highlight the complex interactions between sea ice melt, phytoplankton dynamics, and biodiversity in the Arctic.

Keywords: climate change, time-series clustering, marine, Arctic aquatic communities, Atlantification, Fourier decomposition

Introduction

The Arctic is affected by rapid and drastic environmental changes. For instance, air temperatures rise four times [1] as quickly in the region compared to other regions on Earth [2]. Arctic sea ice is one of the fastest changing components of the Earth system [3]. Over the past decades, the area of Arctic sea ice declined at a rate of ∼1 million km^2 in area extent per decade [3, 4]. There are indications for a 40% decline in ice thickness due to thicker and older ice cover [5]. The geographical extent of warmer and more saline Atlantic water is expected to expand northwards into the Central Arctic Ocean (CAO), which consequently will become warmer and saltier, further accelerating sea-ice decline [6]. This process, called Atlantification of the Arctic Ocean [6], coincides with altered physical conditions. Ecosystems shift towards a more temperate state including the appearance and range expansion of subarctic specie [7-12]. If the temperature increases and the loss of sea-ice continue at their current pace, the Arctic Ocean will likely be seasonally ice-free by 2050 [13]. In such a scenario, sea-ice melt-related processes, such as melt-water stratification of the upper layer of the ocean, that is currently observed in the marginal ice zone (MIZ), might become more important over more prolonged periods throughout the seasonal cycle, and a larger geographic area, with ecological consequences for the Arctic Ocean. The MIZ is usually covered with 15–80% sea ice [14-20] and its distribution, thickness, and melt dynamics are key drivers of productivity [21], carbon export, biogeochemical cycling, and pelagic-benthic coupling. As a result of decreasing sea ice extent and the expected Atlantification, larger areas of the Arctic Ocean might become favorable for pelagic temperate phytoplankton. As a study site, Fram Strait allows us to investigate the combined

Received: 4 February 2024. **Revised:** 21 February 2024. **Accepted:** 21 February 2024

effects of Atlantification and seasonal ice cover on Arctic marine ecosystems. Moorings with a suite of physical and biogeochemical sensors, as well as autonomous sampling systems for molecular biodiversity studies (Remote Access Sampler RAS), are positioned at two different locations in Atlantic Waters of Fram Strait at ~79°N: central Fram Strait (mooring cluster "HG-IV") and in the eastern Fram Strait (mooring cluster "F4") -see Figure 1. F4 is located in the flow path of the West Spitsbergen Current (WSC). HG-IV is located in the vicinity of the interface between the WSC and the East Greenland Current (EGC). The WSC carries relatively warm and salty Atlantic Water via Fram Strait northwards towards the CAO, while the EGC exports cold ice-covered and less saline Polar Water (PW) from the CAO through Fram Strait. In the vicinity of HG-IV, some of the Atlantic Water (AW) is mixed in an eddy-rich area [22] as part of a subduction process [23, 24] with the outflowing colder and fresher water of the EGC. This area is frequently characterized by major sea-ice melt events, as sea-ice coverage regularly extends [25] into the WSC, which carries temperate species towards the CAO. Thus, ecosystem functionality in the vicinity of the MIZ in the WSC might serve as a model for future biodiversity and ecosystem functionality in a seasonally ice-free CAO impacted by Atlantification and thereby inform on the potential of temperate taxa to thrive in a seasonally ice-covered Atlantic-influenced Ocean [8, 26-28].

Over the past few decades, the transport of sea ice in both volume and velocity towards Fram Strait increased in the area of the Transpolar Drift due to the thinning Arctic pack ice [29-31]. This led to a significant south-eastward extension of the MIZ into Fram Strait during certain years of the past decade. In 2017, the MIZ extended into large parts of the WSC during summer, including the two moorings [31]. Conversely, the 2018 ice export was reduced to <40% relative to that between 2000 and 2017.

The associated meltwater-induced stratification promoted a longer phytoplankton bloom with a relatively shallow extent and reduced export flux [32]. The summer of 2018 had a mixed layer regime (MLR) and a shorter, more intense bloom compared to other periods. During the spring of that year, there was also an increased carbon export to the deep sea [33]. The particularly warm year of 2018 may reflect the conditions of the CAO in the future. The native biodiversity of the communities is a key determinant of whether and how a community or an individual organism can respond to changing abiotic conditions [34]. We, therefore, expect that studying the microbial communities and, in particular, comparing the seasonal dynamics between the years 2017 and 2018 can greatly improve our knowledge about the resilience of pelagic and sympagic organisms and how microbial diversity and seasonality scale with the environmental variability. Molecular biodiversity research using ribosomal meta-barcoding has substantially improved our comprehension of marine microbial diversity and distribution patterns during the last 20 years. [35, 36]. As part of the FRAM Infrastructure Program (Frontiers in Arctic Marine Monitoring) and the long-term ecological research site LTER HAUSGARTEN, activities in Fram Strait provide information on Arctic marine eukaryotic microbial biodiversity and biogeography based on annually recurring measurements (since 1999) recently expanded by year-round, continuous sampling since 2016. We hypothesize that biodiversity and seasonal succession in the Fram Strait are strongly impacted by sea-ice melt and the extent of stratification [37].

In this study, we exploit this wealth of data through a combination of statistical and bioinformatic approaches. The continuous data collected over two years were decomposed using a Fourier transformation into a series of sinusoidal functions. Each function represents a specific amplicon sequence variant (ASV) dynamic over time. By clustering the ASVs based on their seasonal fluctuation patterns, it became possible to analyze the impact of different water regimes that occurred in 2017 and 2018, as reported in Appen et al. 2021 [32], on both species and community levels. We could elucidate the effects of sea-ice melt on the seasonal dynamics of the associated eukaryotic microbial communities as key drivers of phytoplankton bloom phenology. By assessing the contribution of polar and temperate phytoplankton taxa to eukaryotic microbial communities in the WSC over the annual cycle, we infer the potential of polar taxa to thrive in ice-free Atlantic water and temperate taxa to expand to areas impacted by sea-ice melt.

Materials and methods

Sampling

The samples analyzed in this study were collected using McLane Remote Access Samplers (RAS) deployed in conjunction with other oceanographic sensors over three individual annual cycles from June 2016–August 2019 on long-term moorings at stations HG-IV (79.0118 N 4.1666E) and F4 (79.0118 N 6.9648E) of the LTER HAUSGARTEN and FRAM in the Fram Strait [38]. This study covers the period from January 2017 to December 2018, i.e., two calendar years. One RAS was deployed at a depth between 24–29 m at HG-IV and another at 23-26 m - at F4. The RAS samplers contained 48 sterile bags, each collecting water samples of 500 mL at programmed sampling events every two weeks. Samples were preserved by adding 700 μl of half-saturated mercuric chloride (7.5% w/v) to the bags prior to sampling. A sample reflects the pool of up to two samples collected one hour apart in two individual bags. Following the recovery of the RAS devices, water samples were filtered using Sterivex filter cartridges with a pore size of 0.22 μm (Millipore, USA). Filters were then stored at −20°C for later processing.

Mooring and satellite data

Temperature, salinity, and dissolved oxygen concentration were measured with a CTD-O_2 attached to the RAS frame. Physical oceanography sensors were manufacturer-calibrated and processed as described in [39]. 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.pangaea.de/10.1594/PANGAEA.941125. For chemical sensors, raw sensor readouts were used. The fraction of Atlantic and Polar Water were computed for each sampling event following [23] and reported along with distance below the surface (due to mooring blowdown). Sea ice concentration derived from the Advanced Microwave Scanning Radiometer sensor AMSR-2 [40] 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. Similar to van Appen et al. 2021 [41], the analyzed datasets consist of ten environmental values for the two locations, F4 and HG-IV, from 01.01.2017 to 31.12.2018. From this dataset, we retrieved the following variables: water temperature (temp °C), fluorescence chlorophyll concentration from in situ sensor (chl_sens ~μg l^{-1}), daylight (daylight h), water depth (depth m), ice concentration (iceConc %), ice distance (IceDist to 20% ice concentration km), mixed layer depth (MLD m), partial pressure of CO_2 (pCO_2_conc

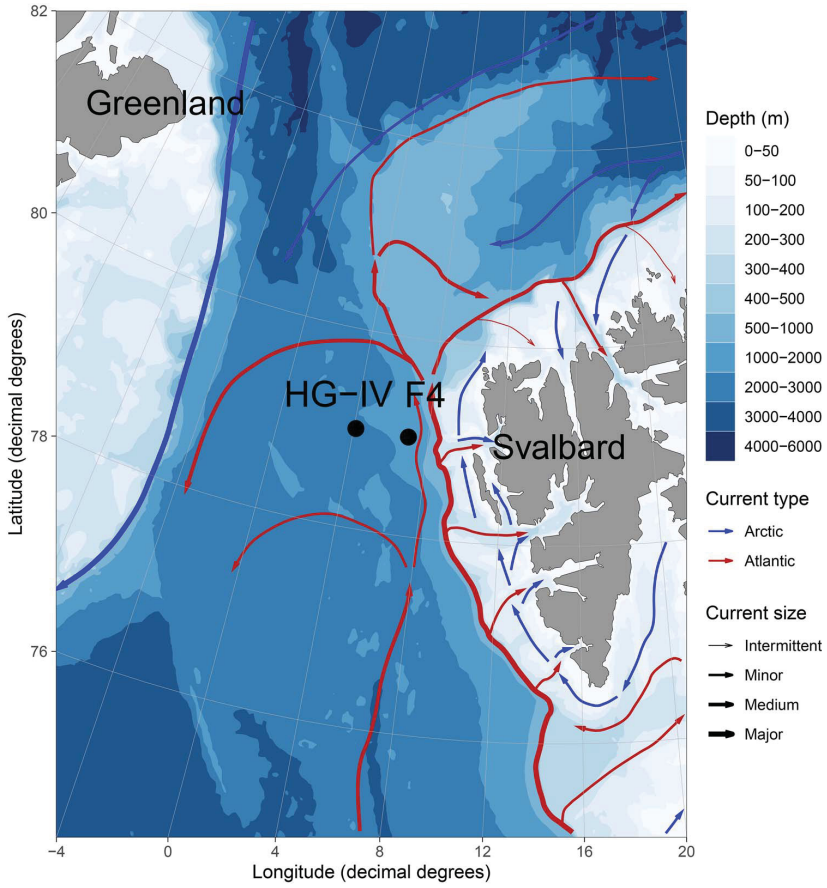

Figure 1. Map of mooring locations, major currents, and water depths in Fram Strait. The main currents in the area are illustrated schematically: West Spitsbergen Current (WSC) in red and East Greenland Current (EGC) in blue. The locations of the moored remote access samplers discussed in this study are marked in black for HG-IV and F4. F4 is located in the WSC and HG-IV west of the WSC. Land is displayed in gray and the different water depths in a white-blue color gradient.

μatm), O_2 concentration (O_2 _conc μmol l^{-1}), polar-water fraction (PW_frac %).

DNA-extraction and Illumina amplicon-sequencing of 18S rRNA genes

Isolation of genomic DNA was carried out using the PowerWater kit (Qiagen, Germany) following the manufacturer's protocol. Obtained DNA was quantified using Quantus (Promega, USA) and stored at −20°C. 18S rRNA gene fragments from the hypervariable V4 region were amplified by polymerase chain reaction (PCR) with primers 528iF (GCGGTAATTCCAGCTCCAA) and 926iR (ACTTTCGTTCTTGATYRR). illuminaNextV4F (TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCGGTAATTCCAGCTCC) and illuminaNextV4R (GTCTCGTGGGCTCG-GAGATGTGTATAAGAGACAGGGCAAATGCTTTCGC) [42]. All PCRs had a final volume of 50 μL and contained 0.02 U Phusion Polymerase (Thermo Fisher, Germany), the 10-fold polymerase buffer according to manufacturer's specification, 0.8 mM each dNTP (Eppendorf, Germany), 0.2 μM L^{-1} of each primer, and 1 μL of template DNA. PCR amplification was performed in a thermal cycler (Eppendorf, Germany) with an initial denaturation (94°C, 2 min) followed by 35 cycles of denaturation (94°C, 20 sec), annealing (58°C, 30 sec), and extension (68°C, 30 sec) with a single final extension (68°C, 10 min). The PCR products were purified from an agarose gel 1% [w/v] with the NucleoSpin Gel Kit (Macherey-Nagel, Germany) and Mini Elute PCR Purification kit (Qiagen, Germany). Subsequently, DNA concentrations were determined using a Quantus Fluorometer (Promega, USA). Prior to library preparation, DNA fragments were diluted with TE buffer to a concentration of 0.2 ng μL^{-1}. Libraries were prepared according to the 16S Metagenomic Sequencing Library Preparation protocol, and sequenced using MiSeq (Illumina, USA) in 2x300 paired-end runs. Sequence data are available under ENA BioProjects PRJEB43889 and PRJEB43890.

Sequence analysis

After primer removal using cutadapt [43], reads were processed into amplicon sequence variants (ASVs) using DADA2 v1.14.1 [39], as described in Wietz et al [44]. Briefly, reads were trimmed based on quality profiles, with filtering settings truncLen = c(250, 200), maxN = 0, minQ = 2, maxEE = c(3, 3), and truncQ = 0. Followed by merging (minOverlap = 20) and chimera removal, reads were

taxonomically classified using PR2 v4.12 [45]. The herein reported data has been processed in the scope of autonomous eDNA biodiversity analyses within the FRAM Observatory, as described under https://github.com/matthiaswietz/FRAM-RAS_eDNA.

Analysis strategy and R packages

All calculations were performed in R version 4.1.3 (2022-03-10). The complete analysis pipeline is available at https://gitlab.com/qtb-hhu/qtb-sda/framstrait_1718. Analysis and plotting tools used for this work are available in a git repository with scripts and an R package. Fourier decomposition was performed with the segmenTier R package [46], available at https://cran.r-project.org/package=segmenTier. The dynamics of eukaryotes were analyzed using the Fourier-transformed time series signals of the relative abundance information. As part of biodiversity, relative species abundance refers to the extent to which a species is common or rare relative to other species in a particular location or community [47]. Relative abundance is the percentage composition of an organism of a given species relative to the total number of organisms in that habitat. The data were interpolated on daily bases.

Time series analysis

The use of Fourier decomposition for time series signals is a common technique to obtain temporal profiles of data that contain seasonal patterns. In this study, we used this technique to identify and describe the seasonality of several species, as also described in Priest et al. [48]. For each amplicon sequence variant we extracted the time series signal from the relative abundance data using a Fourier approach implemented in the R package segmenTier / segmenTools [49]. The Fourier technique is decomposing signals into the sum of their frequency components, characterized by sine and cosine functions. The Fourier Theorem states that any function can be rewritten as the sum of sinusoidal functions. The approximation becomes more accurate with each additional series element. These elements are called Fourier components.

A measurement for seasonality s for the times series t was calculated by the following formula:

$$s(t) = \frac{|f_2(t)|}{|f_0(t)|}, \tag{1}$$

where f_i is the i-th fourier component of the times series t and $|\cdot|$ is the absolute value function [50, 51].

After the Fourier transformation, the frequency, amplitude, and phase information of each particular ASV time signal was extracted. These values indicate the seasonality, abundance strength, and time of occurrence within the measured period.

Cluster definition

Species with similar temporal pattern were grouped into co-occurrence clusters. The choice for the parameter N = 10, the number of clusters for both locations, was chosen to keep the cluster comparable. The metric (Bayesian Information Criterion - BIC) of the applied clustering algorithm proposes a value around 9 and 10 as the optimal cluster number. Groups of species with similar time signals were identified by a clustering approach in the segmenTools R package [49]. The significance of overlapping clusters (shared members by two clusters), illustrated as a color gradient, is calculated based on the negative logarithm of the p-value and the number of overlapping features. All identified clusters were classified into low-light, high-light, and mixed-light clusters depending on the light conditions in which their members show the highest abundance. Further, all clusters were named depending on the mooring (H for HG-IV and F for F4) and numbered in ascending order depending on the phase of the sinusoidal function, which was calculated for each cluster from the average of the cluster members. Therefore, the order of the numbers indicates the order of occurrence within the year.

Co-occurrence of ASVs

In contrast to earlier investigations that depended on Pearson correlation for pairwise comparisons of relative abundance values to deduce co-occurrence patterns our methodology utilized Fourier decomposition of time series data [52-54]. This allowed the extraction of unique temporal profiles for each Amplicon Sequence Variant (ASV). By applying correlation analysis to these individual profiles, we effectively mitigated the inherent bias associated with utilizing Pearson correlation on compositional data [55] .

Conditions preference

To assess the population's annual abundance, we computed the sum of relative abundances for each Amplicon Sequence Variant (ASV) within a specified timeframe. Total abundance values were separately calculated for the F4 and HG-IV locations. Subsequently, entries with zero abundance were excluded to prevent division by zero, and we determined the abundance quotients for 2017 and 2018, as well as the reverse calculation. The log2(quotient) values were categorized as meltwater regime (MWR) or MLR based on whether they were greater than or equal to 1 or less than or equal to −1, respectively. ASVs not meeting either condition were assigned to the unspecified group. To gauge the dissimilarity between locations in a given year for a specific group of ASVs, we defined four quotients as follows:

$$p(x,y) = \frac{\mathrm{MWR}|_x}{\mathrm{MWR}|_y}, \qquad x \in X, y \in Y, \tag{2}$$

$$t(x,y) = \frac{\mathrm{MLR}|_x}{\mathrm{MLR}|_y}, \qquad x \in X, y \in Y, \tag{3}$$

where X = {F417, F418}, Y = {HG-IV17, HG-IV18}, and MWR (MLR) containing all MWR (MLR) ASVs relative two-year abundances. The restriction is defined by selecting only the ASV abundances from the given time and location.

- p(F42017, HG-IV2017) correspond to the ratio of F4 to HG-IV for species preferring the MWR in 2017.
- p(F42018, HG-IV2018) correspond to the ratio of F4 to HG-IV for species preferring the MWR in 2018.
- t(F42017, HG-IV2017) correspond to the ratio of F4 to HG-IV for species preferring the MLR in 2017.
- t(F42018, HG-IV2018) correspond to the ratio of F4 to HG-IV for species preferring the MLR in 2018.

To compare how much the MWR is favoured on average versus a MLR within a given site, we define the following equations:

$$q(z) = \frac{\frac{1}{|\mathrm{MWR}|_z|}\sum_{i\in \mathrm{MWR}|_z i}\mathrm{MWR}|_z}{\frac{1}{|\mathrm{MLR}|_z|}\sum_{h\in \mathrm{MLR}|_z h}\mathrm{MLR}|_z}, \qquad z \in Z, \tag{4}$$

where Z = {F417, F418, HG-IV17, HG-IV18}, and MWR (MLR) containing all MWR (MLR) ASVs relative two-year abundances. The restriction is defined by selecting only the ASV abundances from the given time and location and ${}_i\mathrm{MWR}$ (${}_h\mathrm{MLR}$) is the i-th (h-th) relative two-year abundance from the MWR (MLR) ASV.

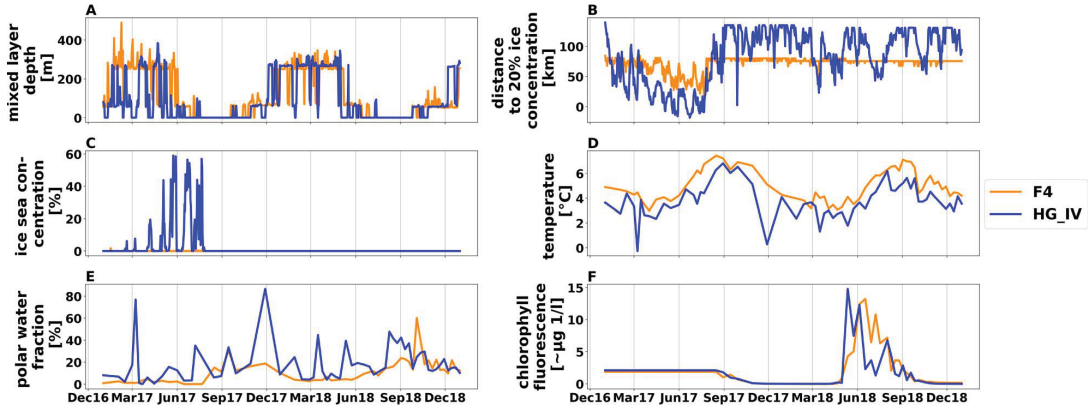

Figure 2. Environmental data for the F4 (dark orange) and HG-IV (blue) location from 2017 to 2018. The x-axis indicates the period from 01.01.2017 to 31.12.2018. The y-axis indicates: **A:** Mixed layer depth (Minimum of the estimated MLD) [m] **B:** Distance to 20% ice concentration (*) [m] **C:** Sea ice concentration [%] **D:** Temperature [°C] **E:** Polar water fraction [%] **F:** Chlorophyll a concentration (**) [μL–1] *Negative values indicate that the ice edge is south east of the mooring points at the blue curve March 2017 to September 2017) **Sensor did not work before August 2017.

- q(F42017) corresponds to the ratio for meltwater preference over mixed-layer in 2017 at station F4.
- q(F42018) corresponds to the ratio for meltwater preference over mixed-layer in 2018 at station F4.
- q(HG-IV2017) corresponds to the ratio for meltwater preference over mixed-layer in 2017 at station HG-IV.
- q(HG-IV2018) corresponds to the ratio for meltwater preference over mixed-layer in 2018 at station HG-IV.

Cross-condition analysis

To investigate how the dynamics of a particular ASV with a preference for a specific water regime change under the conditions of the opposite water regime, we determined and compared the area under the curve (AUC) from the relative abundance within a time range of 365 days. For that, we used on a daily level interpolated abundance data to which we applied a polynomial function and calculated the AUC for each year separately. Afterward, we compared the ratio of the AUC values between the years to illustrate prosperity differences that are related to the environmental conditions of the individual year.

Results and discussion

Environmental conditions

A pronounced extension of the ice edge/MIZ into the WSC during the first half of 2017, compared to 2018, led to different environmental conditions in this part of the eastern Fram Strait. That MLR was similar to that expected for a seasonally ice-free Arctic Ocean, impacted by Atlantification. More specifically, eastern Fram Strait experienced extended sea ice melt during spring and early summer 2017. According to van Appen et al. 2021 [32], there were significant differences in environmental conditions between 2017 and 2018, with station HG-IV exhibiting more pronounced differences compared to the pure Atlantic Water station F4. This is best reflected by variability in the fraction of Polar Water, distance to the ice edge, ice concentration, and water column stratification (Fig. 2).

At HG-IV, the mixed layer depth was overall shallower from January to May 2017 compared to 2018 and F4 due to higher ice concentrations. Moreover, HG-IV was frequently impacted by the intrusion of Polar Water (PW) throughout the annual cycle, which is common for this region. Higher fractions of PW were observed for the period's March, July to August, and November–December of 2017 compared to 2018, according to the RAS data. The intrusion of PW led to lower water temperatures. At HG-IV, temperatures were lower in spring 2017 compared to 2018—ice distances, defined as the distance to 20% ice coverage. At HG-IV, the distance to the ice edge was shorter in 2017 than in 2018 until August but was similar during the remaining months (Figure 2). From mid-August to November; water temperatures were higher in 2017 compared to 2018. In 2017, there was higher ice cover in Fram Strait and subsequent ice melt, resulting in a highly stratified melt water regime (MWR). In contrast, in 2018, an unstratified mixed layer dominated regime (MLR) was present [32].

At F4, ice distances were not significantly different between the two years. However, water temperatures were higher in 2017 compared to 2018 from mid-August to November. In this investigation, Station F4 serves as a reference for typical Atlantic environmental conditions for both years.

In the following section, we examined the behavior of eukaryotic microbes under distinct water regimes, namely meltwater and mixed layer conditions. To achieve this, we employed a top-down structure to delineate the temporal abundance changes for: (i) all ASVs, (ii) specific ASV clusters, and (iii) individual representative species.

Preference of eukaryotic microbes for the different water regimes

There is a remarkable similarity in species composition between the two stations. A total of 50% (583) of all ASVs under inspection were detected at both stations, which we refer to as the core community. In contrast, 22% were unique to F4 (254 ASVs) and 28% to HG-IV (320 ASVs) (Figs S5 and S2). To determine the preferred water regime for microbial eukaryote taxa, we calculated the total relative abundance of each ASV per year and compared them between both years. This comparison was only possible at station HG-IV due to the differing conditions in both years. To achieve this, we sorted the ASVs into three groups based on the preferred water regime: the unstratified MLR, the highly stratified MWR, and an unspecified group. The MLR group comprises all temperate taxa,

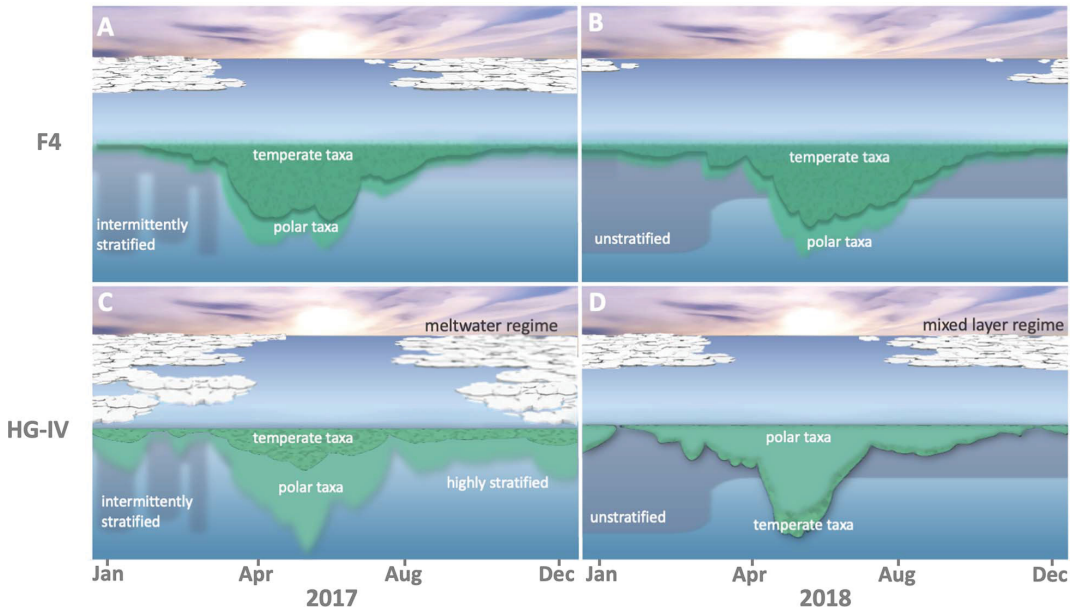

Figure 3. Effects of meltwater and mixed layer conditions on temperate (dark green) and polar (light green) taxa. The x-axis shows the months January through December from 2017 through 2018. The green areas reflect the relative abundances of temperate (dark green) and polar (light green) taxa. Since the data is relative, no quantification is given on the y-axis. The relative abundance curves of A and B were derived from water column samples from cluster F-06, and C and D from cluster H-06. **A:** Polar and temperate taxa are observed in similar abundances in the highly stratified meltwater regime at F4 in 2017. **B:** Similar abundances for polar and temperate taxa in the mixed layer regime at F4 in 2018. **C:** Reduced abundance of temperate taxa in the meltwater regime with high stratification at HG-IV in 2017. **D:** Reduced abundance of polar taxa in the mixed layer regime at HG-IV in 2018.

which were twice as abundant in HG-IV-2018 compared to HG-IV-2017 (n = 67 [11.49% of the core community]). In contrast, ASVs that were twice as abundant in HG-IV-2017 compared to HG-IV-2018 belong to the MWR group (n = 94 [16.12% of the core community]), which are referred to as polar taxa. The remaining ASVs were classified as an unspecified group (n = 422 [72.38% of the core community]). In the following steps, we focused on species that are sensitive to one of the water regimes that occurred. Notably, we identified 161 species in this study that showed a preference for a specific regime. These species were distributed between the MLR group (41.62%) and the MWR group (58.38%) (Table S5]).

Cross spatio-temporal comparison

We compared both groups (MLR & MWR) to identify differences attributed to either location, HG-IV vs. F4 (Fig. 1), or the varying conditions between 2017 and 2018 (Fig. 3). To do so, we conducted two types of comparisons: (i) within each year, we compared the stations to each other and (ii) within each station, we compared the data from 2017 and 2018. First, we compared the relative abundance differences in 2017 between stations. We calculated the median of the MLR group and MWR group, respectively, and compared them. Our results showed that the median differences between the locations (Fig. 3A-D) of species favouring mixed-layer were 1.54 times larger than the median differences of the species favouring meltwater in 2017 (Table S5; see methods formula (2,3)). Furthermore, we confirmed this observation regarding the different medians by comparing the relative abundances of each ASV member in the aforementioned groups (one-sided Kolmogorov–Smirnov test *P*-value: 3.13E-05). In the next step, we repeated the same analysis for the year 2018.

In contrast to 2017, the median differences in 2018 of the meltwater-favouring species were 2.78 times greater than the median differences of the mixed-layer favouring species (Table S5; see methods formula (2,3)). Also, in this case, comparing the relative abundance of the particular ASVs could support this observation (one-sided Kolmogorov–Smirnov test *P*-value: 1.376E-14). Once we had distinguished dissimilarities among the stations, our attention turned to describing dissimilarities over the years (Fig. 3 A-D). This was motivated by the different water regimes observed in 2017 and 2018 [32]. Consequently, this examination enabled us to demonstrate how species abundance is influenced by varying environmental circumstances. Therefore we compared the relative abundance ratio of each group (MLR, MWR) between years (2017 vs. 2018). The difference between the two years (2017 and 2018) for each group was less significant at station F4 (MWR = 1.23 and MLR = 0.60), whereas at HG-IV, the discrepancy was approximately four times higher than that observed at F4 for the same years (MWR = 2.13 and MLR = 0.27), see Table S5. As a result for the following analysis, we used station F4 as a reference for constant environment because it is less influenced by meltwater conditions. In contrast, the HG-IV location offers the opportunity to study the effects of Atlantification in a seasonally ice-covered Arctic Ocean, conditions that are expected for the CAO in the near future [56]. For that, we examined how each other's water regimes affected the relative abundance of the respective ASV. We aimed to determine whether polar or temperate ASVs were more resilient to the opposing condition. For the analysis, we specifically selected ASVs that are known to grow in polar or temperate conditions [57-64].

Seasonal succession of eukaryotic microbes

To understand the seasonal succession of eukaryotic microbes, we analysed the phases obtained from the sinusoidal function after Fourier transformation. This allows us to determine the

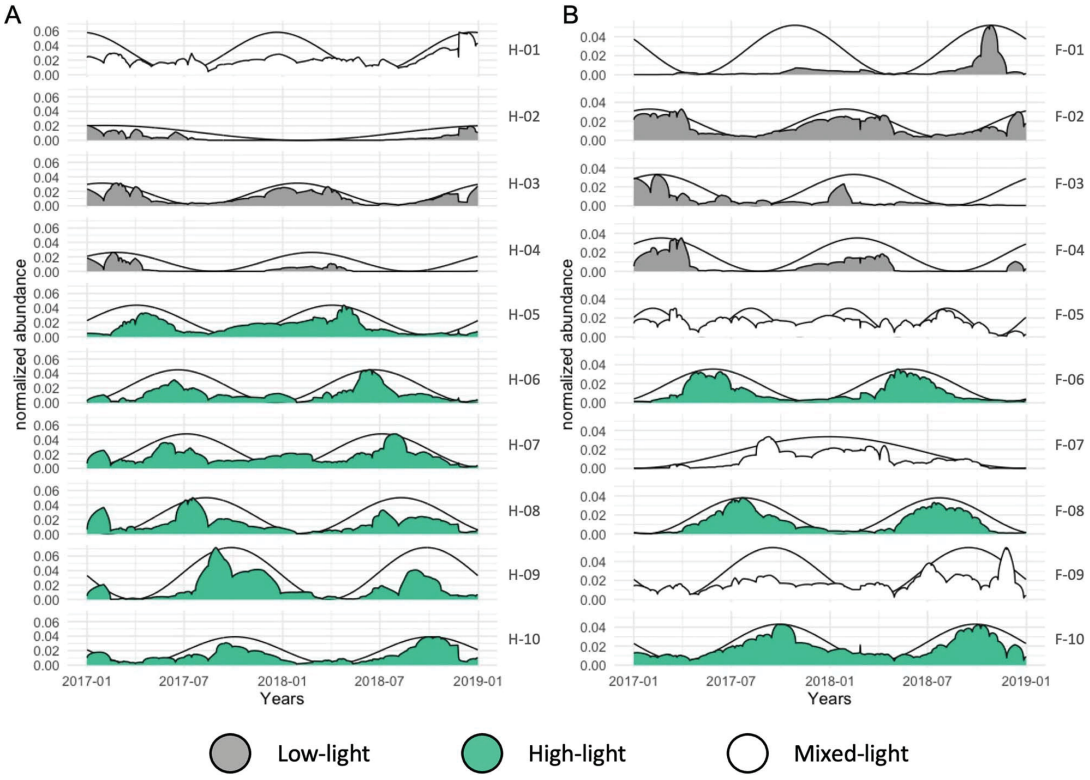

Figure 4. Time-Series Clustering for both moorings spanning the years 2017-2018. The x-axis indicates the period from 01.01.2017 to 31.12.2018. Black sinusoidal curves show the predicted seasonality of the entire cluster based on the dominant Fourier component. The respective relative abundance is shown for each cluster on the left yaxis. Cluster names are shown on the right. The clusters are sorted by phase which illustrates the time of maximal abundance of each community. Clusters are colored according to the three classes HL (green), LL (grey), and NA (white) introduced in the text. **A:** HG-IV, **B:** F4.

chronological timeline of the species in this region. Ten clusters of seasonally synchronized and ordered occurrences of eukaryotic microbial species were identified through community detection analysis of time-series data from the F4 and HG-IV moorings, which included 837 and 903 ASVs, respectively (Fig. 4, Table 1). The frequency obtained from the sinusoidal function (light grey) shows the number of high abundance periods of each community per year. Most clusters (85%) had two maxima, indicating that most organisms exhibit a seasonal occurrence with the highest abundance once a year (Fig. 4, Table 1). We divided the clusters based on their high abundance period into two classes of light conditions. The low-light (LL 0–2 hours sunlight per day) clusters include species with a high abundance phase in the low-light period from October to March when water temperature and distance to the ice edge are low. The high-light class (HL 2–24 hours sunlight per day) includes clusters, in which the high abundance phases coincide with the high-light period from March to October. All other clusters are collected in the mixed light (NA) class. This distinction allowed us to test the succession of the organisms regarding environmental factors per light condition separately. To investigate the commonalities and differences between the two moorings, we compared the species distribution in terms of abundance and seasonality. This analysis also enabled us to assess the succession and prosperity of common species in relation to the varying water regimes. In addition, we compared the time series cluster composition from HG-IV and F4 with each other to identify overlapping communities between both locations. For example, the similarity in cluster composition between the two moorings was highest during the high-light period, particularly between clusters H-06 and F-06 and clusters H-08 and F-08 (Fig. 5). The presence of these common ASVs at both mooring sites can be explained by a similar trend in the transportation of temperate organisms through the northward-flowing warmer Atlantic and the transportation of polar organisms through the intrusion of polar water from EGC. This pattern was also observed for zooplankton [65, 66]. On the other hand, the varying quantities of ASVs reaching each station because of variations in the influence of the two currents may also explain the biodiversity observed at these two locations (Fig. S1).

Low-light period

During the low-light period from October to March, four distinct clusters (F-01, F-02, F-03, F-04 at F4; H-02, H-03, and H-04 at HG-IV) exhibited an ordered appearance, collectively representing around 40% of the total ASVs and 50% of the total reads at both stations. The clusters were dominated by heterotrophic Dinoflagellates, parasitic Syndiniales, and other small heterotrophic flagellates like MAST and Picozoa (Fig. S4). This composition aligns with previous reports of microbial diversity during the low-light period in the Arctic Ocean [66-68], possibly linked to feeding

Table 1. Cluster overview with the 10 clusters for the moorings F4 and HG-IV.

Name	Type	#Peaks	cl_size	cl_size %	s-score	AUC17	AUC18	AUC17/18	AUC18/17	MS(abs)	MS(rel)
H-01	NA	2	72	9	0.18	6.9837	8.8812	0.7863	1.2717	12	16.67
H-02	LL	1	68	8	0.39	1.81	1.0084	1.7949	0.5571	52	76.47
H-03	LL	2	151	18	0.44	4.4387	4.1164	1.0783	0.9274	41	27.15
H-04	LL	2	50	6	0.83	1.7481	0.7188	2.432	0.4112	28	56
H-05	HL	2	76	9	0.33	5.2585	5.1262	1.0258	0.9748	37	48.68
H-06	HL	2	87	10	0.41	4.1827	4.7476	0.881	1.1351	21	24.14
H-07	HL	2	65	8	0.2	6.0155	6.5539	0.9179	1.0895	13	20
H-08	HL	2	113	14	0.32	6.6398	4.405	1.5073	0.6634	23	20.35
H-09	HL	2	33	4	0.53	7.9501	4.5641	1.7419	0.5741	9	27.27
H-10	HL	2	122	15	0.41	5.0174	5.0116	1.0012	0.9988	18	14.75
F-01	LL	2	27	3	0.66	0.6775	2.8243	0.2399	4.1687	26	96.3
F-02	LL	2	144	16	0.39	5.4081	5.1656	1.0469	0.9552	48	33.33
F-03	LL	2	33	4	0.46	3.289	1.228	2.6783	0.3734	18	54.55
F-04	LL	2	168	19	0.75	2.9621	1.8344	1.6148	0.6193	94	55.95
F-05	NA	1	61	7	0.06	6.194	5.2571	1.1782	0.8487	11	18.03
F-06	HL	2	109	12	0.58	3.9653	4.1903	0.9463	1.0567	37	33.94
F-07	NA	4	53	6	0.13	3.4143	3.529	0.9675	1.0336	24	45.28
F-08	HL	2	142	16	0.58	4.8684	4.4017	1.106	0.9041	26	18.31
F-09	NA	2	36	4	0.17	5.6129	7.6206	0.7365	1.3577	8	22.22
F-10	HL	2	130	12	0.34	7.0415	7.0604	0.9973	1.0027	28	21.54

The cluster names, light types (high-light (HL), low-light (LL), mixed-light (NA)), the number of peaks and the total cluster size of ASV and the percent size, the s-score that measures the seasonality, the area under the curve (AUC) for both years (see methods), the quotients of those years and the number of ASV that only occur on this mooring: absolute (MS(abs)) and relative values (in %) (MS(rel)) (MS: mooring specific).

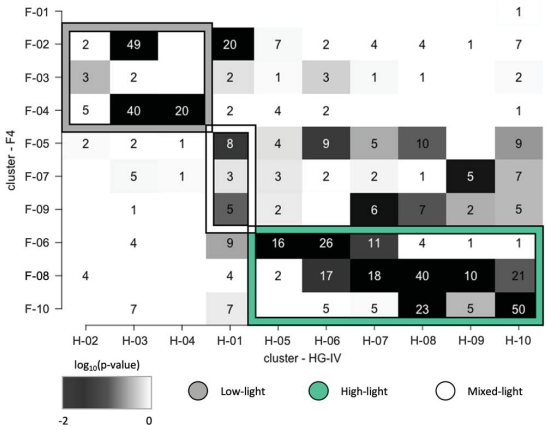

Figure 5. Cluster overlap between F4 and HG-IV locations. The clusters of F4 are plotted on the y-axis against the clusters of HG-IV. The numbers inside the boxes indicate how many ASVs are shared between two clusters The clusters of each location are sorted according to their classes: low-light (grey box frame), mix-light (white box frame) and high-light (green box frame) from top to bottom (F4) and from left to right (HG-IV). The background color of the boxes shows the significance of the overlap from dark (highly significant) to white (non significant).

on bacteria [67]. Notably, diatom ASVs were present in all low-light clusters, exhibiting substantial relative abundances, with higher proportions at HG-IV compared to F4 (Fig. S4). These diatoms, including ice-associated genera such as *Melosira arctica*, *Navicuales sp.*, or *Attheya sepentrionalis* (Fig. S4 Table S4), are adapted to low light and colder temperatures [69] or residing under the ice [70]. The source of these diatoms in the water column during winter at HG-IV is attributed to physical exchange processes at the water-sea ice interface and advection. The persistence of diatoms, particularly Bacillariophyceae, during the polar night in ice-covered waters has been observed previously [67] and their survival strategies, possibly involving resting stages like spores or cysts [71]), influence the composition of Arctic phytoplankton during early spring. Thistaxon-specific survival contributes to diatoms gaining a competitive advantage in the Arctic phytoplankton community when sunlight returns, facilitated by their chlorophyll storage throughout the polar night [72].

High-light period

The high-light period (March to October), distinct clusters (F-06, F-08, F-10 at F4;H-05, H-06, H-07, H-08, H-09, H-10 at HG-IV) sequentially emerged, collectively constituting ~50% of mooring-specific ASVs (Table 1). The community composition of the earlier high-light clusters in 2017 at HG-IV resembled that of early high-light in 2018 at F4 (Fig. 5), suggesting a shared community initiation (Table S3, Table S4). Throughout this period, diatoms, alongside dinoflagellates and other autotrophic taxa, were prevalent (Fig. S4, Table S3). Diatom sequences exhibited a sequential appearance during spring, aligning with Arctic diatoms like *Fragilariopsis cylindrus*, *Bacillaria paxilifer*, *Chaetoceros neogracilis*, and *Grammonema striatula* [73-75] (Table S3) Their major contribution to the pelagic spring bloom emphasized the polar character of the spring bloom community at HG-IV [7, 27, 76]. Notably, *Grammononema striatula* and *C. neogracilis*, polar taxa, were abundant in the first high-light cluster (F-06) at. In contrast to the Arctic diatoms dominating the spring bloom at HG-IV, the temperate diatom *Odontella aurita* [77] ranked among the five most abundant diatoms in the early spring cluster at F4. This suggests the influence of Atlantic Water, transporting organisms from warmer, temperate waters (Fig. 6). *O. aurita*, a key contributor to spring blooms in the German Bight [78], further supports the idea that it thrives in warm, nutrient-rich waters.

Differences in diatoms community composition between F4 and HG-IV became more pronounced in the late summer clusters (H-08 and F-10), which peaked after July. Cluster H-08 at HG-IV was dominated by sea-ice-associated diatoms like *Melosira arctica* and related taxa [70], comprising 57% of the total diatom abundance in this cluster (Table S4 H-08). In contrast, the late cluster F-10 at

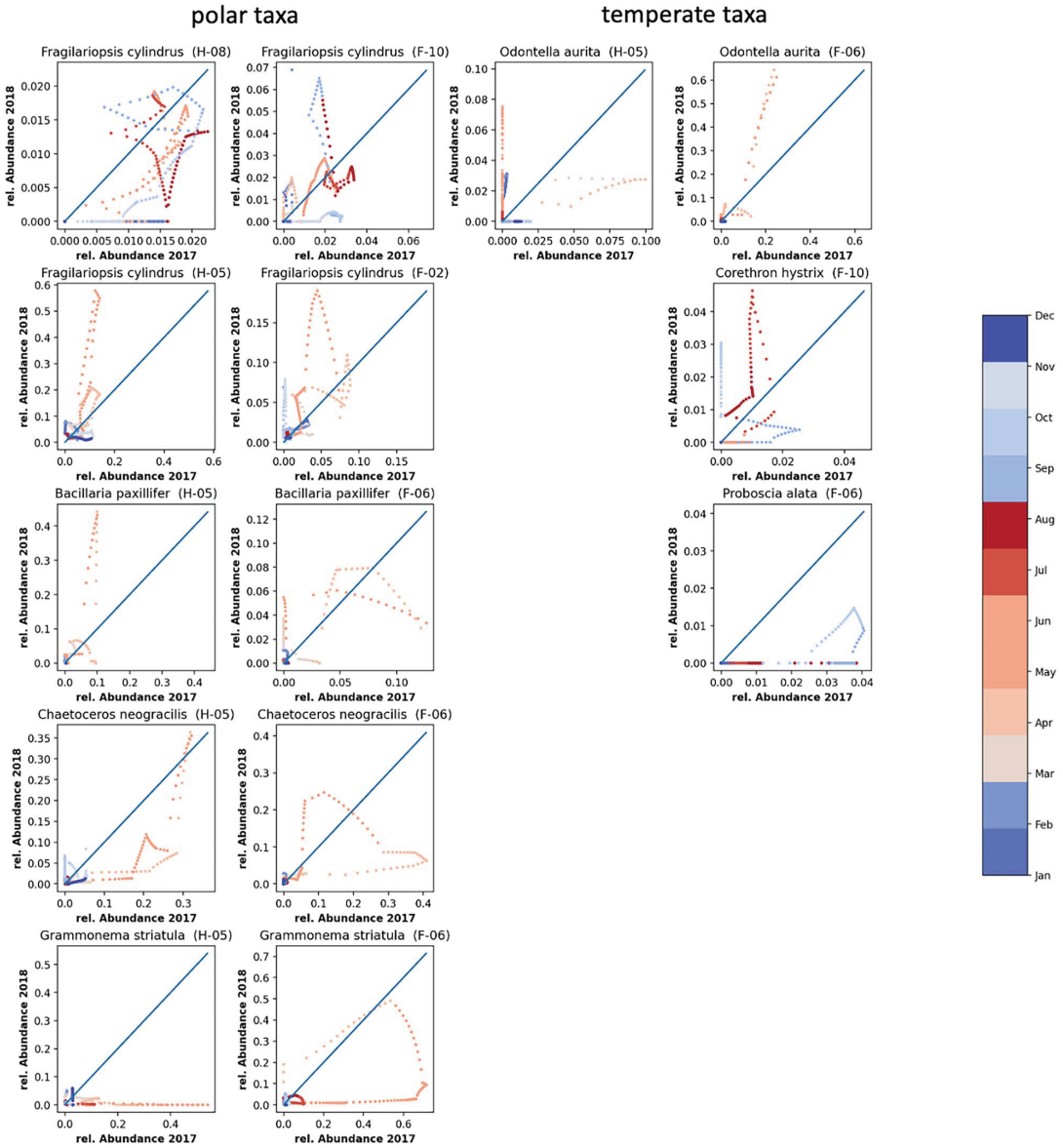

Figure 6. Correlation between the relative abundances of selected ASVs in 2017 vs 2018: The diagonal (blue line) indicates the line on which abundances in 2017 and 2018 would be identical. On the left side (first and second columns) selected polar taxa are displayed, where the first column shows the species at HG-IV and the second column the same ASV at F4. The right side shows selected temperate taxa, where the third column displays species at HG-IV and the fourth column the same ASV at F4. The dots indicate: *Fragilariopsis cylindrus* (ASV207: H-08, F-10), *Fragilariopsis cylindrus* (ASV16: H-05, F-02), *Bacillaria paxillifer* (ASV98: H-03, F-06), Chaetoceros neogracilis (ASV17: H-05, F-06), *Grammonema stratula* (ASV33: H-05, F-06), Odontella aurita (ASV96:H-05, F-06), Corethron hystrix (ASV172: F-10), *Proboscia alata* (ASV947: F-06), Color bar and colored dots indicate month of the year from blue (winter) to red (summer).

F4 was dominated by *Pseudonitzschia sp.* representing 38% of the total diatom abundance (Table S4 F-10). The late summer cluster F-10 at F4 also contained significant quantities of *Corethron hystrix* and *Proboscia alata*, two diatoms thriving in temperate waters [79, 80] highlighting the influence of Atlantic Water on the diatom community at F4. Cluster H-09 at HG-IV, representing 28% of the total abundance of diatoms (Table S3), was dominated by the genus Pseudonitzschia, known for year-round blooms with peaks in late August or early September [81]. Studies have shown that this diatom undergoes blooming throughout the year, typically exhibiting a minor bloom in June, followed by a more substantial bloom in late August or early September [81]. Other major Arctic pelagic autotrophs, such as *Phaeocystis sp.*, *Chaetoceros socialis* and *Micromonas sp.*, were predominantly found in clusters with high

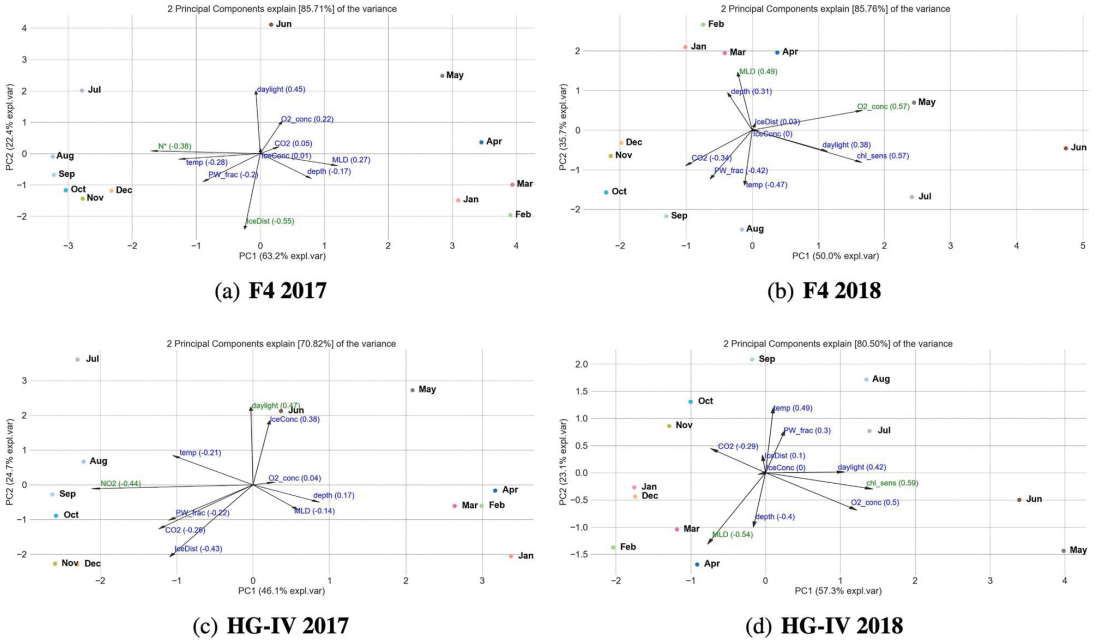

Figure 7. Biplots for both mooring and both years: First PCA component plotted versus the second PCA component of the abundance values for the ASV and the environment conditions on monthly bases aggregated. The arrows correspond to the most relevant features. The length of the arrow corresponds to the feature importance associated with this arrow. **A:** HG-IV 2017, **B:** HG-IV2018, **C:** F4 2017 and **D:** F4 2018.

light levels (Table S3). Specifically, *Phaeocystis pouchetii* was most abundant in early spring clusters (H-05 and F-06) accounting for 16% and 11% of the total abundance, respectively (Table S3 H-05, F-06), consistent with findings from the Western Spitsbergen Current (WSC) and under the ice north of Svalbard [82, 83].

Impact of sea-ice melt on seasonal phytoplankton dynamics and consequences for bloom phenology in Atlantic waters

The different environmental conditions observed in 2017 and 2018 did not seem to affect the order of the annually recurring community clusters at F4 and HG-IV. Instead, changes in environmental conditions resulted in differences in their persistence, abundance amplitude, and integrated abundances (Fig. 4 A and B). At F4, environmental conditions during the high light periods of 2017 and 2018 were similar. In consequence, the integrated seasonal cluster abundance, reflected by the area under the curve, did not significantly change from one year to the other (Table 1). In contrast, we observed differences between both years for HL and LL periods at HG-IV (Fig. 7). According to our data, the changes in environmental conditions, associated with sea-ice melt in spring and summer 2017 at HG-IV, might have significantly affected the communities during high-light periods. For example, these changes can be observed in the high-light cluster H-09 (Table 1). The last period of the cluster (2018) shows a 1.7-fold decrease in abundance compared to the first period (2017). Despite the area under the curve of the early high-light clusters (H-05, H-06, and H-07) showing almost no difference between the two years at HG-IV, the amplitude was much lower in 2017 compared to 2018 (Table S2). This observation suggests that the growth rates in 2017 were lower. It is important to note that the organism abundances only reflect relative proportions of the filtered samples. However, in 2017 the RAS was below the productive layer for at least the first half of the high-light period [35], which may explain the lower relative abundances.

Polar pelagic taxa, such as *C. neogracilis* and *Grammonema striatula*, were dominant (compared to other Bacillariophyta) in the first clusters of the high-light period at both stations (H-05, F-06, Table S3). These species are more robust to variation in ice coverage. In contrast, the contribution of *Fragillariopsis cylindrus* to the spring cluster H-05 was greater at HG-IV than at F-06, as indicated in Table S3. During the spring of 2017 at HG-IV, lower relative abundances of *F. cylindrus* may suggest lower growth rates, which could be attributed to higher ice coverage at this station. *F. cylindrus* and *Bacillaria pacillifer* were among the ASVs with the ten highest relative abundances at both stations. They had higher relative abundances during the spring at HG-IV compared to F4 in the observation period, as shown in Table S4 and Fig. 6. This was likely because they benefited from lower ice concentrations and comparatively higher water temperatures at HG-IV during the spring of 2018 compared to 2017 (Fig. 6). This observation suggests that these polar taxa are not strictly dependent on polar conditions and can tolerate or benefit from Atlantic influence.

O. aurita, a temperate taxon occurring at both stations, benefits at both stations from warmer temperatures. The contribution of this temperate species in cluster H-05 was negligible, accounting for only 1% of total abundance, with a further decrease in 2017 to 0.81%, indicating that it struggles to thrive under the ice. In contrast, at mooring F4, its contribution to the spring cluster F-06 was high in both years (Table S3) as temperatures were in a similar range. During the later part of the season in HG-IV, the area under the curve of cluster H-08 showed a 1.5-fold increase in 2017 compared to 2018, as indicated in Table 1. This cluster mainly comprised typical sea-ice-associated diatoms like

M. arctica, *Fragillariopsis sublineata* and *-cylindrus*, and *Chaetoceros rostratus*. Interestingly, these diatoms did not contribute significantly to the phytoplankton community at F4 during the same time of the year. This indicates a sea-ice melt-related release of sea-ice-associated taxa. The environmental conditions existing at this time, especially meltwater stratification, promoted their bloom in the Atlantic Water of Fram Strait (Tables S3 and S4).

During the specified time frame, there was a notable decrease in the prevalence of polar spring phytoplankton species at the start of the season, accompanied by a corresponding increase in the abundance of ice-associated phytoplankton species during the autumn of 2017. It is worth noting that the peak abundance of ice-associated phytoplankton species usually occurs later in the season in the CAO [84-86]. Ice-associated phytoplankton is less present at HG-IV in 2018 (ice-free year) and does not significantly contribute to the autumn community at ice-free station F4 in either year.

Conclusion

In this study, we compared the dynamics of phytoplankton ASVs from two locations in the Fram Strait (moorings HG-IV and F4) as recorded in 2017 and 2018 (Fig. 3). Although data from only 2 years are not necessarily representative of the long-term development of environmental parameters, these particular years exhibit conditions that make them appear ideal for comparing current conditions with those expected in the future in an Atlantified CAO. This comparison supports a new perspective on how the eukaryotic microbial community in the Central Arctic Ocean might change in the near future. Climate change will likely lead to an ice-free Central Arctic Ocean in summer but ice-covered in winter, as suggested by some climate model scenarios [13].

In our analysis, we could show that a MWR can strongly influence arctic micro-eukaryotes on several levels and that phytoplankton bloom phenology in 2017 is a result of an increased sea ice melt [32]. We could extend previous observations about the influence of sea-ice melt on community dynamics and carbon export. We propose that sea ice melt and the resulting environmental conditions are putative key drivers of microbial eukaryotic community composition and bloom phenomenology. Our observations suggest that polar pelagic and ice-associated taxa (such as *F. cylindrus* or *M. arctica*) are relatively tolerant of more Atlantic oceanographic conditions. In contrast, temperate taxa (such as *O. aurita* or *P. alata*) have limited potential to persist in colder ice-impacted waters. Thus, we hypothesize that sea-ice melt in the MIZ may hinder the northward expansion of temperate Atlantic taxa towards the CAO. This trend will continue even as Atlantic oceanographic conditions move further northwards.

Acknowledgements

We thank Theresa Hargesheimer, Jana Bäger and Daniel Scholz for technical support of the RAS deployment. 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. Our special thanks go to S. Neuhaus for bioinformatic support, K. Oetjen and S. Ziemann for excellent technical support in the laboratory, Eva-Maria Nöthig for critical reading of the manuscript, Martina Löbl for coordination of the FRAM-project.

Supplementary material

Supplementary material is available at *ISME Communications* online.

Author contributions

EO conducted the data analyses. WJvA, CB, MW, STV and KM are responsible for the sampling design. STV contributed nutrient data. WJvA contributed oceanographic data. EO, KM and OP interpreted the data, conceptualized and drafted the manuscript. All authors contributed to improving the final manuscript, by contributions to the scientific interpretation of the data and the discussion of results.

Conflicts of interest

The authors declare no competing interests.

Funding

This study was accomplished in the framework of the HGF Infrastructure Program FRAM and institutional funds of the Alfred-Wegener-Institute Helmholtz Centre for Polar and Marine Research. This work was further 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.).

Data availability

Raw data can be obtained from the European Nucleotide Archive (ENA) at https://www.ebi.ac.uk/ena/browser/view/PRJEB43905. The datasets generated during and/or analysed during the current study are available in the gitlab repository, https://gitlab.com/qtb-hhu/qtb-sda/framstrait_1718.

References

1. Rantanen M, Karpechko AY, Lipponen A *et al.* The Arctic has warmed nearly four times faster than the globe since 1979. *Commun Earth Environ* 2022;**3**:168. https://doi.org/10.1038/s43247-022-00498-3.
2. Zhang L, Risser MD, Molter EM *et al.* Accounting for the spatial structure of weather systems in detected changes in precipitation extremes. *Weather Climate Extremes* 2022;**38**:100499. https://doi.org/10.1016/j.wace.2022.100499.
3. Perovich D, Smith M, Light B *et al.* Meltwater sources and sinks for multiyear Arctic Sea ice in summer. *Cryosphere* 2021;**15**:4517–25. https://doi.org/10.5194/tc-15-4517-2021.
4. Richter-Menge J, Jeffries M, Osborne E. 5. The arctic. *Bull Am Meteorol Soc* 2018;**99**:S143–3.
5. Kwok R. Arctic Sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). *Environ Res Lett* 2018;**13**:105005. https://doi.org/10.1088/1748-9326/aae3ec.
6. Asbjørnsen H, Årthun M, Skagseth Ø *et al.* Mechanisms underlying recent Arctic atlantification. *Geophys Res Lett* 2020;**47**:e2020GL088036. https://doi.org/10.1029/2020GL088036.

7. Wassmann P, Ratkova T, Andreassen I *et al.* Spring bloom development in the marginal ice zone and the central Barents Sea. *Mar Ecol* 1999;**20**:321–46. https://doi.org/10.1046/j.1439-0485.1999.2034081.x.
8. Reigstad M, Gabrielsen T, Amargant M *et al.* Seasonal cruise Q3: cruise report. *The Nansen Legacy Report Series* 2022;**27**:pp.9,26,45ff. https://doi.org/10.7557/nlrs.6407.
9. Kortsch S, Primicerio R, Fossheim M *et al.* Climate change alters the structure of arctic marine food webs due to poleward shifts of boreal generalists. *Proc R Soc B Biol Sci* 1814;**282**:20151546. https://doi.org/10.1098/rspb.2015.1546.
10. Polyakov IV, Pnyushkov AV, Alkire MB *et al.* Greater role for Atlantic inflows on sea-ice loss in the Eurasian Basin of the Arctic Ocean. *Science* 2017;**356**:285–91. https://doi.org/10.1126/science.aai8204.
11. Vihtakari M, Welcker J, Moe B *et al.* Black-legged kittiwakes as messengers of Atlantification in the Arctic. *Sci Rep* 2018;**8**:1–11. https://doi.org/10.1038/s41598-017-19118-8.
12. Polyakov IV, Alkire MB, Bluhm BA *et al.* Borealization of the Arctic Ocean in response to anomalous advection from sub-Arctic seas. *Front Mar Sci* 2020;**7**:491. https://doi.org/10.3389/fmars.2020.00491.
13. Collins M *et al. Long-Term Climate Change: Projections, Commitments and Irreversibility*, Cambridge University Press 2013.
14. Flocco D, Feltham DL, Turner AK. Incorporation of a physically based melt pond scheme into the sea ice component of a climate model. *J Geophys Res* 2010;**115**:115. https://doi.org/10.1029/2009JC005568.
15. Horvat C, Tziperman E. A prognostic model of the sea-ice floe size and thickness distribution. *Cryosphere* 2015;**9**:2119–34. https://doi.org/10.5194/tc-9-2119-2015.
16. Tsamados M, Feltham D, Petty A *et al.* Processes controlling surface, bottom and lateral melt of Arctic Sea ice in a state of the art sea ice model. *Philos Trans R Soc A Math Phys Eng Sci* 2015;**373**:20140167. https://doi.org/10.1098/rsta.2014.0167.
17. Lee CM *et al.* An autonomous approach to observing the seasonal ice zone in the western Arctic. *Oceanography* 2017;**30**:56–68. https://doi.org/10.5670/oceanog.2017.222.
18. Strong C, Foster D, Cherkaev E *et al.* On the definition of marginal ice zone width. *J Atmos Ocean Technol* 2017;**34**:1565–84. https://doi.org/10.1175/JTECH-D-16-0171.1.
19. Boutin G, Lique C, Ardhuin F *et al.* Towards a coupled model to investigate wave–sea ice interactions in the Arctic marginal ice zone. *Cryosphere* 2020;**14**:709–35. https://doi.org/10.5194/tc-14-709-2020.
20. Lester CW, Wagner TJW, McNamara DE *et al.* The influence of meltwater on phytoplankton blooms near the sea-ice edge. *Geophys Res Lett* 2021;**48**:e2020GL091758. https://doi.org/10.1029/2020GL091758.
21. Pachauri RK *et al.* Climate change 2014: Synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change. In: Cambridge: Ipcc, 2014.
22. Wekerle C, Wang Q, von Appen WJ *et al.* Eddy-resolving simulation of the Atlantic water circulation in the Fram Strait with focus on the seasonal cycle. *J Geophys Res* 2017;**122**:8385–405. https://doi.org/10.1002/2017JC012974..
23. von Appen WJ, Wekerle C, Hehemann L *et al.* Observations of a submesoscale cyclonic filament in the marginal ice zone. *Geophys Res Lett* 2018;**45**:6141–9. https://doi.org/10.1029/2018GL077897.
24. Hofmann Z, von Appen WJ, Wekerle C. Seasonal and mesoscale variability of the two Atlantic water recirculation pathways in Fram Strait. *J Geophys Res* 2021;**126**:e2020JC017057.
25. Timmermans M-L, Labe Z. *Sea Surface Temperature*, United States, National Oceanic and Atmospheric Administration. Series: NOAA technical report OAR ARC. 2022. https://doi.org/10.25923/p493-2548.
26. Lind S, Ingvaldsen RB, Furevik T. Arctic warming hotspot in the northern Barents Sea linked to declining sea-ice import. *Nat Clim Chang* 2018;**8**:634–9. https://doi.org/10.1038/s41558-018-0205-y.
27. Lafond A *et al.* Late spring bloom development of pelagic diatoms in Baffin Bay. In: Jody W. Deming, Michel C. (ed), Elementa: Science of the Anthropocene, 2019, 7. https://doi.org/10.1525/elementa.382.
28. Søreide JE *et al. Environmental Status of Svalbard Coastal Waters: Coastscapes and Focal Ecosystem Components*, In SESSreport 2020 - The State of Environmental Science in Svalbard - anannual report (S. 142–175). Svalbard Integrated Arctic Earth Observing System. 2020. https://doi.org/10.5281/zenodo.4293849.
29. von Eye, M., A. von Eye, and J. Rodrigues, *Global Warming and Changes in Sea Ice in the Greenland Sea: 1979–2007*. Cambridge, InterStat: Available online at http://interstat.statjournals.net/YEAR/2009/abstracts/0905003.php, 2009.
30. Hansen J, Sato M, Russell G *et al.* Climate sensitivity, sea level and atmospheric carbon dioxide. *Philos Trans R Soc A Math Phys Eng Sci* 2001;**371**:20120294. https://doi.org/10.1098/rsta.2012.0294.
31. Sumata H, de Steur L, Divine DV *et al.* Regime shift in Arctic Ocean sea ice thickness. *Nature* 2023;**615**:443–9. https://doi.org/10.1038/s41586-022-05686-x.
32. von Appen W-J, Waite AM, Bergmann M *et al.* Sea-ice derived meltwater stratification slows the biological carbon pump: results from continuous observations. *Nat Commun* 2021;**12**:7309. https://doi.org/10.1038/s41467-021-26943-z.
33. Lampe V, Nöthig E-M, Schartau M. Spatio-temporal variations in community size structure of Arctic protist plankton in the Fram Strait. *Front Mar Sci* 2021;**7**:579880. https://doi.org/10.3389/fmars.2020.579880.
34. Wingfield JC. Ecological processes and the ecology of stress: the impacts of abiotic environmental factors. *Funct Ecol* 2013;**27**:37–44. https://doi.org/10.1111/1365-2435.12039.
35. Kilias E *et al. The Successful Spreading Over the Arctic Ocean*. Picoplankton, 2012.
36. Metfies K, von Appen WJ, Kilias E *et al.* Biogeography and photosynthetic biomass of arctic marine pico-eukaroytes during summer of the record sea ice minimum 2012. *PLoS One* 2016;**11**:e0148512. https://doi.org/10.1371/journal.pone.0148512.
37. Fadeev E *et al. Sea-Ice Retreat May Decrease Carbon Export and Vertical Microbial Connectivity in the Eurasian Arctic Basins*, 2020. https://doi.org/10.21203/rs.3.rs-101878/v1.
38. Soltwedel T, Bauerfeind E, Bergmann M *et al.* HAUSGARTEN: multidisciplinary investigations at a deep-sea, long-term observatory in the Arctic Ocean. *Oceanography* 2005;**18**:46–61. https://doi.org/10.5670/oceanog.2005.24.
39. Callahan BJ, McMurdie PJ, Rosen MJ *et al.* DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016;**13**:581–3. https://doi.org/10.1038/nmeth.3869.
40. Spreen G, Kaleschke L, Heygster G. Sea ice remote sensing using AMSR-E 89-GHz channels. *J Geophys Res: Oceans* 2008;**113**:113. https://doi.org/10.1029/2005JC003384..
41. von Appen W-J. *Report on Mooring Processing of PS99. 2/PS100/PS101 Recoveries*, 2017.
42. Metfies K, Hessel J, Klenk R *et al.* Uncovering the intricacies of microbial community dynamics at Helgoland roads at the end of a spring bloom using automated sampling and 18S metabarcoding. *PLoS One* 2020;**15**:e0233921. https://doi.org/10.1371/journal.pone.0233921.

43. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal* 2011;**17**:10–2. https://doi.org/10.14806/ej.17.1.200.
44. Wietz M, Bienhold C, Metfies K *et al.* The polar night shift: seasonal dynamics and drivers of Arctic Ocean microbiomes revealed by autonomous sampling. *ISME Communications* 2021;**1**: 1–12. https://doi.org/10.1038/s43705-021-00074-4.
45. Guillou L, Bachar D, Audic S *et al.* The Protist ribosomal reference database (PR2): a catalog of unicellular eukaryote small subunit rRNA sequences with curated taxonomy. *Nucleic Acids Res* 2012;**41**:D597–604. https://doi.org/10.1093/nar/gks1160.
46. Machné R, Murray DB, Stadler PF. Similarity-based segmentation of multi-dimensional signals. *Sci Rep* 2017;**7**:12355. https://doi.org/10.1038/s41598-017-12401-8.
47. Hubbell, S.P., *The unified neutral theory of biodiversity and biogeography (MPB-32)*, in *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. 2011, Princeton: Princeton University Press, https://doi.org/10.1515/9781400837526.
48. Priest T, von Appen WJ, Oldenburg E *et al.* Atlantic water influx and sea-ice cover drive taxonomic and functional shifts in Arctic marine bacterial communities. *ISME J* 2023;**17**:1612–25. https://doi.org/10.1038/s41396-023-01461-6.
49. Machné R, Murray DB. The yin and yang of yeast transcription: elements of a global feedback system between metabolism and chromatin. *PLoS One* 2012;**7**:e37906. https://doi.org/10.1371/journal.pone.0037906.
50. Cooley JW, Tukey JW. An algorithm for the machine calculation of complex Fourier series. *Math Comput* 1965;**19**:297–301. https://doi.org/10.1090/S0025-5718-1965-0178586-1.
51. Lipovetsky S. Numerical recipes: the art of scientific computing. *Technometrics* 2009;**51**:481.
52. Ma B, Wang H, Dsouza M *et al.* Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. *ISME J* 2016;**10**:1891–901. https://doi.org/10.1038/ismej.2015.261.
53. Lima-Mendez G, Faust K, Henry N *et al.* Determinants of community structure in the global plankton interactome. *Science* 2015;**348**:1262073. https://doi.org/10.1126/science.1262073.
54. Ma B, Wang Y, Ye S *et al.* Earth microbial co-occurrence network reveals interconnection pattern across microbiomes. *Microbiome* 2020;**8**:1–12. https://doi.org/10.1186/s40168-020-00857-2.
55. Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. *PLoS Comput Biol* 2012;**8**:e1002687. https://doi.org/10.1371/journal.pcbi.1002687.
56. Burki F, Sandin MM, Jamy M. Diversity and ecology of protists revealed by metabarcoding. *Curr Biol* 2021;**31**:R1267–80. https://doi.org/10.1016/j.cub.2021.07.066.
57. Mock T, Hoch N. Long-term temperature acclimation of photosynthesis in steady-state cultures of the polar diatom *Fragilariopsis cylindrus*. *Photosynth Res* 2005;**85**:307–17. https://doi.org/10.1007/s11120-005-5668-9.
58. Bayer-Giraldi M, Weikusat I, Besir H *et al.* Characterization of an antifreeze protein from the polar diatom *Fragilariopsis cylindrus* and its relevance in sea ice. *Cryobiology* 2011;**63**:210–9. https://doi.org/10.1016/j.cryobiol.2011.08.006.
59. Raghu Prasad R, Nair P. Observations on the distribution and occurrence of diatoms in the inshore waters of the Gulf of Mannar and Palk Bay. *Indian Journal of Fisheries* 1960;**7**:49–68.
60. Comeau AM, Philippe B, Thaler M *et al.* Protists in Arctic drift and land-fast sea ice. *J Phycol* 2013;**49**:229–40. https://doi.org/10.1111/jpy.12026.
61. Tisserand L, Dadaglio L, Intertaglia L *et al.* Use of organic exudates from two polar diatoms by bacterial isolates from the Arctic Ocean. *Phil Trans R Soc A* 2020;**378**:20190356. https://doi.org/10.1098/rsta.2019.0356.
62. Lacour T, Larivière J, Ferland J *et al.* Photoacclimation of the polar diatom *Chaetoceros neogracilis* at low temperature. *PLoS One* 2022;**17**:e0272822. https://doi.org/10.1371/journal.pone.0272822.
63. Barbieri ES, Villafañ VE, Helbling EW. Experimental assessment of UV effects on temperate marine phytoplankton when exposed to variable radiation regimes. *Limnol Oceanogr* 2002;**47**: 1648–55. https://doi.org/10.4319/lo.2002.47.6.1648.
64. Lomas MW, Baer SE, Acton S *et al.* Pumped up by the cold: elemental quotas and stoichiometry of cold-water diatoms. *Front Mar Sci* 2019;**6**:286. https://doi.org/10.3389/fmars.2019.00286.
65. Kraft A, Berge J, Varpe Ø *et al.* Feeding in Arctic darkness: mid-winter diet of the pelagic amphipods *Themisto abyssorum* and *T. libellula*. *Mar Biol* 2013;**160**:241–8. https://doi.org/10.1007/s00227-012-2065-8.
66. Egge E, Elferink S, Vaulot D *et al.* An 18S V4 rRNA metabarcoding dataset of protist diversity in the Atlantic inflow to the Arctic Ocean, through the year and down to 1000 m depth. *Earth System Science Data* 2021;**13**:4913–28. https://doi.org/10.5194/essd-13-4913-2021.
67. Sherr EB, Sherr BF, Wheeler PA *et al.* Temporal and spatial variation in stocks of autotrophic and heterotrophic microbes in the upper water column of the Central Arctic Ocean. *Deep-Sea Res I Oceanogr Res Pap* 2003;**50**:557–71. https://doi.org/10.1016/S0967-0637(03)00031-1.
68. Marquardt M, Vader A, Stübner EI *et al.* Strong seasonality of marine microbial eukaryotes in a high-Arctic fjord (Isfjorden, in West Spitsbergen, Norway). *Appl Environ Microbiol* 2016;**82**: 1868–80. https://doi.org/10.1128/AEM.03208-15.
69. Cota GF. Photoadaptation of high Arctic ice algae. *Nature* 1985;**315**:219–22. https://doi.org/10.1038/315219a0.
70. von Quillfeldt CH. Distribution of diatoms in the northeast water polynya. *Greenland J Marine Syst* 1997;**10**:211–40. https://doi.org/10.1016/S0924-7963(96)00056-5.
71. Smetacek V. Role of sinking in diatom life-history cycles: ecological, evolutionary and geological significance. *Mar Biol* 1985;**84**: 239–51. https://doi.org/10.1007/BF00392493.
72. van de Poll WH, Abdullah E, Visser RJW *et al.* Taxon-specific dark survival of diatoms and flagellates affects Arctic phytoplankton composition during the polar night and early spring. *Limnol Oceanogr* 2020;**65**:903–14. https://doi.org/10.1002/lno.11355.
73. Von Quillfeldt C. Common diatom species in Arctic spring blooms: Their distribution and abundance. 2020;**43**(6):499–516. https://doi.org/10.1515/BOT.2000.050.
74. Stachura-Suchoples K, Enke N, Schlie C *et al.* Contribution towards a morphological and molecular taxonomic reference library of benthic marine diatoms from two Arctic fjords on Svalbard (Norway). *Polar Biol* 2016;**39**:1933–56. https://doi.org/10.1007/s00300-015-1683-2.
75. Balzano S, Percopo I, Siano R *et al.* Morphological and genetic diversity of Beaufort Sea diatoms with high contributions from the *Chaetoceros neogracilis* species complex. *J Phycol* 2017;**53**: 161–87. https://doi.org/10.1111/jpy.12489.
76. Lalande C, Grebmeier JM, McDonnell AMP *et al.* Impact of a warm anomaly in the Pacific Arctic region derived from time-series export fluxes. *PLoS One* 2021;**16**:e0255837. https://doi.org/10.1371/journal.pone.0255837.
77. Hartley B, Ross R, Williams DM. A check-list of the freshwater, brackish and marine diatoms of the British Isles and adjoining coastal waters. *J Mar Biol Assoc U K* 1986;**66**:531–610. https://doi.org/10.1017/S0025315400042235.

78. Schlüter MH, Kraberg A, Wiltshire KH. Long-term changes in the seasonality of selected diatoms related to grazers and environmental conditions. *J Sea Res* 2012;**67**:91–7. https://doi.org/10.1016/j.seares.2011.11.001.
79. Sukhanova I, Flint MV, Whitledge TE *et al.* Mass development of the planktonic diatom *Proboscia alata* over the Bering Sea shelf in the summer season. *Oceanology* 2006;**46**:200–16. https://doi.org/10.1134/S000143700602007X.
80. Crawford RM, Hinz F, Honeywill C. Three species of the diatom genus *Corethron castracane*: structure, distribution and taxonomy. *Diatom Res* 1998;**13**:1–28. https://doi.org/10.1080/0269249X.1998.9705432.
81. Bates SS, Garrison DL, Horner RA. Bloom dynamics and physiology of domoic-acid-producing pseudo-nitzschia species. *NATO ASI series G ecological sciences* 1998;**41**:267–92.
82. Nöthig E-M, Bracher A, Engel A *et al.* Summertime plankton ecology in Fram Strait—a compilation of long-and short-term observations. *Polar Res* 2015;**34**:23349. https://doi.org/10.3402/polar.v34.23349.
83. Kauko HM, Pavlov AK, Johnsen G *et al.* Photoacclimation state of an Arctic underice phytoplankton bloom. *J Geophys Res Oceans* 2019;**124**:1750–62. https://doi.org/10.1029/2018JC014777.
84. Biggs TE, Alvarez-Fernandez S, Evans C *et al.* Antarctic phytoplankton community composition and size structure: importance of ice type and temperature as regulatory factors. *Polar Biol* 2019;**42**:1997–2015. https://doi.org/10.1007/s00300-019-02576-3.
85. Dalpadado P, Arrigo KR, van Dijken GL *et al.* Climate effects on temporal and spatial dynamics of phytoplankton and zooplankton in the Barents Sea. *Prog Oceanogr* 2020;**185**:102320. https://doi.org/10.1016/j.pocean.2020.102320.
86. Nardelli SC, Gray PC, Stammerjohn SE *et al.* Characterizing coastal phytoplankton seasonal succession patterns on the West Antarctic peninsula. *Limnol Oceanogr* 2023;**68**:845–61. https://doi.org/10.1002/lno.12314.

Chapter 3

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

In this chapter, we present our use case concerning analysis of atlantic water influx to bacteria communities. Here we investigates how Atlantification affects Arctic microbial communities by analyzing changes in composition, structure, and function in the Fram Strait mixing zone (Figure 3.1).

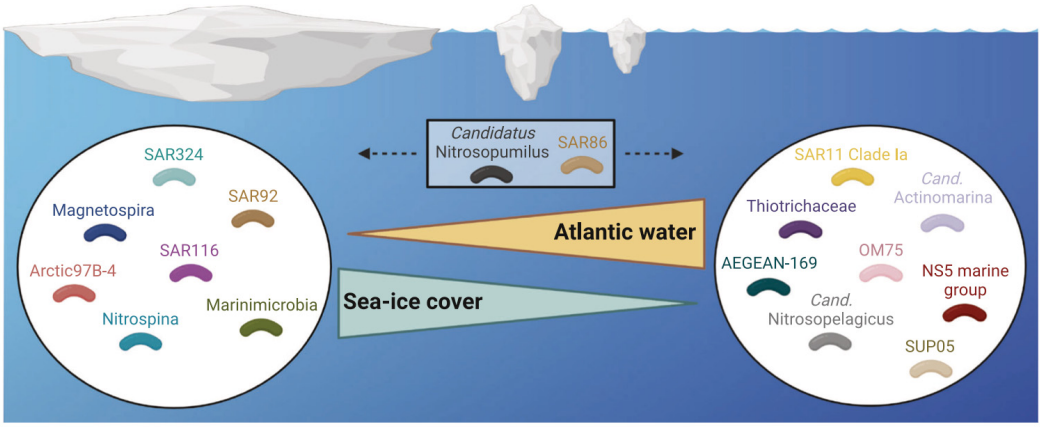

Figure 3.1: **Bacterial communities under contrasting AW influx and ice cover conditions.** Illustration showing the ten taxonomic groups with highest average relative abundances under Atlantic vs. Arctic conditions, derived from the relative abundances of Int-ASVs (sPLS cluster C1) and Res-ASVs (sPLS cluster C8), respectively. Adapted from Priest et al., 2022.

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

In this section, we provide an overview of the contributions and impact of our paper (Priest et al., 2022):

Taylor Priest , Wilken-Jon von Appen , Ellen Oldenburg , Ovidiu Popa , Sinhué Torres-Valdés ,Christina Bienhold, Katja Metfies , Bernhard M. Fuchs , Rudolf Amann , Antje Boetius ,Matthias Wietz

"Variations in Atlantic water influx and sea-ice cover drive taxonomic and functional shifts in Arctic marine bacterial communities"
In: *The ISME Journal, 2023, volume 17*

Main Results in Simple Terms

This study examines the impact of the increasing influx of Atlantic water into the Arctic Ocean, known as Atlantification, on microbial communities. The study focused on areas where Pacific water meets Atlantic water in Fram Strait, with the aim of understanding how changes in sea-ice cover and Atlantic water influx impact these communities. The detailed analysis revealed significant changes in the types and functions of bacteria in response to these environmental changes. Areas with thick sea ice exhibited stable microbial communities, while areas with less ice and more Atlantic water exhibited fluctuating populations linked to phytoplankton. Additionally, specific groups of bacteria were identified that were associated with different environmental conditions, indicating how they adapt to their surroundings. Our findings suggest that as Atlantification continues, bacterial populations in the Arctic will shift, with consequences for the entire ecosystem.

Summary/Abstract

The Arctic Ocean is undergoing unprecedented changes due to climate warming, necessitating in-depth analysis of its biological communities to comprehend present and future ecosystem transformations. In this study, we compiled a comprehensive dataset spanning four years, including high-resolution amplicon data and PacBio HiFi read metagenomes from the East Greenland Current (EGC). These datasets were integrated with datasets from various spatiotemporal scales (Tara Arctic and Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC)) in order to investigate the impact of the influx of Atlantic water and changes in sea-ice cover on bacterial communities in the Arctic Ocean. Our findings revealed that densely ice-covered polar waters host a stable, resident microbiome, whereas increased Atlantic water influx and reduced sea-ice cover lead to the dominance of fluctuating populations, indicating a process akin to "replacement" driven by advection, mixing, and environmental sorting. We identified bacterial signature populations associated with distinct environmental conditions, including polar night and high-ice cover, and evaluated their

ecological roles. Furthermore, our analyses demonstrated consistent dynamics of signature populations across the broader Arctic region. For instance, populations linked with dense ice cover and winter in the EGC were prevalent in the central Arctic Ocean during winter. Population- and community-level assessments unveiled metabolic disparities between bacteria adapted to Arctic and Atlantic conditions. Arctic-affiliated bacteria showed enhanced capabilities to utilise various substrates. The present study offers novel insights into Arctic ecology by elucidating the spatiotemporal dynamics of bacterial communities. It highlights the ongoing biological Atlantification of the warming Arctic Ocean and its implications for food webs and biogeochemical cycles.

Personal Contribution

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.

Importance of the Research and Contribution to this Thesis

Therefore, it answers our second research question: This study addresses our research question by demonstrating the significant impact of Atlantic water influx and changes in sea-ice cover on bacterial communities in the Arctic Ocean. Through comprehensive analysis spanning four years and integrating datasets from different spatiotemporal scales, we reveal how these environmental factors influence the composition, structure, and functionality of microbial populations, providing insights into the progressing Biological Atlantification of the warming Arctic ecosystem.

www.nature.com/ismej

ISME

ARTICLE OPEN

Check for updates

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

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

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 th,0e 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

[1]Max Planck Institute for Marine Microbiology, Bremen 28359, Germany. [2]Physical Oceanography of the Polar Seas, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven 27570, Germany. [3]Institute for Quantitative and Theoretical Biology, Heinrich Heine University Düsseldorf, Düsseldorf 40225, Germany. [4]Deep-Sea Ecology and Technology, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven 27570, Germany. [5]Polar Biological Oceanography, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven 27570, Germany. [6]School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom. [7]School of Computing Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom. [8]MARUM – Center for Marine Environmental Sciences, University of Bremen, Bremen 28359, Germany. ✉email: tpriest@mpi-bremen.de; 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 segmenTier [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 FEMTOpulse. 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 $<1^{-10}$. 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-amsr-eamsr2, 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

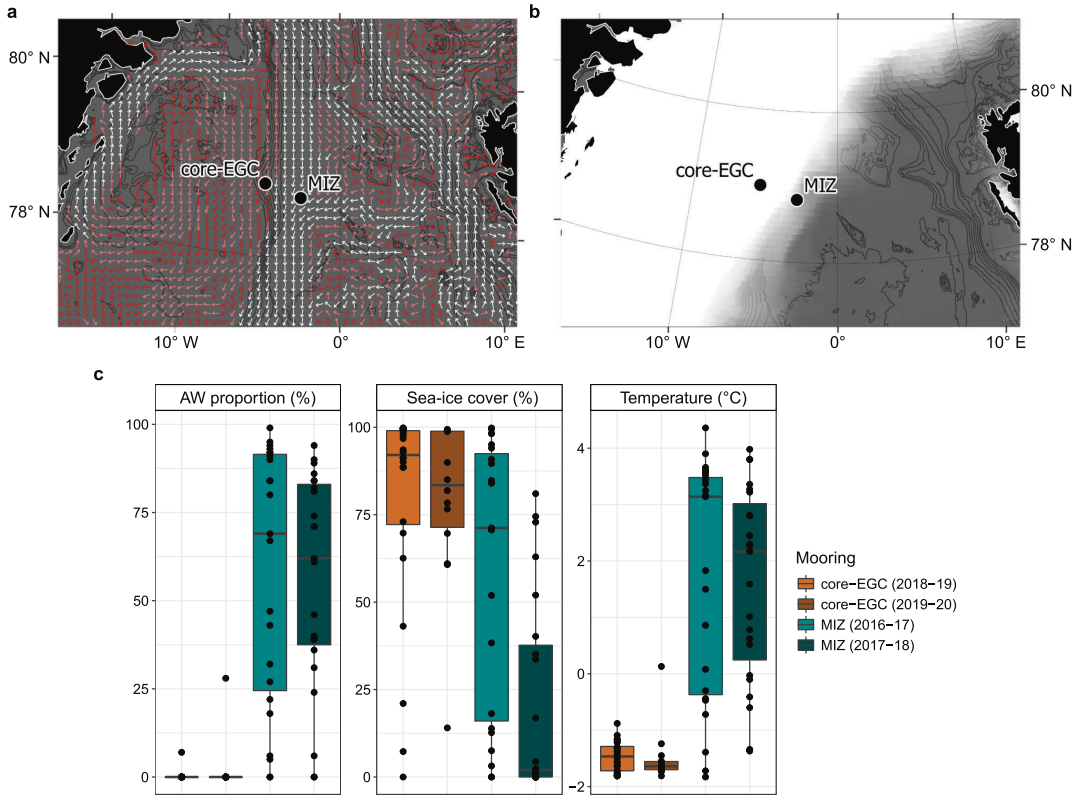

Fig. 1 Location of seafloor moorings and environmental conditions in the MIZ (2016–2018) and core-EGC (2018–2020). a Example representation of monthly average (January 2020) current velocities at the approximate depth of sampling (78 m). White and dark red arrows indicate strongest and weakest velocities, 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.

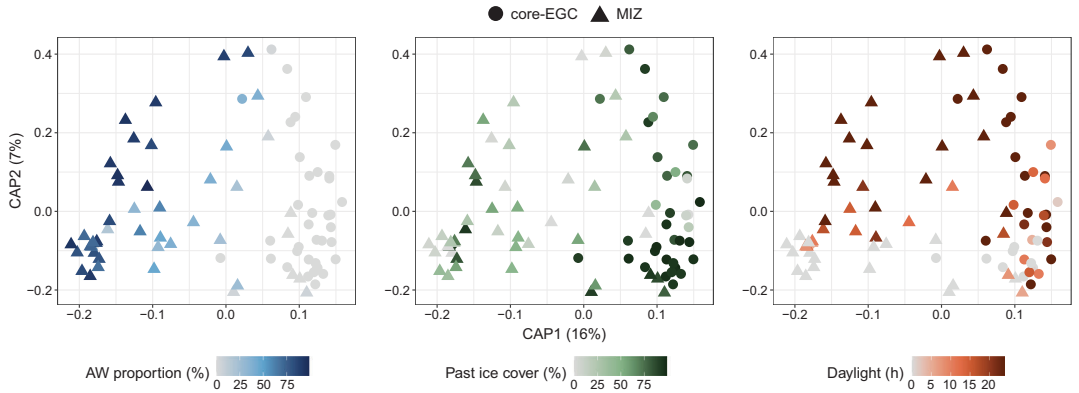

Fig. 2 Community structure across water mass, sea-ice and daylight conditions. Distance-based redundancy analysis based on Bray-Curtis dissimilarities of community composition along with AW proportion (blue), past ice cover (green) and daylight (orange) as constraining factors. The factors were selected 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.

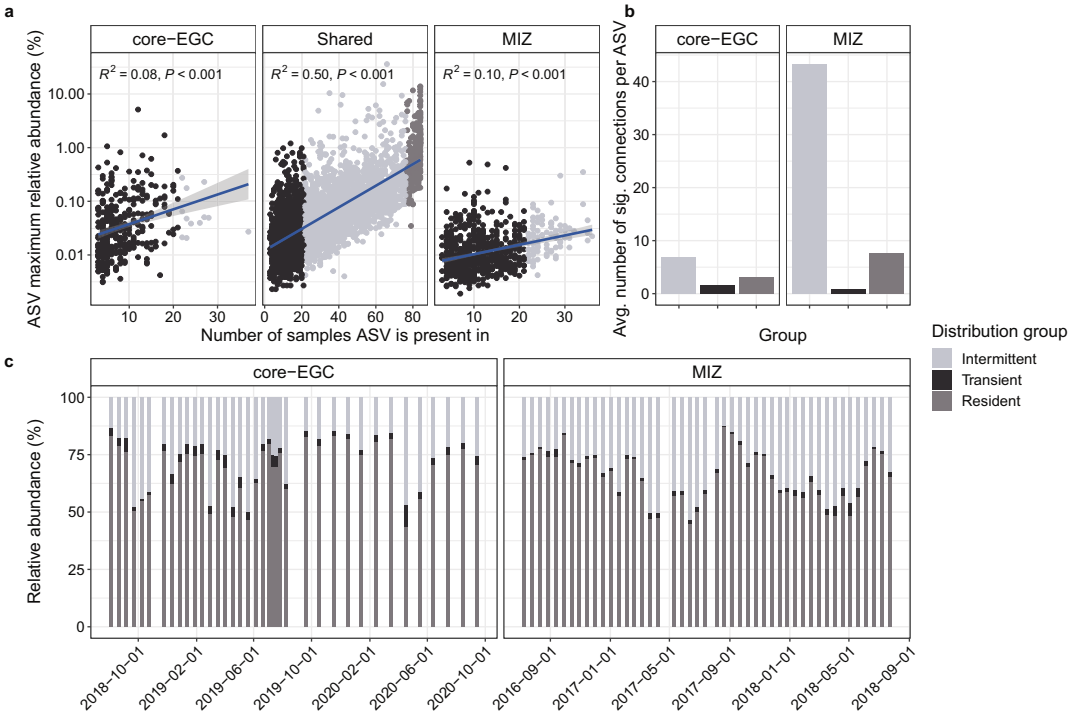

Fig. 3 Distribution dynamics and co-occurrence of ASVs. a Occurrence of ASVs across samples in relation to their maximum relative abundances, along with categorisation into resident, intermittent and transient. **b** Average number of connections within the co-occurrence networks for resident, intermittent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

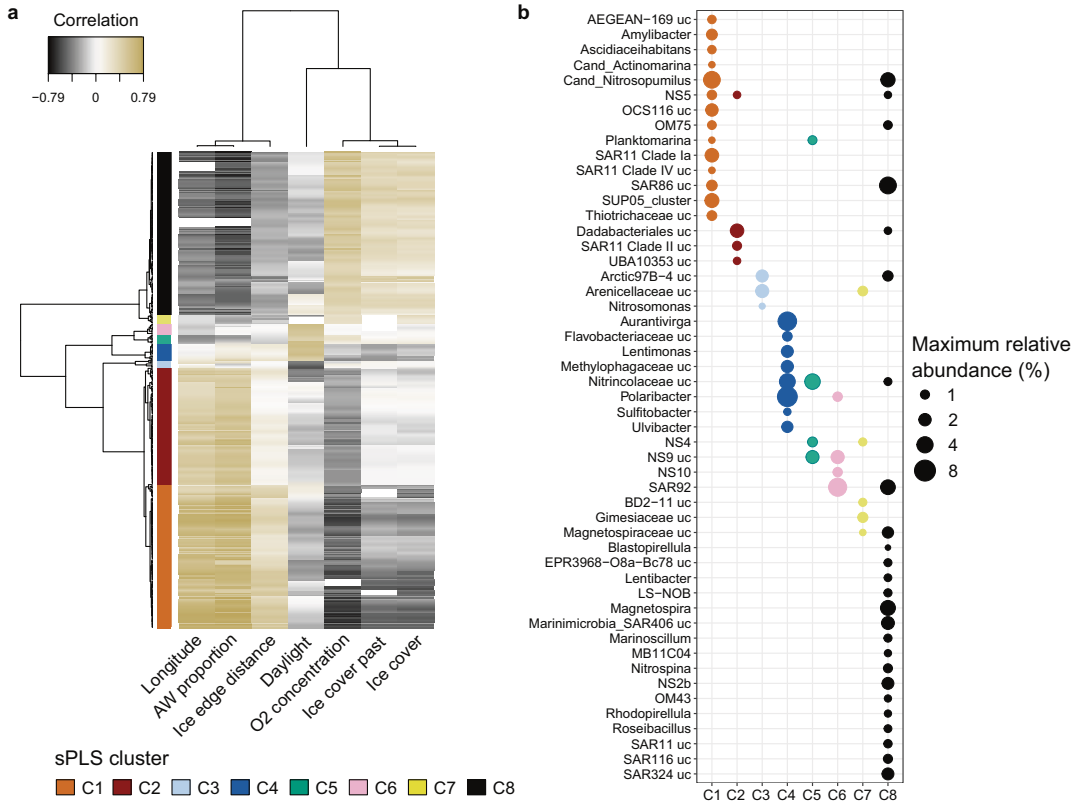

Fig. 4 Sparse partial least square regression (sPLS) linking community structure and environmental parameters. a Heatmap showing eight major sPLS clusters, encompassing 430 ASVs with significant correlations to environmental conditions. **b** Representation of the most prominent genera per cluster. ASVs 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.

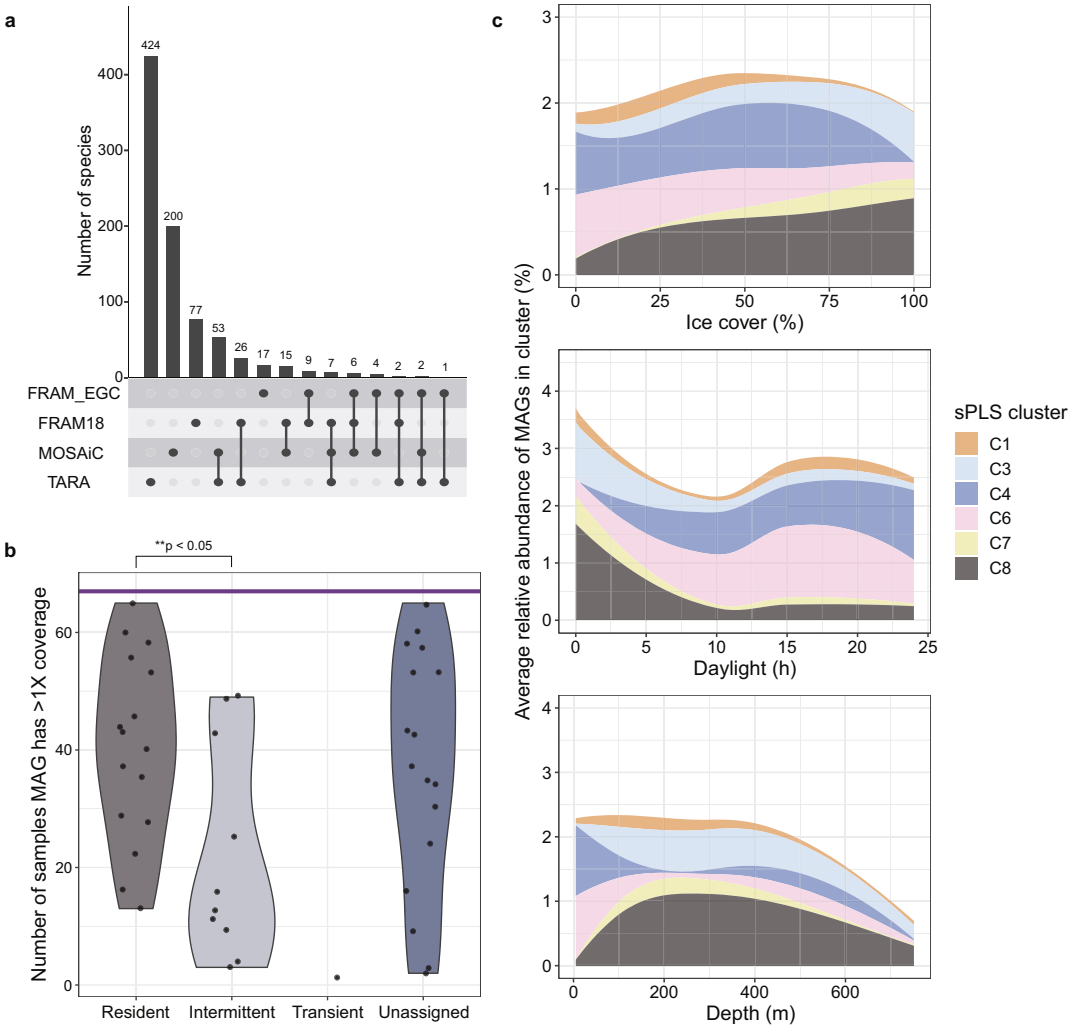

Fig. 5 Comparison and dynamics of MAGs across the Fram Strait and Arctic Ocean. We compared metagenome-assembled genomes (MAGs) generated in this study (FRAM_EGC), and from samples previously collected in the Fram Strait (FRAM18) [23], in the Arctic Ocean during summer (TARA) [34], and in the Arctic 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

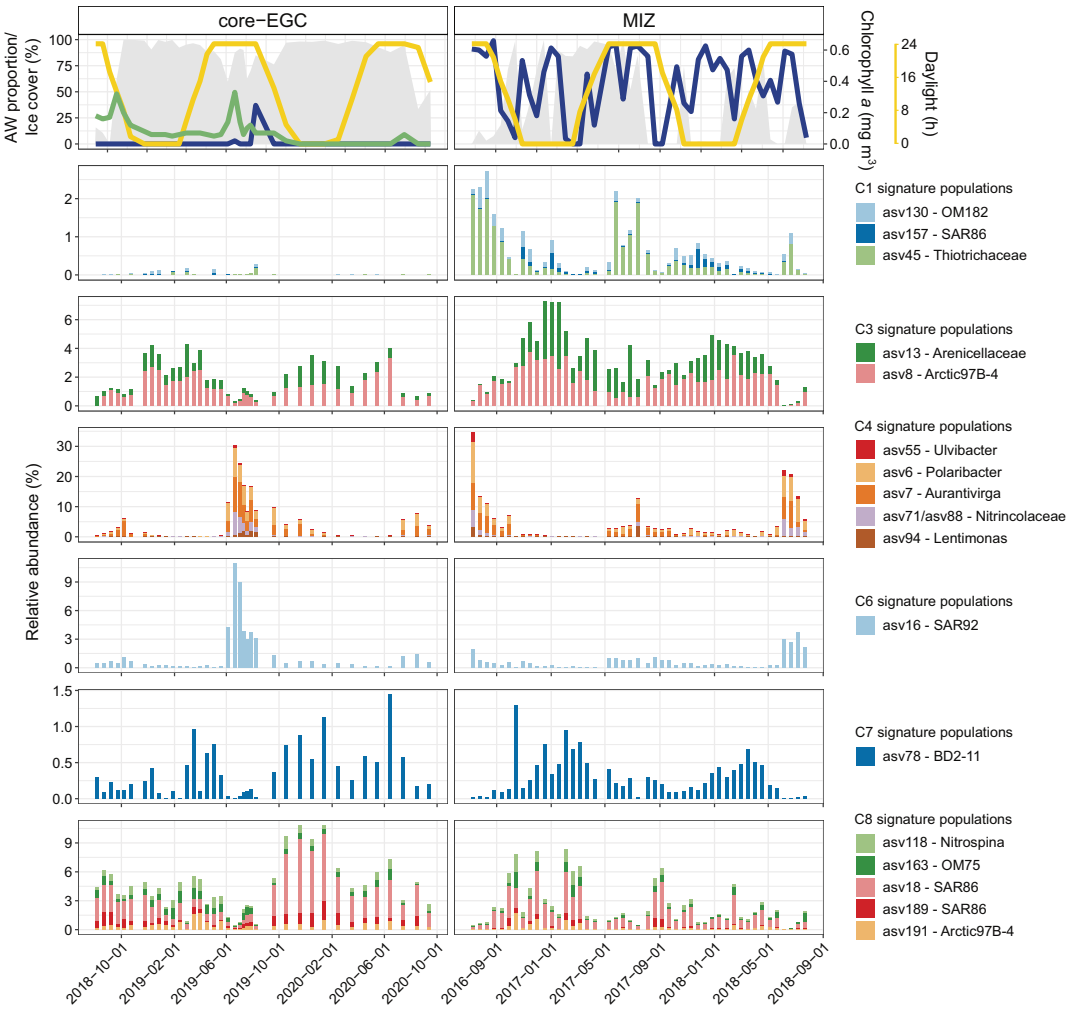

Fig. 6 Temporal dynamics of signature populations. Signature populations were identified as ASV representatives from sPLS clusters that a corresponding MAG was recovered for (based on 100% identity threshold competitive read recruitment). The temporal dynamics visualised are derived from ASV data. The 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, α-galactose- and β-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 physiochemical 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

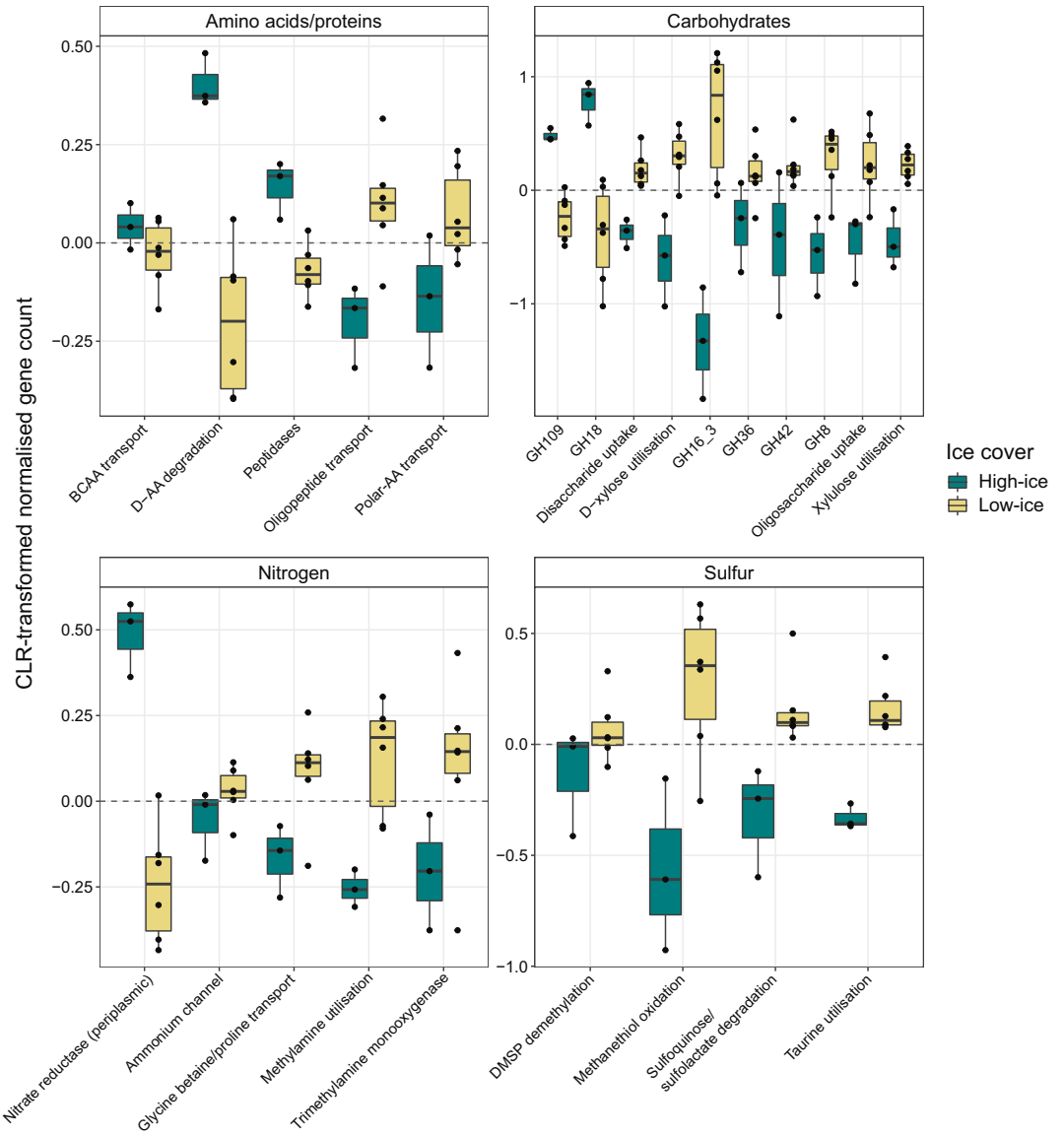

Fig. 7 Selected genes involved in the metabolism of organic and inorganic compounds enriched under high- and low-ice conditions. Enrichment is displayed as centred-log ratio transformed normalised gene counts. Where several genes of a single pathway or mechanism were identified as enriched, they were 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].

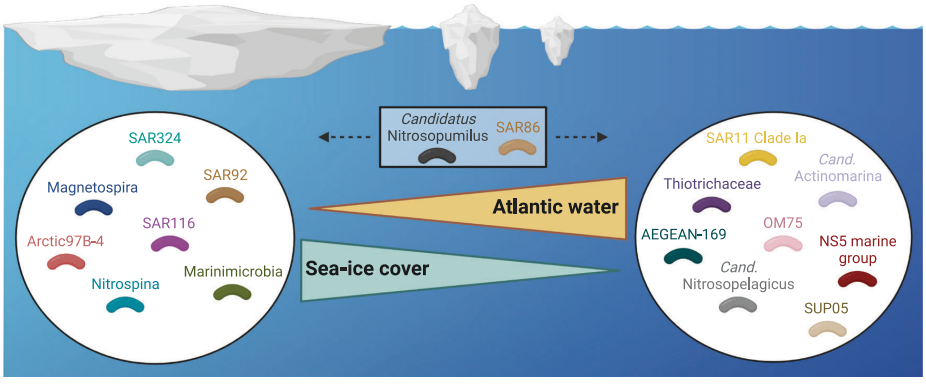

Fig. 8 Bacterial communities under contrasting AW influx and ice cover conditions. Illustration showing the ten taxonomic groups with highest average relative abundances under Atlantic vs. Arctic conditions, derived from the relative abundances of Int-ASVs (sPLS cluster C1) and Res-ASVs (sPLS cluster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

REFERENCES

1. 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.
2. Kwok R. Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). Environ Res Lett. 2018;13:105005.
3. 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.
4. Notz D, SIMIP Community. Arctic sea ice in CMIP6. Geophys Res Lett. 2020;47:e2019GL086749.
5. 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.
6. 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.
7. 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.
8. Lewis KM, Dijken GL, van, Arrigo KR. Changes in phytoplankton concentration now drive increased Arctic Ocean primary production. Science. 2020;369:198–202.
9. 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.
10. Ardyna M, Arrigo KR. Phytoplankton dynamics in a changing Arctic Ocean. Nat Clim Chang. 2020;10:892–903.
11. 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.
12. 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.
13. 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.
14. 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.
15. Yergeau E, Michel C, Tremblay J, Niemi A, King TL, Wyglinski J, et al. Metagenomic survey of the taxonomic and functional microbial communities of seawater and sea ice from the Canadian Arctic. Sci Rep. 2017;7:42242.
16. Niemi A, Meisterhans G, Michel C. Response of under-ice prokaryotes to experimental sea-ice DOM enrichment. Aquat Micro Ecol. 2014;73:17–28.
17. 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.
18. 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 HAUSGARTEN. Ecol Indic. 2016;65:89–102.
19. 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.
20. 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.
21. 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.
22. 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.
23. 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 microbiomes revealed by autonomous sampling. ISME Commun. 2021;1:1–12.
24. 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.
25. 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.
26. 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.
27. 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.
28. 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.
29. RStudio Team. RStudio: Integrated development of R. 2020. PBC, Boston, MA.
30. 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.
31. 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.
32. 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.
33. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer; 2016.
34. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32:2847–9.
35. Machné R, Murray DB, Stadler PF. Similarity-based segmentation of multi-dimensional signals. Sci Rep. 2017;7:12355.
36. 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.
37. 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.
38. 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.
39. 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.
40. 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.
41. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12:59–60.
42. Shen W, Ren H. TaxonKit: a practical and efficient NCBI taxonomy toolkit. J Genet Genomics. 2021;48:844–50.
43. Seemann T. Prokka: rapid prokaryotic genome annotation. J Bioinform. 2014;30:2068–9.
44. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinform. 2009;10:421.
45. Eddy SR. Accelerated profile HMM searches. PLoS Comput Biol. 2011;7:e1002195.
46. 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.
47. 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.
48. 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.
49. 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.
50. 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.
51. 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.
52. 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 maximuml 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, Yadhukumar, 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 RI. Verrucomicrobiota 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, Kyrpides 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. Multiomics in the central Arctic Ocean for benchmarking biodiversity change. PLoS Biol. 2022;20:e3001835.
76. Hoarfrost A, Nayfach S, Ladau J, Yooseph S, Arnosti C, Dupont CL, et al. Global ecotypes in the ubiquitous marine clade SAR86. ISME J. 2020;14:178–88.
77. Francis B, Urich 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:S343–S362.
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. Gordeev 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 Ebenhöh 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.

Chapter 4

Bacteria: Seasonal recurrence and modular assembly of an Arctic pelagic marine microbiome

In this chapter, we present our use case concerning analysis of pelagic Arctic Ocean microbiome dynamics over seasonal and interannual scales. Here we examine taxonomic and functional dynamics with high temporal resolution data collected over four years using autonomous samplers and in situ sensors.

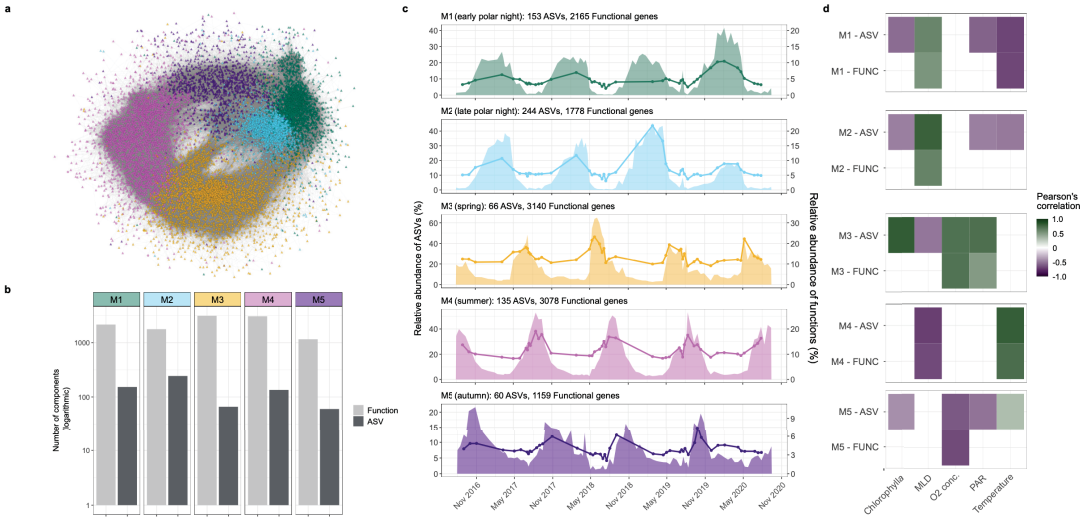

Figure 4.1: **Ecosystem modules, their temporal dynamics and association with environmental conditions a)** Co-occurrence network of prokaryotic ASVs and functional cluster oscillation signals (Pearson correlations >0.7). Circle nodes represent ASVs, triangle nodes represent functional clusters. **b)** Number of components in the modules. **c)** Temporal abundance dynamics of module ASVs (area) and functional clusters (line). **d)** Significant Pearson's correlations between combined abundance of module ASVs and functional clusters against environmental factors ($p < 0.05$ after multiple testing correction) (Priest et al., 2024).

4.1 Seasonal recurrence and modular assembly of an Arctic pelagic marine microbiome

In this section, we provide an overview of the contributions and impact of our paper (Priest et al., 2024):

Taylor Priest, Ellen Oldenburg, Ovidiu Popa, Bledina Dede, Katja Metfies, Wilken-Jon von Appen, Sinhué Torres-Valdés, Christina Bienhold, Bernhard M. Fuchs, Rudolf Amann, Antje Boetius, Matthias Wietz
"Seasonal recurrence and modular assembly of an Arctic pelagic marine microbiome"
In: *biorxiv, 2024 (Submitted to Nature Communications)*

Main Results in Simple Terms

This paper explores how different types of tiny organisms, called microbes, come together in the ocean, particularly in the Arctic. We used special tools to track these microbes over four years, focusing on how they change over time. We found that certain groups of microbes tend to come and go in a pattern each year, and this pattern is influenced by the seasons and environmental factors like sunlight and nutrients. For example, in the winter, we often see a certain type of microbe that helps break down ammonia, while in the summer, we find different microbes that help break down sulfur compounds. We also noticed that from year to year, the types of microbes and their functions can vary, showing how the environment affects them. This study gives us important insights into how these tiny organisms organize themselves in the ocean, especially in places like the Arctic where there hasn't been much research done before.

Summary/Abstract

An understanding of how microbial communities assemble in different environments is essential for comprehending the functioning of ocean ecosystems. Currently, the diversity of functions exhibited by these microbial communities and the relationship between their classification and their activities, particularly over time, remain poorly understood. In this study, we employed specialized equipment to track the alterations in both the types of microbes and their activities in the pelagic Arctic Ocean over four years, with considerable detail. The study revealed that the primary types of microscopic organisms exhibit predictable fluctuations in abundance and genetic composition on an annual basis. These organisms form five distinct patterns throughout the seasons, reflecting varying ecological conditions. For instance, during the early stages of the polar night, we frequently observe a specific microbe and its associated function in ammonia breakdown, along with other specific organisms and oceanic conditions. In contrast, late summer brings different microbes and functions, such as the breakdown of sulfur compounds and various types of algae. Furthermore, it was observed that these patterns change from year to year, indicating that the environment plays a significant role in shaping them. For instance, in spring, there are fewer types of microbes, but they perform similar functions, and there is a great deal of variation in their genetics, which suggests that the environment selects

for certain functions over time. By integrating data on the types of microbes present, their functions, and the environmental context, our research contributes to a better understanding of the organization of microbial communities in the Arctic Ocean. This is crucial because these ecosystems are undergoing rapid changes and have been relatively understudied.

Personal Contribution

TP designed and executed the ASV and metagenomic analysis pipelines and wrote the manuscript, with input from MW, **EO** and BD. **EO** and OP performed time-series and network analyses. MW processed amplicon raw data into ASVs, co-designed the sampling and mooring strategy, and coordinated data analysis. WJvA contributed quality-controlled oceanographic data and coordinated the mooring operations. KM coordinated the processing of samples and sequencing and provided 18S rRNA gene sequence data. STV provided quality-controlled chlorophyll sensor data. CB, KM and AB co-designed the sampling and mooring strategy, and contributed to the interpretation. BF and RA contributed to study design and interpretation. **All authors** contributed to the final manuscript.

Importance of the Research and Contribution to this Thesis

Therefore, it answers our third research question: This study provides fundamental insights into the seasonal and interannual structuring of prokaryotic and microeukaryotic communities in an Arctic pelagic ocean ecosystem. By demonstrating the prevalence of annually recurrent dynamics of populations and community gene content, we identify five distinct seasonal modules. These modules represent unique ecological states within the prokaryotic microbiome each year, connected to specific microeukaryotic populations and environmental conditions. Furthermore, our findings demonstrate that environmental selection varies across these states, exerting differential pressures on both organismal and functional levels. Our results contribute to the growing body of knowledge on microbial community structuring across pronounced environmental gradients in rapidly changing and understudied ocean regions.

Seasonal recurrence and modular assembly of an Arctic pelagic marine microbiome

Taylor Priest[1,2]* & Ellen Oldenburg[3], Ovidiu Popa[3], Bledina Dede[1,4], Katja Metfies[5], Wilken-Jon von Appen[5], Sinhué Torres-Valdés[5], Christina Bienhold[1,5], Bernhard M. Fuchs[1], Rudolf Amann[1], Antje Boetius[1,5,6], Matthias Wietz[1,5]*

[1] Max Planck Institute for Microbiology, Bremen, Germany
[2] Institute of Microbiology, ETH Zurich, Zurich, Switzerland
[3] Institute for Quantitative and Theoretical Biology, Heinrich Heine University Düsseldorf, Düsseldorf, 40225, Germany
[4] Ecologie Systématique Evolution, CNRS, Université Paris-Saclay, AgroParisTech, Gif-sur-Yvette, France
[5] Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, 27570, Germany
[6] MARUM – Center for Marine Environmental Sciences, University of Bremen, Bremen, 28359, Germany

***Corresponding authors**:

Taylor Priest – tpriest@ethz.ch

Matthias Wietz – matthias.wietz@awi.de

ABSTRACT

Deciphering how microbial communities are shaped by environmental variability is fundamental for understanding the structure and function of ocean ecosystems. Thus far, we know little about the structuring of community functionality and the coupling between taxonomy and function over seasonal environmental gradients. To address this, we employed autonomous sampling devices and *in situ* sensors to investigate the taxonomic and functional dynamics of a pelagic Arctic Ocean microbiome over a four-year period. We demonstrate that the dominant prokaryotic and microeukaryotic populations exhibit recurrent, unimodal fluctuations each year, with community gene content following the same trend. The recurrent dynamics within the prokaryotic microbiome are structured into five temporal modules that represent distinct ecological states, characterised by unique taxonomic and metabolic signatures and connections to specific microeukaryotic populations and oceanographic conditions. For instance, *Cand.* Nitrosopumilus and the machinery to oxidise ammonia and reduce nitrite are signatures of early polar night, along with Radiolarians.

In contrast, late summer is characterised by *Amylibacter,* sulfur compound metabolism and diverse Haptophyta lineages. Exploring the composition of modules further along with their degree of functional redundancy and the structuring of genetic diversity within functions over time revealed seasonal heterogeneity in environmental selection processes. In particular, we observe strong selection pressure on a functional level in spring while late polar night features weaker selection pressure that likely acts on an organismal level. By integrating taxonomic, functional, and environmental information, our study provides fundamental insights into how microbiomes are structured under pronounced environmental variability in understudied, yet rapidly changing polar marine ecosystems.

INTRODUCTION

Bacteria, archaea and microeukaryotes are the dominant life forms in ocean environments and comprise an immense taxonomic, functional and physiological diversity. These microbes drive and respond to changes in their surrounding environment, such as bottom-up (e.g. resource availability) and top-down (e.g. grazing and viral infection) factors and physicochemical conditions (e.g. temperature), which results in the assembly of distinct communities over spatial and temporal scales[1,2]. The assembled communities subsequently perform essential trophic roles and mediate the biogeochemical cycling of biologically important elements[3–5]. Deciphering how microbial community dynamics are shaped across environmental gradients is thus fundamental for understanding the structure and function of ecosystems and how they respond to change.

Long-term observations have uncovered recurrent and transient dynamics in microbial communities across daily, seasonal and annual timescales. In temperate ecosystems, communities are predominantly structured by seasonal variability, with broadly recurrent fluctuations of taxa on annual scales[6–11]. These patterns have led to the conclusion that microbial responses to biological and environmental shifts are predictable[12]. However, recent evidence indicates that microeukaryotes exhibit weaker temporal structuring than prokaryotes[13], suggesting different controlling mechanisms. Furthermore, high-frequency sampling has shown that population dynamics are highly ephemeral in nature, undergoing rapid, short-lived fluctuations that transpire over days[14–16]. Thus, while microbial communities are structured over time, they also undergo constant flux, reflecting the dynamic nature of ocean ecosystems.

Despite the wealth of knowledge gained from long-term observations, the focus primarily on taxonomic dynamics and on temperate and sub-tropical ecosystems has left many questions unanswered. In particular, it remains unclear how compositional shifts across environmental gradients translate to changes in the functionality of microbial communities. Since distantly related organisms can perform similar metabolic functions[17,18], taxonomic information alone does not inform about ecological landscapes or ecosystem function. Therefore, long-term observations that integrate taxonomic, functional, and environmental information are greatly needed. This is particularly important in the polar oceans, where long-term observations are rare and unprecedented changes are taking place because of climate warming.

To address this, we investigated the dynamics of prokaryotic (here used operationally to refer to bacteria and archaea) and microeukaryotic communities from a taxonomic and functional perspective over

a four-year period in an Arctic Ocean ecosystem, the West Spitsbergen Current (WSC). The WSC constitutes the primary entry route for Atlantic water into the Arctic Ocean and is characterised by pronounced seasonal variability in environmental conditions, archetypal of high-latitude ecosystems. However, unlike other Arctic Ocean ecosystems, the WSC remains ice-free year-round, due to the warmth of the North Atlantic water. Therefore, the WSC represents a model system for investigating the dynamics of microorganisms and ecosystem function in the context of pronounced seasonal variability. In addition, it can provide valuable insights into potential future shifts in Arctic Ocean ecosystems, given the progressive northward expansion of Atlantic water influence[19].

We hypothesised that the Arctic epipelagic microbiome shows a seasonally recurring assembly primarily structured by polar day/night cycles. To investigate this, we employed moorings fitted with autonomous sampling and measuring devices to continuously track taxonomic, functional and environmental dynamics. Through 16S and 18S rRNA gene amplicon and PacBio HiFi metagenome sequencing, we generated a high-quality, temporally resolved microbiome catalogue. Using a Fourier transformation-based approach, we demonstrate that prokaryotic and microeukaryotic communities exhibit annually recurrent, seasonally structured dynamics. Within the prokaryotic microbiome, these dynamics assemble into five distinct temporal modules that feature unique taxonomic and metabolic signatures, are associated with specific microeukaryotic populations and are subject to different environmental selection pressures. Our study provides the first multi-year ecosystem catalogue from the Arctic that integrates taxonomic, functional and environmental information, and provides fundamental insights into the dynamics and structuring of microbiomes across pronounced environmental gradients.

RESULTS & DISCUSSION

The West Spitsbergen Current harbours an ecosystem with pronounced temporal structuring

We first investigated how environmental conditions in the WSC are structured over intra- and inter-annual scales. For this, we combined data collected from *in situ* sensors attached to the mooring, including temperature, salinity and oxygen saturation, with chlorophyll *a* measurements and satellite-derived values of photosynthetically active radiation (PAR) (Supplementary Table S1). Our measurements were derived from the epipelagic layer but varied between 20 – 100 m in depth due to the movement of the mooring by currents. Owing to the inclusion of multiple CTD sensors at different depths, we were also able to determine the lower bound of the mixed layer depth (MLD). Each annual cycle was characterised by pronounced shifts in environmental conditions (Figure 1). As expected, these shifts followed the transition between polar night and polar day, which is a major force stimulating biological dynamics in the Arctic[20]. The end of polar night was marked by an increase in PAR in April, which continued to rise until a maximum, on average, of 38 µmol photons m^{-2} s^{-1} in June. The increasing solar radiation onset warming, with temperatures rising until a peak of ~7°C in August/September before decreasing again to <4 °C between December – May. Changes in MLD were inversely related to temperature (Pearson correlation: R = -0.47, p < 0.05), but were

characterised by abrupt events of shallowing to <5 m in June and deepening to >200 m between December-January. Chlorophyll concentrations showed a lagged association with PAR, peaking between June-August. However, the magnitude and timing of the chlorophyll peak varied across years, from 3.35 µg l^{-1} in July 2019 to 13.22 µg l^{-1} in June 2018, indicating differences in phytoplankton bloom phenologies. The WSC thus exhibits pronounced intra-annual shifts in environmental conditions, presenting an ideal ecosystem to study seasonally driven biological dynamics.

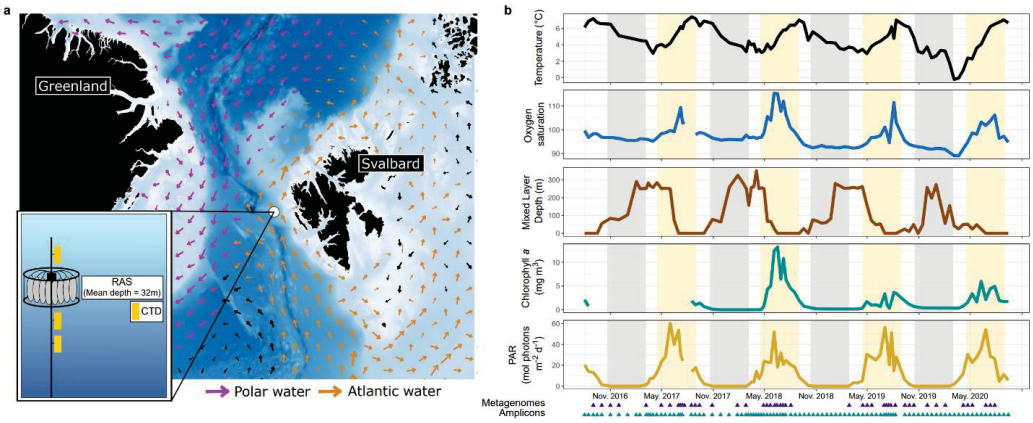

Figure 1. The WSC mooring site and environmental conditions between July 2016 and July 2020. a) Bathymetric map of the Fram Strait region with the mooring location within the West Spitsbergen Current. Arrows represent average current velocities over the four-year sampling period at the average depth of the moored autonomous sampler (32 m). **b)** Water temperature, oxygen saturation, mixed layer depth, chlorophyll concentrations measured from mooring-attached sensors and photosynthetically active radiation (PAR) derived from AQUA-MODIS satellite data. The shaded grey and yellow represent the periods of polar night and day, respectively.

The intra- and inter-annual temporal structuring of communities

We next investigated the temporal structuring of prokaryotic and microeukaryotic communities from a taxonomic perspective. Using autonomous Remote Access Samplers, we collected 97 samples at, on average, fortnightly resolution that were used for 16S and 18S rRNA gene sequencing. From these, we recovered 3629 bacterial, 119 archaeal and 3019 microeukaryotic Amplicon Sequence Variants (ASVs) (Supplementary Table S2 and S3).

The alpha diversity of prokaryotic and microeukaryotic communities exhibited distinct trends within each annual cycle (Figure 2a and 2b). For prokaryotic communities, we observed a tight coupling between Species Richness (*R*), Evenness (*E*) and Shannon Diversity (*H'*), evidenced through significant positive Pearson's correlations (Figure 2c and Supplementary Table S4 and S5). The alpha diversity metrics followed a unimodal fluctuation within each annual cycle for prokaryotic communities, reaching a peak during polar night (December - March) with a mean *R* of 1210 ± 208, *E* of 0.78 ± 0.04 and *H'* of 5.3 ± 0.36. The enriched diversity in winter observed here mirrors previous observations in temperate and polar regions[21–23] and is associated with the deepening of MLD, which drives mixing and dilution of previously

stratified communities[24]. The influx of PAR and rapid shallowing of MLD at the end of polar night coincided with a sharp drop in alpha diversity, with lowest values observed in June (R = 762 ± 156, H' = 4.50 ± 0.34, E = 0.68 ± 0.03). Hence, shifts in prokaryotic alpha diversity were tightly coupled to MLD, supported by strong positive Pearson's correlations; Richness (r = 0.71, 95% CI [0.59,0.79], p = 4 x 10^{-16}), Evenness (r = 0.43, 95% CI [0.25,0.58], p = 1.08 x 10^{-5}) and Shannon diversity (r = 0.58, 95% CI [0.43,0.70], p = 5.68 x 10^{-10}) (Supplementary Table S5). This observation is in contrast to previous reports of temperature[25–27], ocean currents[28] and day length[10,22] as key drivers of epipelagic bacterial diversity. The greater role of MLD in shaping prokaryotic diversity observed here may be a feature unique to high-latitude ocean ecosystems, where seasonal shifts in MLD are more pronounced[29] and can be influenced by sea-ice dynamics[30]. However, MLD has rarely been measured or incorporated before, and thus its considerations in future studies will help to ascertain whether its influence varies across latitude.

In contrast to prokaryotes, microeukaryotic communities exhibited a bimodal fluctuation in alpha

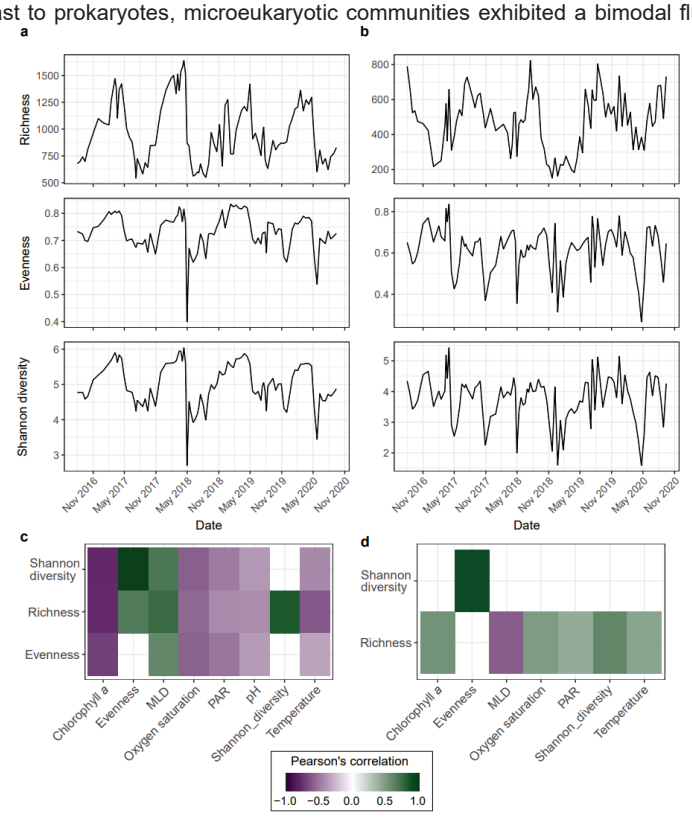

Figure 2. The diversity of prokaryotic and microeukaryotic communities is structured differently over time. All values of diversity shown represent the mean value after performing 100 iterations of rarefying and metric calculation. Richness, Evenness and Shannon diversity of **a)** prokaryotic and **b)** microeukaryotic communities. Statistically significant ($p < 0.05$) correlations between **(c)** prokaryotic and **(d)** microeukaryotic alpha-diversity measures and environmental parameters after multiple testing correction.

diversity in each annual cycle. The bimodal pattern was reflected in a peak in H' in both polar-night (February-March) and polar-day (July-August). However, during polar night, the increased H' was

underpinned by a reduced R and increased E, while in polar day, it was driven by an increased R. These temporal fluctuations in microeukaryotic diversity are in contrast to observations from the temperate San Pedro time-series (SPOT), where *H'* was shown to be invariable over time across the top 500 m of the water column[13]. The difference may indicate more seasonal structuring of microeukaryotic communities in high-latitude ocean ecosystems as a consequence of the pronounced environmental variability, supported by the negative correlation of MLD and Richness (Pearson's R = -0.48, 95% CI [-0.62,-0.31], $p = 7.2 \times 10^{-7}$). Furthermore, the distinct trends observed between prokaryotes and microeukaryotes provides evidence that the diversity of these communities is shaped by different forces.

We next assessed how the composition of communities was structured over time. We observed a coherent structuring of prokaryotic and microeukaryotic communities based on the month of sampling. That is, communities sampled from the same month across years were often more similar than from other months in the same year (Figure 3a and 3b), reflecting an annual ecosystem clock. For prokaryotic communities, this pattern aligns with previous observations of month-based clustering at the Banyuls Bay Microbial Observatory in the Mediterranean[31] as well as lowest pairwise beta-diversity at 12-month intervals reported from the temperate English Channel L4[32] and SPOT[13] time-series. However, in contrast, microeukaryotic communities showed weaker temporal structuring over multi-annual scales in the SPOT time-series[13]. By comparing the within- and between-month dissimilarities across years through the convex hull areas within the NMDS ordination (Figure 3 and Supplementary Table S6), we demonstrate a clear distinction in the temporal recurrence of prokaryotic and microeukaryotic composition. Prokaryotic communities were more cohesive across years in February-March and more variable during June-July, with convex hull areas of ~0.025 and ~0.13 respectively. In contrast, microeukaryotic communities were more cohesive during August and more variable during January-March, with convex hull areas of 0.06 and ~0.43 respectively. However, maximal inter-annual differences in microeukaryotic communities were observed in April, in the phase before the spring bloom. This indicates that the assembly of the spring bloom is less predictable, and only later in the summer, the increased richness of microeukaryotic populations (high R) assembles into a cohesive structure each year. In previous years, a high inter-annual variability during polar day has been observed in microeukaryotic communities in this region[33,34], so this pattern may change with time and reflect climate-induced or natural decadal variations[34]. The recurrent structuring of microeukaryotic communities during the productive season could be anticipated to stimulate predictable dynamics in prokaryotes, owing to specialised substrate niches and specific interactions[35,36] as observed in temperate systems[7,8]. Indeed, it triggers a pronounced response of a minority of prokaryotic populations, evidenced through reduced R and E, but their emergent population structure shows high inter-annual variability, which may be a result of selection on a functional level and stochastic processes. In contrast, polar night conditions manifest species-rich but compositionally cohesive prokaryotic communities, suggesting that vertical mixing drives replenishment back to a "standing stock".

Temporal dynamics of ASVs and community gene content

To gain a deeper understanding of ecosystem dynamics in the WSC, we complemented the amplicon dataset with long-read metagenomes and systematically investigated the temporal fluctuations of ASVs and community gene content. From 47 PacBio HiFi read metagenomes that spanned the four years, we obtained 48.5 million open-reading frames that could be classified as bacterial or archaeal, by comparing to species-representative genomes in the GTDB. The predicted prokaryotic gene sequences were subsequently grouped into 704,158 non-singleton clusters based on a 95% average nucleotide identity cut-off. This, combined with the ASV data, forms a rich multi-year ecosystem catalogue that can be used to elucidate community assembly processes and temporal dynamics on a taxonomic and functional level.

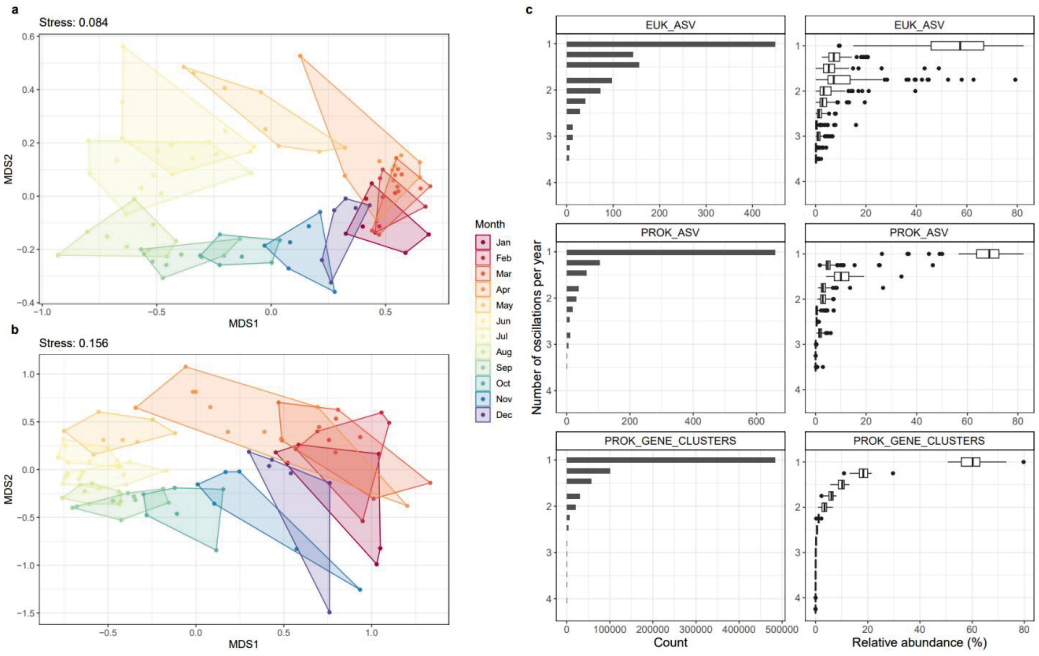

Figure 3. Structuring of community composition and the oscillations of ASVs and genes across years. Non-metric multi-dimensional scaling of Bray-Curtis dissimilarities calculated after rarefying and Hellinger transforming ASV count data for **(a)** prokaryotic and **(b)** microeukaryotic communities over four years, where colours indicate months. **c)** The count (left) and relative abundance (right) of prokaryotic ASVs, microeukaryotic ASVs, and prokaryotic gene clusters for each oscillation signal. The relative abundances of ASVs and genes was subject to Fourier transformation and oscillation signals determined based on amplitude of peak/trough dynamics within each annual cycle. An oscillation of 1 indicates a single peak/trough, reflecting a unimodal annual fluctuation in abundance.

To unravel the temporal dynamics of ASVs and gene clusters, we employed an approach based on Fourier transformations and the determination of oscillation signals. Fourier transformations convert abundance data into frequencies, resulting in wave-like signals that can be evaluated in terms of peak/trough dynamics, hereon termed oscillation signals. We determined that 18% of prokaryotic ASVs, 15% of microeukaryotic ASVs and 69% of gene clusters exhibited a single oscillation each year, reflecting a unimodal fluctuation in

abundance with a single peak and trough (Figure 3c). Although only capturing a fraction of the diversity, these annually oscillating ASVs and gene clusters comprised the majority portion of communities across all time points, with an average relative abundance of 67% for prokaryotic ASVs, 55% for microeukaryotic ASVs, and 60% for gene clusters. These findings expand on previous reports of seasonally recurrent patterns in temperate prokaryotic communities at more coarse-grained resolutions, such as OTUs[10] and higher taxonomic levels[13,22], and are in line with those on microdiversity temporal dynamics from a coastal temperate region[8] – suggesting comparable ecological forcing in high-latitude oceans.

The dominance of annually oscillating ASVs raises the question, “To what extent are these patterns deterministic?”. More than 15 years ago, Fuhrman and colleagues proposed that “annually recurrent microbial communities can be predicted from ocean conditions”[12]. However, only more recently, with technological advancements and increasing quantity and resolution of data, is predictive ecology becoming more feasible. To contribute to this, we assessed the timing and order of ASV oscillations across each annual cycle. We found that 20% of prokaryotic ASVs consistently oscillated in the same order each year, while 51% reached their peak within the same 30-day window. Therefore, while recurrent oscillations are largely bound within temporal windows, the composition of co-occurring populations can vary across years, likely reflecting the influence of trophic interactions and the ephemeral nature of population dynamics[14,37]. A similar pattern was recently described from prokaryotic communities in the NW Mediterranean, where seasonally recurrent taxa exhibited inter-annual changes in the composition of their neighbours within co-occurrence networks[31]. Despite these variations, the periodic timing of recurrent population dynamics within narrow temporal windows provides strong support towards deterministic patterns in ocean ecosystems.

Seasonal recurrence is underpinned by transitions across distinct ecological states

We investigated how the annual recurrence of ASVs and genes translates to ecological and functional shifts within prokaryotic communities. For this, we first grouped the prokaryotic genes into functions. Of the 482,923 annually oscillating gene clusters, 85% were assigned to an orthologous group in the EGGNOG database and were subsequently grouped based on the functional annotation of the matching ortholog. This resulted in 11,320 unique gene functions that captured between 41 – 72% of community gene content across all time points.

To investigate temporal structuring on a taxonomic and functional level, we used the oscillations of prokaryotic ASVs and gene functions to build a correlation-based network. Correlation co-occurrence networks have proven powerful for disentangling community dynamics and organismal interactions over spatial and temporal scales[38]. Here, we compared the oscillation signals of ASVs and gene functions through Pearson’s correlation and retained only those with a strong, positive coefficient (R > 0.7, FDR-

based adjusted $p < 0.05$). We subsequently built a network using these correlation coefficients as the edges and the respective ASVs and gene functions as the nodes. Using the Louvain algorithm, the network was partitioned into five modules comprised of co-oscillating ASVs and gene functions (Figure 4a) that represent

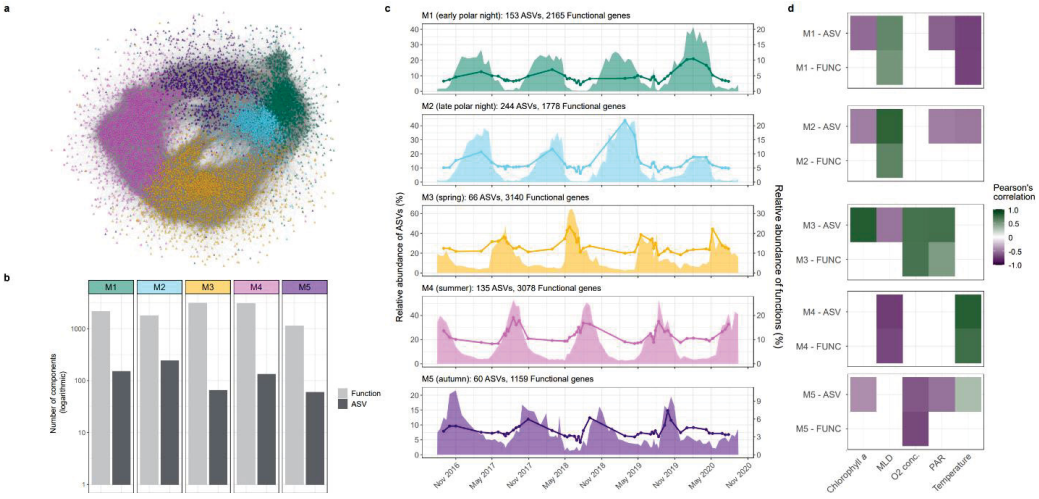

Figure 4. Ecosystem modules, their temporal dynamics and association with environmental conditions. a) Co-occurrence network constructed from significant positive Pearson correlations (>0.7) between prokaryotic ASVs and functional cluster oscillation signals. Circle nodes = ASVs, Triangle nodes = functional clusters. **b)** The temporal relative abundance dynamics of module ASVs (area) and functional clusters (line). **c)** Significant Pearson's correlations between the combined relative abundance of module ASVs and module functional clusters against measured environmental factors (only those with a p <0.05 after multiple testing correction are shown).

distinct temporal periods in the WSC. Each annual cycle was thus characterised by a succession across these five temporal modules (Figure 4b and Supplementary Table S7).

To uncover the ecological shifts associated with the module succession, we compared the composition of modules, the environmental conditions they prevail under, and their association with microeukaryotes. The modules differed markedly in the number of their taxonomic and functional components, from a low number of ASVs (n = 66) and high number of functions (n = 3140) in the spring module M3 to a low number of both functions (n = 1159) and ASVs (n = 60) in the autumn module M5. Beyond quantitative differences, the modules comprised unique taxonomic (Figure 5a) and metabolic signatures (Figure 5b). Module M1, which prevailed from early to mid-polar night, was taxonomically distinguished by *Cand.* Nitrosopumilus, Arctic97B-4 (Verrucomicrobiota) and BD2-11 (Gemmatimonadota) and functionally distinguished by ammonia oxidation (*amoABC*), nitrite reduction (*nirK* and *norBC*) and carbon fixation (hydroxypropionate/hydroxybutyrate cycle[39]). The numerous *Cand.* Nitrosopumilus ASVs detected and their high network connectivity to gene functions within module M1 (Figure 5a) complements previous findings from southward flowing Arctic waters[40] and reports of increased ammonia oxidation rates during polar night in Antarctic coastal waters[41]. Taken together, these observations indicate that ammonia-oxidising Archaea are likely keystone members of prokaryotic communities during winter in high-latitude

ocean ecosystems, irrespective of water mass origin. Microeukaryotic ASVs that co-oscillated with this early polar night module included members of Radiolarians, Nassophorea and MOCH-1 (Figure 6). A shift in dominance from module M1 to M2 marked a transition in the prevailing ecological state during the mid- to late-polar night period. This transition saw the emergence of *Thalassobius* and *Arenicellaceae* members dominating the prokaryotic communities along with an increase in C1 and nitrogen metabolic machinery. The onset of solar radiation and rapid transition from module M2 to M3 signified the end of polar night. The positive correlation of module M3 with PAR and chlorophyll (Figure 4c and Supplementary Table S8) as well as its association with Bacillariophyta, Prymnesiophyceae and Dinophyceae (Figure 6) indicates its representation of the spring phytoplankton bloom period. As could be expected, this module was dominated by members of heterotrophic bacteria known as primary responders to phytoplankton blooms and their carbohydrate exudates in temperate regions[7,35], including *Polaribacter*, *Aurantivirga*, *Formosa* and SAR92. Functionally, module M3 was enriched in carbohydrate-active enzymes, amino and nucleotide sugar metabolism and organosulfur compound utilisation, including DMSP demethylation, DMSO to DMS (*dmsBC*) and methanesulfonate to sulfate (*msmAB*, *sorAB*/SUOX sulfite oxidase). The summer module M4 that preceded the spring bloom was also enriched in sulfur metabolism but included both inorganic and organic sulfur utilisation, such as sulfur oxidation (*soxABCXYZ*), sulfite reduction (*soeABC*), DMSP demethylation, methanethiol oxidation (*MeSH*), and taurine utilisation (*tauABC*). The prokaryotic community in module M4 was dominated by a single *Amylibacter* ASV. *Amylibacter* are key contributors to organosulfur compound metabolism in temperate coastal waters during summer[42,43]. However, we could also attribute *Amylibacter* to the machinery to oxidise sulfur (SOX system) and reduce sulfite (sulfite reducatese), suggesting their specialisation on more diverse sulfur sources[44]. Module M4 was also signified by motility

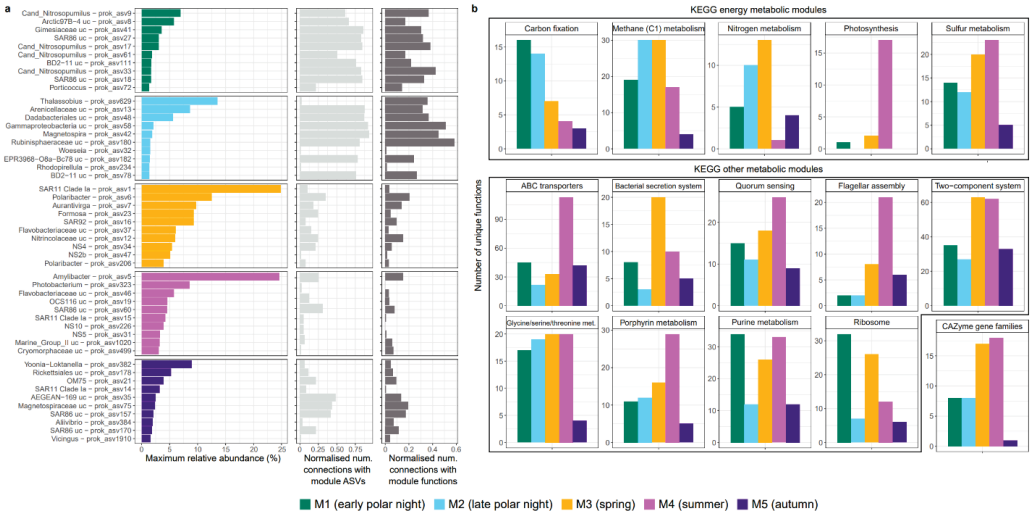

Figure 5. Modules are phylogenetically and functionally distinct. a) Ten most abundant ASVs identified in each module, along with the number of connections they share with other ASVs and functions from the same network module. The number of network connections was normalised by the total number of nodes in each module. **b)** Composition of KEGG metabolic pathways that exhibited the largest variance in the number of assigned functions between modules along with the number of carbohydrate-active enzyme gene families.

machinery, quorum sensing and photosynthesis. The photosynthesis machinery was associated with *Synechococcus*, which represented up to 1.7% of the prokaryotic community during this period. *Synechococcus* is a significant contributor to primary production in tropical and temperate oceanic regions but is largely absent from polar waters. Our observations of increased *Synechococcus* abundances in late summer mirrors recent findings from the WSC[45] and suggests that progressing Atlantification is driving their northward expansion into the Arctic.

The taxonomic and metabolic signatures of each module demonstrate that the prokaryotic microbiome is structured by a succession across five distinct ecological states within each annual cycle. Previous taxonomic-centred analyses of prokaryotic communities from temperate and tropical ecosystems have also reported recurrent dynamics structured by seasonality. However, the proportion of ASVs exhibiting seasonal recurrence and the number of temporal modules appears to be higher in the WSC. For instance, in the North West Mediterranean, only 4% of prokaryotic ASVs, constituting a relative abundance of 47%, were shown to exhibit seasonal recurrence and could be grouped into three distinct seasonal clusters[46]. In a coastal temperate region that lacks pronounced phytoplankton blooms but experiences large environmental variability, recurrent dynamics of prokaryotes are primarily partitioned into summer and

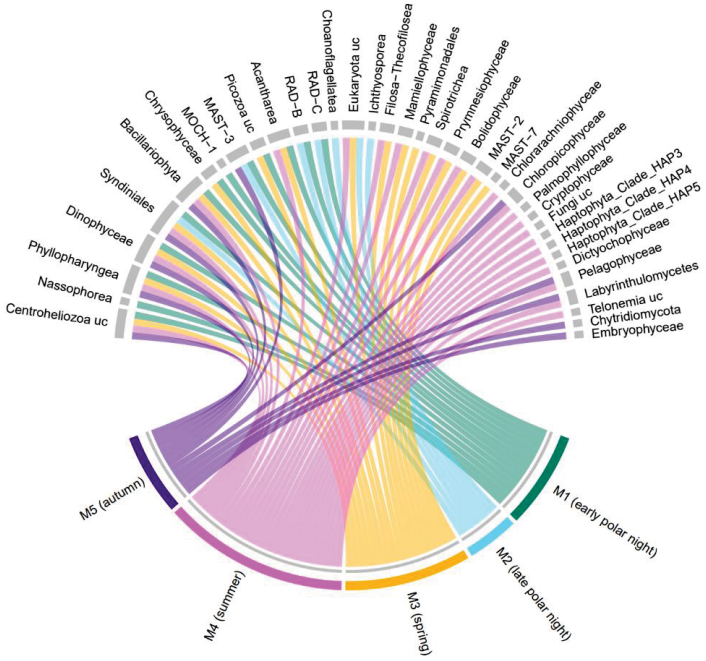

Figure 6. Correlations between prokaryotic and microeukaryotic ASV oscillations. Pairwise Pearson's correlations were computed between module associated prokaryotic ASVs and microeukaryotic ASVs. Correlations with a coefficient > 0.7 and p <0.05, after multiple testing correction, were used to construct the chord diagram. Microeukaryotic ASVs are grouped at higher taxonomic ranks.

winter groups[47]. The higher prevalence of recurrent dynamics and their organisation into narrower temporal

modules in the WSC may indicate stronger selective pressure arising from the more pronounced seasonal environ mental variability.

Selection pressure varies across ecological states

The recurrent dynamics of ASVs, genes and functions and their assembly into cohesive, ecological modules suggests that the WSC microbiome is predominantly shaped by deterministic processes. The primary mechanisms that contribute to driving deterministic structuring of microbiomes are environmental selection and organismal interactions. Environmental selection is considered the primary force shaping the distribution of populations and the assembly of microbial communities in ocean ecosystems[48]. However, we lack a mechanistic understanding of how environmental selection operates. In particular, it is unclear whether the environment primarily selects for a function or for a specific organism with a function. Recent evidence has shown that the taxonomic composition of a microbiome can change while the gene content remains conserved[49,50], reflecting the prevalence of metabolic redundancy across microbial taxa. In addition, studies on surface ocean microbial communities have demonstrated incongruencies in structuring on a taxonomic and functional level, with strong selection pressure on functional groups but weak selection pressure on the taxonomic composition within functional groups[18]. These observations provide evidence for the decoupling of taxonomy and function within microbiomes and suggests that selection may act on a functional level.

Here, we showcase that selection acts heterogeneously across seasonal periods, with differential pressures on a taxonomic and functional level. To demonstrate this, we focus on the late polar night module M2 and spring module M3, as they represent two contrasting scenarios. The spring module M3 comprises a nearly two-fold larger diversity of gene functions than the polar night module M2, but fourfold less ASVs. Hence, spring is underpinned by annually recurrent dynamics predominantly on a functional level, compared to both taxonomic and functional recurrence during late polar night. Hence, selection pressure may be stronger on the function than organismal level in spring. However, despite this, both modules contained more functional redundancy than the other seasonal modules, with 10% more multi-gene cluster compared to single-gene cluster functions. Amongst these redundant functions, we identified a strong positive linear relationship between function abundance and the diversity of the contained gene clusters (Figure 7b) in ~55% of cases. As such, the oscillations of these redundant functions are, in part, driven by the concurrent oscillations of metabolically overlapping prokaryotes as opposed to selection for specific populations. To explore this further, we compared the structuring of gene cluster diversity within redundant functions over time. In module M2, 60% of the redundant functions were dominated by the same gene cluster during each annual oscillation, compared to 40% in module M3 (Figure 7c). The increased genetic variability within functions of module M3 supports the notion of stronger functional selection and weaker organismal selection during spring.

Integrating the observations on ASV dynamics further supports seasonal variations in environmental selection pressure. The late polar night features high species richness and evenness and a low variability in community composition across years. As such, the late polar night is characterised by a well-mixed community that is taxonomically and functionally conserved across years. Interestingly, we observed three ASVs during this period that reached a large fraction of the microbial community (Figure 5), which included members of known sub-surface taxa, such as *Dadabacteriales*. Therefore, the deeper vertical mixing during polar night drives a recurrent homogenisation of epipelagic communities, with potentially a weaker environmental pressure that selects for only a few populations upwelled from subsurface waters. Conversely, spring exhibits low species richness and evenness and high inter-annual variability in community composition (Figure 2 and 3). Combined with the observations on gene functions, spring can thus be characterised by the emergence of a prokaryotic community dominated by few populations that can compositionally vary but remain functionally conserved across years. Therefore, selection pressure appears stronger on a functional rather than organismal level in spring.

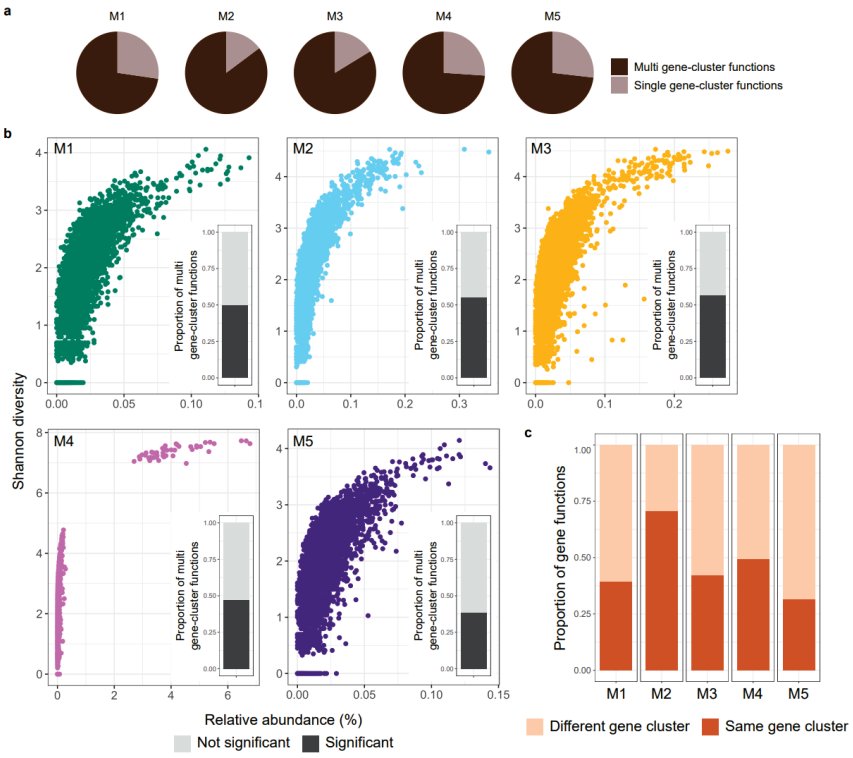

Figure 7. Diversity, abundance and structuring of gene clusters within functions. a) Proportion of functions in each module comprised of multiple or single gene clusters. For those functions that contain >1 gene cluster, we calculated the **b)** Shannon diversity of gene clusters within functions against the relative abundance of the function across all timepoints. Each point represents an individual function. The inserted barplots illustrate the proportion of multi gene clusters functions that exhibit a significant positive linear relationship between Shannon diversity and abundance. **c)** Proportion of multi gene clusters functions that are dominated by the same or different gene cluster during each annual oscillation.

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

CONCLUSION

Our study provides fundamental insights into the seasonal and interannual structuring of prokaryotic and microeukaryotic communities in an Arctic pelagic ocean ecosystem. We demonstrate the prevalence of annually recurrent dynamics of populations and community gene content, which are organised into five distinct seasonal modules. Each of the modules represents distinct ecological states that prevail within the prokaryotic microbiome each year and are connected to specific microeukaryotic populations and environmental conditions. We further provide evidence that environmental selection is heterogeneous across these ecological states, with differential pressures on an organismal and functional level. Our findings provide new insights into understudied yet rapidly changing ocean regions and advances our understanding of how microbial communities are structured across pronounced environmental gradients.

MATERIALS AND METHODS

Sample collection and processing

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

Mooring and satellite data

Attached to the RAS were Seabird SBE37-ODO CTD sensors that measured temperature, depth, salinity, and oxygen concentration. Sensor measurements were averaged over 4 h around each seawater sampling

event. Physical sensors were manufacturer-calibrated and processed in accordance with https://epic.awi.de/id/eprint/43137. Employing multiple CTD sensors along the mooring depths enabled the determination of the minimum mixed layer depth (MLD) at each sampling time point. For instance, if two CTDs showed the same temperature and salinity measurements, the MLD was at least the depth of the deeper CTD. Chlorophyll concentrations were measured via Wetlab Ecotriplet sensors. Surface water Photosynthetically Active Radiation (PAR) data, with a 4 km grid resolution, was obtained from AQUA-MODIS (Level-3 mapped; SeaWiFS, NASA) and extracted in QGIS v3.14.16 (http://www.qgis.org).

SSU rRNA gene amplicon and metagenome sequencing

Filtered seawater samples from 97 time points were subjected to DNA extraction using the DNeasy PowerWater Kit (QIAGEN, Hilden, Germany). 16S and 18S rRNA gene fragments were PCR-amplified using the primers 515F–926R[51] and 528iF–964iR[52], respectively. Sequencing libraries were constructed from rRNA gene products according to the "16S Metagenomic Sequencing Library Preparation" protocol (Illumina, San Diego, CA) and sequenced on an Illumina MiSeq platform in 2 x 300 bp, paired-end mode. Amplicon sequencing took place at the Alfred Wegener Institute. The extracted DNA from 47 timepoints was additionally used to generate PacBio HiFi metagenomes. Sequencing libraries were prepared following the protocol "Procedure & Checklist – Preparing HiFi SMRTbell Libraries from Ultra-Low DNA Input" (PacBio, Menlo Park, CA) followed by inspection with a FEMTOpulse. The libraries were multiplexed and sequenced on 8M SMRT cells (7 - 8 samples per cell) on a PacBio Sequel II platform for 30 h with sequencing chemistry 2.0 and binding kit 2.0. Metagenomes were sequenced at the Max Planck Genome Centre, Cologne, Germany.

PacBio HiFi metagenome analysis

A total of 48 Gbp of PacBio HiFi reads were generated, with an average of 1 Gbp per sample. Gene sequences were predicted on HiFi reads using Fraggenescan (v1.31; parameters:-complete=1 -train=sanger_5)[53]. The genes were subsequently clustered using cd-hit v4.8.1[54] at a 95% identity threshold – these comprise the 'gene clusters'. To facilitate comparisons between metagenomes, gene cluster counts were normalised by the estimated number of prokaryotic genomes in each sample, determined from the average sequencing depth of 16 single-copy ribosomal proteins, as described previously[55]. The longest sequence from each cluster was used as the representative for functional assignment against the EGGNOG v5.0 database[56] using the eggnog-mapper tool v2[57]. To build the 'functional clusters' used in our analysis, we grouped gene clusters based on functional annotations of matching seed orthologs. As the seed orthologs of EGGNOG have been functionally annotated using numerous resources, we grouped gene clusters firstly based on annotations against the carbohydrate-active enzyme database, followed by KEGG and then PFAM.

Taxonomic diversity analyses

The 16S and 18S amplicons were processed into Amplicon Sequence Variants (ASV) using DADA2[58] in RStudio v4.1.3[59]. ASVs were taxonomically classified using the SILVA SSU v138 (16S) and PR2 v4.12 (18S) databases using the *assignTaxonomy* function of DADA2. The ASV dataset was filtered to remove those with <3 counts in <3 samples. For alpha diversity, the ASV count table was subject to 100 iterations of rarefying followed by the calculation of Richness, Shannon diversity (*diversity* function in *vegan*) and Evenness. The mean and standard deviation was calculated for each metric, with the mean being used for visualisations and statistical comparisons to environmental metadata. For beta diversity analysis, the ASV count table was rarefied a single time, followed by Hellinger transformation before calculation of Bray-Curtis dissimilarities (using *vegan*). Dissimilarities were ordinated using Non-Metric Multi-Dimensional Scaling.

Time-series and network analysis

The temporal analysis of ASVs and gene clusters as well as the construction of co-occurrence networks was performed using the same workflow as that presented recently by authors of this paper[60]. In brief, for each ASV and gene cluster, we calculated a time-series signal using the signal using the *segmenTier*/*segmenTools* (https://github.com/raim/segmenTools) packages https://www.nature.com/articles/s41598-017-12401-8) using relative (ASV) and normalized (gene cluster) abundance data. The recovered signals indicate the frequency, amplitude and phase of abundances in each annual cycle, which we termed oscillation signals. Oscillation signals were extracted for each ASV and gene cluster using the *fprocessTimeseries* function from the *segmenTier* package. The ASVs and genes with an oscillation signal of one, i.e. a single peak and trough in each annual cycle, were deemed as "annually oscillating" and retained for further analysis. Phase-Rectified Signal Averaging (PRSA) was used to visualize periodic patterns of ASVs, using phase-rectified data to remove phase variability. Then, the phase-rectified data was averaged for the final PRSA plot. The oscillation signals of ASVs and gene clusters were compared through pairwise Pearson's correlation, with multiple testing corrections using the FDR method. Those with a statistically significant ($p < 0.05$) positive correlation coefficient of >0.7 were used to build a co-occurrence network, with edges as correlation coefficients. By using oscillation signals, focusing on ASVs and gene clusters with a defined oscillation, and using only positive correlations, we minimise noise in the dataset and prevent potential network topology distortion of negative correlations. Co-occurrence correlation networks were constructed using the *igraph* package[61] in R and visualised in Cytoscape v3.7.2[62] using the Edge-weighted Spring-Embedded Layout. ASVs and gene clusters in the network were clustered using the Louvain algorithm[63]. All steps outlined above were performed in R v4.1.3.

DATA AVAILABILITY

Mooring data are available under https://doi.pangaea.de/10.1594/PANGAEA.904565 (2016-2017), https://doi.pangaea.de/10.1594/PANGAEA.904534 (2017-2018), https://doi.pangaea.de/10.1594/PANGAEA.941126 (2018-2019), and https://doi.pangaea.de/10.1594/PANGAEA.946508 (2019-2020). ENA accession numbers for 16S rRNA

amplicons are: PRJEB43890 (2016-2017), PRJEB43889 (2017-2018), PRJEB67813 (2018-2019), PRJEB66202 (2019-2020). ENA accession numbers for 18S rRNA amplicons are: PRJEB43504 (2016-2017), PRJEB43885 (2017-2018), PRJEB66212 (2018-2019), PRJEB66220 (2019-2020). Raw metagenomic reads are available under PRJEB67368. Code for reproducing workflow and figures is available at https://github.com/tpriest0/Fram_Strait_WSC_time_series_2016-2020 while the accompanying derived data files are available at https://doi.org/10.17617/3.CA8MQY.

ACKNOWLEDGEMENTS

We thank Jana Bäger, Theresa Hargesheimer, Rafael Stiens and Lili Hufnagel for RAS operations; Daniel Scholz for RAS and sensor operations; Normen Lochthofen, Janine Ludszuweit, Lennard Frommhold and Jonas Hagemann for mooring operations; Jakob Barz, Swantje Ziemann and Anja Batzke for DNA extraction, library preparation and sequencing; and Bruno Huettel and the technicians at the Max Planck Genome Centre in Cologne for sequencing. The captain, crew and scientists of RV Polarstern cruises PS99.2, PS107, PS114, PS121 and PS126 are gratefully acknowledged. This project has received funding from Polarstern grants AWI_PS99_00, AWI_PS107_05, AWI_PS114_01, AWI_PS121_07, AWI_PS126_05, and AWI_PS126_07. Further support came from the European Research Council (ERC) under the European Union's Seventh Framework Program (FP7/2007-2013) research project ABYSS (Grant Agreement no. 294757) to AB, from the Helmholtz Association, specifically for the FRAM infrastructure, and from the Max Planck Society.

AUTHOR CONTRIBUTIONS

TP designed and executed the ASV and metagenomic analysis pipelines and wrote the manuscript, with input from MW, EO and BD. EO and OP performed time-series and network analyses. MW processed amplicon raw data into ASVs, co-designed the sampling and mooring strategy, and coordinated data analysis. WJvA contributed quality-controlled oceanographic data and coordinated the mooring operations. KM coordinated the processing of samples and sequencing and provided 18S rRNA gene sequence data. STV provided quality-controlled chlorophyll sensor data. CB, KM and AB co-designed the sampling and mooring strategy, and contributed to the interpretation. BF and RA contributed to study design and interpretation. All authors contributed to the final manuscript.

REFERENCES

1. Fuhrman, J. A., Cram, J. A. & Needham, D. M. Marine microbial community dynamics and their ecological interpretation. *Nat. Rev. Microbiol.* **13**, 133–146 (2015).
2. Rodriguez-Brito, B. *et al.* Viral and microbial community dynamics in four aquatic environments. *ISME J.* **4**, 739–751 (2010).
3. Falkowski, P. G., Barber, R. T. & Smetacek, V. Biogeochemical controls and feedbacks on ocean primary production. *Science* **281**, 200–206 (1998).
4. Azam, F. *et al.* The ecological role of water-column microbes in the sea. *Mar. Ecol. Prog. Ser.* **10**, 257–263 (1983).

5. Behrenfeld, M. J. & Boss, E. S. Resurrecting the ecological underpinnings of ocean plankton blooms. *Annu. Rev. Mar. Sci.* **6**, 167–194 (2014).
6. Eiler, A., Hayakawa, D. & Rappé, M. Non-random assembly of bacterioplankton communities in the subtropical North Pacific Ocean. *Front. Microbiol.* **2**, (2011).
7. Teeling, H. *et al.* Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *eLife* **5**, e11888 (2016).
8. Chafee, M. *et al.* Recurrent patterns of microdiversity in a temperate coastal marine environment. *ISME J.* **12**, 237–252 (2018).
9. Cram, J. A. *et al.* Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *ISME J.* **9**, 563–580 (2015).
10. Gilbert, J. A. *et al.* Defining seasonal marine microbial community dynamics. *ISME J.* **6**, 298–308 (2012).
11. Käse, L. *et al.* Rapid succession drives spring community dynamics of small protists at Helgoland Roads, North Sea. *J. Plankton Res.* **42**, 305–319 (2020).
12. Fuhrman, J. A. *et al.* Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc. Natl. Acad. Sci.* **103**, 13104–13109 (2006).
13. Yeh, Y.-C. & Fuhrman, J. A. Contrasting diversity patterns of prokaryotes and protists over time and depth at the San-Pedro Ocean Time series. *ISME Commun.* **2**, 1–12 (2022).
14. Needham, D. M. & Fuhrman, J. A. Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nat. Microbiol.* **1**, 16005 (2016).
15. Martin-Platero, A. M. *et al.* High resolution time series reveals cohesive but short-lived communities in coastal plankton. *Nat. Commun.* **9**, 266 (2018).
16. Berdjeb, L., Parada, A., Needham, D. M. & Fuhrman, J. A. Short-term dynamics and interactions of marine protist communities during the spring–summer transition. *ISME J.* **12**, 1907–1917 (2018).
17. Martiny, J. B. H., Jones, S. E., Lennon, J. T. & Martiny, A. C. Microbiomes in light of traits: A phylogenetic perspective. *Science* **350**, aac9323 (2015).
18. Louca, S., Parfrey, L. W. & Doebeli, M. Decoupling function and taxonomy in the global ocean microbiome. *Science* **353**, 1272–1277 (2016).
19. Polyakov, I. V. *et al.* Greater role for Atlantic inflows on sea-ice loss in the Eurasian Basin of the Arctic Ocean. *Science* **356**, 285–291 (2017).
20. Berge, J. *et al.* In the dark: A review of ecosystem processes during the Arctic polar night. *Prog. Oceanogr.* **139**, 258–271 (2015).
21. Ghiglione, J. F. & Murray, A. E. Pronounced summer to winter differences and higher wintertime richness in coastal Antarctic marine bacterioplankton. *Environ. Microbiol.* **14**, 617–629 (2012).
22. Raes, E. J. *et al.* Seasonal bacterial niche structures and chemolithoautotrophic ecotypes in a North Atlantic fjord. *Sci. Rep.* **12**, 15335 (2022).
23. Wietz, M. *et al.* The polar night shift: seasonal dynamics and drivers of Arctic Ocean microbiomes revealed by autonomous sampling. *ISME Commun.* **1**, 1–12 (2021).
24. Whitney, F. A. & Freeland, H. J. Variability in upper-ocean water properties in the NE Pacific Ocean. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **46**, 2351–2370 (1999).
25. Sunagawa, S. *et al.* Structure and function of the global ocean microbiome. *Science* **348**, 1261359 (2015).
26. Ibarbalz, F. M. *et al.* Global Trends in Marine Plankton Diversity across Kingdoms of Life. *Cell* **179**, 1084-1097.e21 (2019).
27. Fuhrman, J. A. *et al.* A latitudinal diversity gradient in planktonic marine bacteria. *Proc. Natl. Acad. Sci.* **105**, 7774–7778 (2008).
28. Hernando-Morales, V., Ameneiro, J. & Teira, E. Water mass mixing shapes bacterial biogeography in a highly hydrodynamic region of the Southern Ocean. *Environ. Microbiol.* **19**, 1017–1029 (2017).
29. Zhang, Y., Xu, H., Qiao, F. & Dong, C. Seasonal variation of the global mixed layer depth: comparison between Argo data and FIO-ESM. *Front. Earth Sci.* **12**, 24–36 (2018).

30. von Appen, W.-J. *et al.* Sea-ice derived meltwater stratification slows the biological carbon pump: results from continuous observations. *Nat. Commun.* **12**, 7309 (2021).
31. Lambert, S., Lozano, J.-C., Bouget, F.-Y. & Galand, P. E. Seasonal marine microorganisms change neighbours under contrasting environmental conditions. *Environ. Microbiol.* **23**, 2592–2604 (2021).
32. Hatosy, S. M. *et al.* Beta diversity of marine bacteria depends on temporal scale. *Ecology* **94**, 1898–1904 (2013).
33. Kubiszyn, A. M., Piwosz, K., Wiktor, J. M., Jr & Wiktor, J. M. The effect of inter-annual Atlantic water inflow variability on the planktonic protist community structure in the West Spitsbergen waters during the summer. *J. Plankton Res.* **36**, 1190–1203 (2014).
34. Nöthig, E.-M. *et al.* Summertime plankton ecology in Fram Strait—a compilation of long- and short-term observations. *Polar Res.* **34**, 23349 (2015).
35. Sarmento, H., Morana, C. & Gasol, J. M. Bacterioplankton niche partitioning in the use of phytoplankton-derived dissolved organic carbon: quantity is more important than quality. *ISME J.* **10**, 2582–2592 (2016).
36. Seymour, J. R., Amin, S. A., Raina, J.-B. & Stocker, R. Zooming in on the phycosphere: the ecological interface for phytoplankton–bacteria relationships. *Nat. Microbiol.* **2**, 1–12 (2017).
37. Choi, D. H. *et al.* Daily variation in the prokaryotic community during a spring bloom in shelf waters of the East China Sea. *FEMS Microbiol. Ecol.* **94**, fiy134 (2018).
38. Lima-Mendez, G. *et al.* Determinants of community structure in the global plankton interactome. *Science* **348**, 1262073 (2015).
39. Könneke, M. *et al.* Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation. *Proc. Natl. Acad. Sci.* **111**, 8239–8244 (2014).
40. Priest, T. *et al.* Atlantic water influx and sea-ice cover drive taxonomic and functional shifts in Arctic marine bacterial communities. *ISME J.* **17**, 1612–1625 (2023).
41. Tolar, B. B. *et al.* Contribution of ammonia oxidation to chemoautotrophy in Antarctic coastal waters. *ISME J.* **10**, 2605–2619 (2016).
42. Liu, Y., Blain, S., Crispi, O., Rembauville, M. & Obernosterer, I. Seasonal dynamics of prokaryotes and their associations with diatoms in the Southern Ocean as revealed by an autonomous sampler. *Environ. Microbiol.* **22**, 3968–3984 (2020).
43. O'Brien, J. *et al.* The microbiological drivers of temporally dynamic dimethylsulfoniopropionate cycling processes in Australian coastal shelf waters. *Front. Microbiol.* **13**, (2022).
44. González, J. M., Hernández, L., Manzano, I. & Pedrós-Alió, C. Functional annotation of orthologs in metagenomes: a case study of genes for the transformation of oceanic dimethylsulfoniopropionate. *ISME J.* **13**, 1183–1197 (2019).
45. Paulsen, M. L. *et al.* Synechococcus in the Atlantic Gateway to the Arctic Ocean. *Front. Mar. Sci.* **3**, (2016).
46. Auladell, A. *et al.* Seasonal niche differentiation among closely related marine bacteria. *ISME J.* **16**, 178–189 (2022).
47. Ward, C. S. *et al.* Annual community patterns are driven by seasonal switching between closely related marine bacteria. *ISME J.* **11**, 1412–1422 (2017).
48. Aalto, N. J., Schweitzer, H. D., Krsmanovic, S., Campbell, K. & Bernstein, H. C. Diversity and Selection of Surface Marine Microbiomes in the Atlantic-Influenced Arctic. *Front. Microbiol.* **13**, (2022).
49. Turnbaugh, P. J. *et al.* A core gut microbiome in obese and lean twins. *Nature* **457**, 480–484 (2009).
50. Louca, S. *et al.* High taxonomic variability despite stable functional structure across microbial communities. *Nat. Ecol. Evol.* **1**, 1–12 (2016).
51. Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* **18**, 1403–1414 (2016).

52. Fadeev, E. *et al.* Microbial communities in the east and west Fram Strait during sea ice melting season. *Front. Mar. Sci.* **5**, (2018).
53. Rho, M., Tang, H. & Ye, Y. FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res.* **38**, e191 (2010).
54. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinforma. Oxf. Engl.* **28**, 3150–3152 (2012).
55. Priest, T., Vidal-Melgosa, S., Hehemann, J.-H., Amann, R. & Fuchs, B. M. Carbohydrates and carbohydrate degradation gene abundance and transcription in Atlantic waters of the Arctic. *ISME Commun.* **3**, 1–13 (2023).
56. Huerta-Cepas, J. *et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314 (2019).
57. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: Functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* **38**, 5825–5829 (2021).
58. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
59. RStudio Team. RStudio: Integrated development of R. (RStudio Inc., 2015).
60. 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).
61. Csardi, G. & Nepusz, T. The igraph software package for complex network research. (2006).
62. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
63. 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).

Chapter 5

Interactions, Stability and Occurence

In this chapter, we present our use case concerning keystone prediction: Keyspecies identification framework. We first briefly introduce the topic "Image Classification using Machine Learning" and subsequently present our approach to identify keystone species based on their 18S abundance data using a framework consisting on three mathematical approaches Figure 5.1.

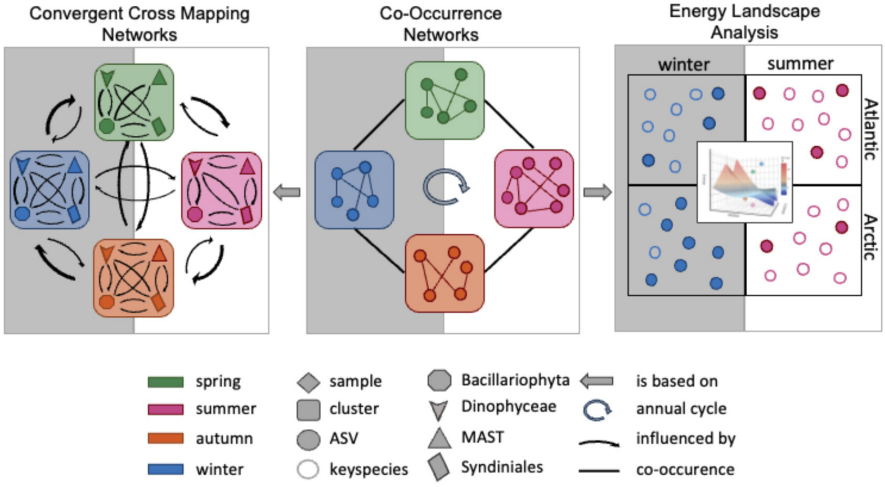

Figure 5.1: **Workflow of the framework.** The middle panel displays the results of clustering achieved through analysis of the Co-Occurrence Network. Each season is presented as a distinct cluster, including spring, summer, autumn and winter. Following this, a Convergence Cross Mapping Network is calculated, based on the connections within the network. It reveals the normalized mutual information score of the four taxonomic class levels within and between clusters, as depicted in the right panel. The ASVs (circles) that represent the samples (diamonds) are organized into seasonal clusters for an Energy Landscape Analysis. Keystone species are identified based on their occupancy within Stable States (rings) in the left panel. Finally, the last panel shows the stable sets that correspond to the typical habitats during winter and summer in the Arctic and Atlantic regions, determined by assessing the abundance of our Atlantic communities (Oldenburg et al., 2024a).

5.1 A novel time series analysis framework predicts seasonal keystone species

In this section, we provide an overview of the contributions and impact of our paper (Oldenburg et al., 2024a):

Ellen Oldenburg, Raphael M. Kronberg, Katja Metfies, Wilken-Jon von Appen, Matthias Wietz, Christina Bienhold, Ovidiu Popa and Oliver Ebenhöh
"Beyond blooms: A novel time series analysis framework predicts seasonal keystone species and sheds light on Arctic pelagic ecosystem stability"
In: *biorxiv, 2024 (Submitted to Communications Earth & Environment)*

Main Results in Simple Terms

In the Arctic Ocean, which is experiencing significant challenges due to climate change, it is crucial to comprehend the interactions between living organisms in this region. Scientists aim to identify the most critical species for maintaining a healthy ecosystem. However, traditional methods are inadequate for this purpose. To address this issue, we have applied three novel techniques to study microorganisms. These methods help to determine the most important species in different seasons and how they interact in the ecosystem. The research discusses the identification of 38 key species that are crucial for the Arctic food web using Co-Occurrence Networks, Convergent Cross Mapping, and Energy Landscape Analysis. The research demonstrates that the types of microbes in the Arctic vary depending on the season, and these key species can provide insight into how environmental changes affect the ecosystem. Microbes can be more active during either the bright summer months or the dark winter. Our research has shown that summer microbes have a significant impact on winter microbes, but not vice versa. In the spring, as the ice begins to melt, two main groups of microbes emerge: those that consume organic matter and those that use sunlight for growth. These groups do not directly affect each other, indicating that each winter acts as a de novo start, allowing certain microbes to flourish. Energy Landscape Analysis revealed that the microbial communities in winter are more stable than those in summer.

Summary/Abstract

A thorough understanding of ecosystem functioning in the Arctic Ocean, a region under severe threat by climate change, requires detailed studies on inhabiting biological communities. The identification of keystone species with special ecological relevance is of great importance, yet difficult to achieve with established community assessments. In the case of microbes, metabarcoding and metagenomics offer fundamental insights into community structure and function, yet remain limited regarding conclusions about the role of individual species within the ecosystem. To overcome this limitation, we have developed an analytical approach based on three different methods: Co-Occurrence Networks, Convergent Cross Mapping, and Energy Landscape Analysis. These methods enable the identification of seasonal communities in microbial

ecosystems, elucidate their interactions, and predict potential stable community configurations under varying environmental conditions. Combining the outcomes of these three methods allowed us to define 38 keystone species that are representative for the different trophic modes that build the local food web. They may serve as indicator species for monitoring the consequences of environmental change in Arctic marine ecosystems. Our research reveals a clear seasonal pattern in the composition of the microbial phytoplankton community, with distinct assemblages characterizing the carbon fixation (light) and consumption (dark) phases. Species interactions exhibit strong seasonality, and we observed summer communities with significant influence on winter communities but not vice versa. During spring thaw, two distinct groups are present: consumers (heterotrophs), strongly linked to the dark phase, and photoautotrophs (mainly Bacillariophyta), initiating growth (photoautotrophic Bacillariophyta). These groups are not causally related, suggesting a "winter reset" with selective effects that facilitates a new blooming period, allowing survivors of the dark phase to emerge. Investigating the fragility of these ecological systems using Energy Landscape Analysis we demonstrate that winter communities are more stable than summer communities. In summary, the ecological landscape of the Fram Strait can be categorized by two distinct phases: a production phase governed by specialized organisms that are highly responsive to environmental changes, and a consumption phase dominated by generalist species with enhanced resilience.

Personal Contribution

WJvA, CB, MW and KM are responsible for the sampling design. WJvA contributed oceanographic data. **EO** devised the project, the main conceptual ideas and study outline. **EO** and RMK designed the model and the computational framework, analysed the data and wrote the initial draft. RMK carried out the implementation. **EO**, RMK and OP interpreted the data, conceptualized and drafted the manuscript. OE and OP advised data evaluation and data interpretation and revised and finalized the manuscript. **All authors** contributed to improving the final manuscript, by contributions to the scientific interpretation of the data and the discussion of results.

Importance of the Research and Contribution to this Thesis

Comprehending the functioning of the delicate Arctic ecosystem requires an understanding of the structure and dynamics of its food web. Our research aims to gather comprehensive data on various components, including eukaryotes, bacteria, genomes, zooplankton, and environmental factors, to piece together a clearer picture.

Here, we created a framework to identify co-occurrence, interactions, and stability within microbial communities. This framework allows us to model how these communities change in response to shifts in environmental conditions. Our aim is to effectively address our fourth research question, particularly in identifying keystone species that are vital to Arctic ecosystem health.

To demonstrate the effectiveness of our framework, we validated our findings with existing scientific literature. Our analysis encompassed a broader array of components from the food

webs, not solely eukaryotic data. This research expansion enhances our understanding of ecosystem dynamics, aiding in the identification of keystone species crucial for maintaining Arctic marine ecosystems.

Beyond blooms: A novel time series analysis framework predicts seasonal keystone species and sheds light on Arctic pelagic ecosystem stability

Ellen Oldenburg [1,2,3,◇,*], **Raphael M. Kronberg** [3,4,◇], **Katja Metfies** [5,6], **Matthias Wietz** [3,7,8], **Wilken-Jon von Appen** [9], **Christina Bienhold** [3,7], **Ovidiu Popa** [1,△] **and Oliver Ebenhöh** [1,2,△]

[1] *Institute of Theoretical and Quantitative Biology, Heinrich Heine University, Düsseldorf, 40225, North Rhine-Westphalia, Germany*
[2] *Cluster of Excellence on Plant Sciences, Heinrich Heine University, Düsseldorf, 40225, North Rhine-Westphalia, Germany*
[3] *Deep-Sea Ecology and Technology, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany*
[4] *Mathematical Modelling of Biological Systems, Heinrich Heine University, Düsseldorf, 40225, North Rhine-Westphalia, Germany*
[5] *Polar Biological Oceanography, Alfred-Wegener-Institut Helmholtz Center for Polar and Marine Research*
[6] *Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg, Oldenburg, Germany*
[7] *Max Planck Institute for Marine Microbiology, Bremen, Germany*
[8] *Institute for Chemistry and Biology of the Marine Environment,University of Oldenburg, Oldenburg, Germany*
[9] *Physical Oceanography of the Polar Seas, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany*

Correspondence*:
Ellen Oldenburg - Universitätstraße 1, 40225 Düsseldorf, Germany
Ellen.Oldenburg@hhu.de

◇ Equal contribution and first authorship: These authors contributed equally to this work and share first authorship.
△ Last authorship: These authors share last authorship.

ABSTRACT

A thorough understanding of ecosystem functioning in the Arctic Ocean, a region under severe threat by climate change, requires detailed studies on linkages between biodiversity and ecosystem stability. The identification of keystone species with special relevance for ecosystem stability is of great importance, yet difficult to achieve with established community assessments. In the case of microbes, metabarcoding and metagenomics offer fundamental insights into community structure and function, yet remain limited regarding the ecological relevance of individual taxa. To overcome this limitation, we have developed an analytical approach based on three different methods: Co-Occurrence Networks, Convergent Cross Mapping, and Energy Landscape Analysis. These methods enable the identification of seasonal communities in microbial ecosystems, elucidate their interactions, and predict potential stable community configurations under varying environmental conditions. Combining the outcomes of these three methods allowed us to define 38 keystone species in the Arctic Fram Strait that represent different trophic modes within the food web, and might signify indicator for ecosystem functionality under the impact of environmental change. Our research reveals a clear seasonal pattern in phytoplankton composition, with distinct assemblages characterizing the phases of carbon fixation (polar day) and consumption (polar night). Species interactions exhibited strong seasonality, with significant influence of summer communities on winter communities but not vice versa. Spring harbored two distinct groups: consumers (heterotrophs), strongly linked to polar night, and photoautotrophs (mainly Bacillariophyta). These groups are not causally related, suggesting a "winter reset" with selective effects that facilitates a new blooming period, allowing survivors of the dark phase to emerge. Energy Landscape Analysis showed that winter communities are more stable than summer communities. In summary, the ecological landscape of the Fram Strait can be categorized into two distinct phases: a production phase governed by specialized organisms that are highly responsive to environmental variability, and a heterotrophic phase dominated by generalist species with enhanced resilience.

Keywords: Cross Convergence Mapping, Co-Occurrence Network, Energy Landscape Analysis, Keystone species, Eukaryotes

1 INTRODUCTION

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 average (Meredith et al., 2019), linked to reduced in sea-ice and snow cover, which exacerbates warming trends. In particular, the extent of Arctic sea ice has declined (Meredith et al., 2019). These environmental changes have a wide range of consequences, including profound shifts in biodiversity (Sala et al., 2000), 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 northwards (Kraft et al., 2013), 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 stability (Meredith et al., 2019). 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 urgently 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 (Eppley and Peterson, 1979; Lin et al., 2003), 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 metazoans (Bluhm et al., 2017). Recent decades have seen a remarkable increase in pelagic phytoplankton and primary production in the Arctic Ocean, a direct consequence of global warming (Arrigo et al., 2015; Lewis et al., 2020; Nöthig et al., 2020).

In the Arctic Ocean (CAO), sea-ice algae rather than phytoplankton account for much of the primary production (Gosselin et al., 1997; Fernández-Méndez et al., 2015)as they have the potential to initiate pelagic blooms beneath the ice (van Leeuwe et al., 2022). Typically, phytoplankton growth starts mainly within the marginal ice zone in spring, co-occurring with increased solar radiation and meltwater-induced stratification (Clement Kinney et al., 2020). Over the past three decades,

increasing evidence has documented the occurrence of under-ice blooms in the Arctic Ocean (Strass and Nöthig, 1996; Fortier et al., 2002; Leu et al., 2011; Assmy et al., 2017), while phytoplankton in the water column below the ice shows significant differences from the microalgal communities in the sea ice (Hardge et al., 2017). However, changes in biodiversity 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 variable (Zhao et al., 2023). Nevertheless, the underlying mechanisms remain a subject of debate and limited understanding (Loreau and De Mazancourt, 2013). The presence of nearly 2,000 phytoplankton taxa and 1,000 ice-associated protists in the Arctic (Bluhm et al., 2011) indicates the relevance of identifying keystone species in this wealth of Arctic marine microbial diversity that account for ecosystem stability (Frey, 2017; Barber et al., 2015; Richter-Menge and Farrell, 2013).

Understanding biological and ecological dynamics across seasonal environmental gradients is substantially fostered by novel statistical approaches. In polar ecosystems, these gradients, including 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 (Priest et al., 2023; Ma et al., 2016). 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 another (Ives and Carpenter, 2007). Cross-convergence mapping (CCM) helps to identify 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 (Sugihara et al., 2012; Suzuki et al., 2020, 2021).

In this study we establish a mathematical methodology to reveal seasonal patterns, suggest causal ecological relationships and identify microbial key species in Western Fram Strait. This major ‚gateway between Arctic and Atlantic Oceans has been studied for over 20 years within the framework of the Longterm ecological research site HAUSGARTEN and FRAM observatories (Soltwedel et al., 2016). 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 (Metfies et al., 2017; Wietz et al., 2021; von Appen et al., 2021; Cardozo-Mino et al., 2023; Wietz et al., 2024). Based on a four-year metabarcoding dataset of microeukaryotic taxa in context of rich oceanographic data, sampled year-round in approx. biweekly intervals, we develop scenarios of their long-term resilience. Additionally, we predict taxa that play a crucial role in maintaining stable communities within the 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 significantly improve our understanding of the effects of climate change on this ecosystem.

2 METHODS

2.1 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 (Soltwedel et al., 2005; Oldenburg et al., 2024). Each RAS contains 48 sterile bags, each collecting water samples of 500 mL at programmed sampling intervals . The samples collected from 2016 to 2018 reflects the pool of up to two samples collected one hour apart in two individual bags. Since 2018, we pooled samples taken 7 to 8 days apart from two consecutive weeks. The samples were preserved by adding 700 µl of mercuric chloride (7.5% w/v) to the bags prior to sampling. Following RAS recovery, water 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 528iF (GCGGTAATTCCAGCTCCAA) and 926iR (ACTTTCGTTCTTGATYRR). The resulting amplicon sequence variants (ASVs) were classified using the PR2 4.12 database (see Supplementary - Methods). We normalised raw ASV counts for CON and CCM using the Hellinger transformation but did not for the energy landscape analysis; hence a different normalisation is introduced in the for the rELA implementation (Suzuki et al., 2021).

Temperature, salinity and oxygen concentration were measured with a CTD-O_2 attached to the RAS. Physical oceanography sensors were manufacturer-calibrated and processed as described under (von Appen et al., 2021). 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.pangaea.de/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 following (von Appen et al., 2018) 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-2 (Spreen et al., 2008) 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 15km 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 $\sim \mu g\ l^{-1}$), PAR (μ mol photons $m^{-}2d^{-}1$), Salinity (PSU), oxygen concentration (O2_conc μmol l^{-1}) and sampling depth (depth m) (von Appen et al., 2021),.

2.2 Co-Occurrence Network

The abundance of species over the full observation period were converted into temporal profiles by employing Fourier transformation techniques to time-series signals. These temporal profiles rely on the 14 Fourier coefficients. We chose 14 coefficients because they reflect the majority of observed species abundance peaks within the four years. To investigate the similarity of temporal profiles between species pairs, we performed pairwise correlations between the individual temporal profiles, where pairs with higher Pearson correlation value show also a similar temporal profile. Pairs with at least 0.7 ($p<0.05$) Pearson correlation were then visualized in an undirected graph. Only positive correlations were retained to later focus on co-operative relationships. To identify strongly connected components that reflect the existing communities of co-occurring species, we applied the Louvain community detection algorithm (Blondel et al., 2008) 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 Layout (Shannon et al., 2003). The whole co-occurrence network construction is described in Supplementary - Methods Co-Occurrence Network.

2.2.1 Distance between Clusters

To measure the distance between previously defined Louvain communities (clusters), we applied UMAP 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 - Results Figures S4, S5, 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.

2.3 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 species allows predicting the time series of an another species. 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/polsys/ennemi` by Petri Laarne and the Convergent Cross Mapping by Implementation from Javier, Prince `https://github.com/PrinceJavier/causal_ccm` (Javier et al., 2022) to measure the strength of the causal relationship considering also non-linear relations. We could show that the implementation of Normalized Mutual Information (NMI) results into similar findings as the original implementation based on Pearson correlation (Veilleux, 1979; Sugihara et al., 2012) (see Supplementary - Methods).

Using a permutation approach (Ma et al., 2016) 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 - Results). The whole CCM network construction and validation are described in Supplementary - Methods and Supplementary - Results (Convergent Cross Mapping) of the Supplementary Information.

2.3.1 Aggregation on cluster level

We simplify the network of interactions between single species 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 species belonging to the respective clusters. This process effectively reduces the number of items in the node cloud, representing clusters through a unified composite node.

2.4 Energy Landscape Analysis

Energy Landscape Analysis is a method based on statistical physics. From data for many points in time, which contain species 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. Moreover, because this function also depends on environmental variables, we can predict how stable a given community is under perturbed environmental conditions. Assessing stable community states and how they change across environmental shifts is crucial for comprehending the resilience and adaptability of ecosystems in the face of environmental challenges. Our analysis focused on the Top 100 most abundant ASVs within each cluster. Outliers were excluded solely for the purpose of plotting. Details of our analysis, including parameters and thresholds applied, are described in in 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. Details of our analysis, including parameters and thresholds applied, are described in in Supplementary - Methods (Energy

Landscape Analysis). Our analysis focused on the Top 100 ASVs within each cluster. Outliers were excluded solely for the purpose of plotting.

2.5 Keystone Species Definition

After collecting attributes from a 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 i) a significant influence on other organism in the network (significant NMI value), ii) a high centrality (closeness) value within its co-occurrence community and iii) 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 (Guimera and Nunes Amaral, 2005; Joyce et al., 2010).

2.6 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 maximum month 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.

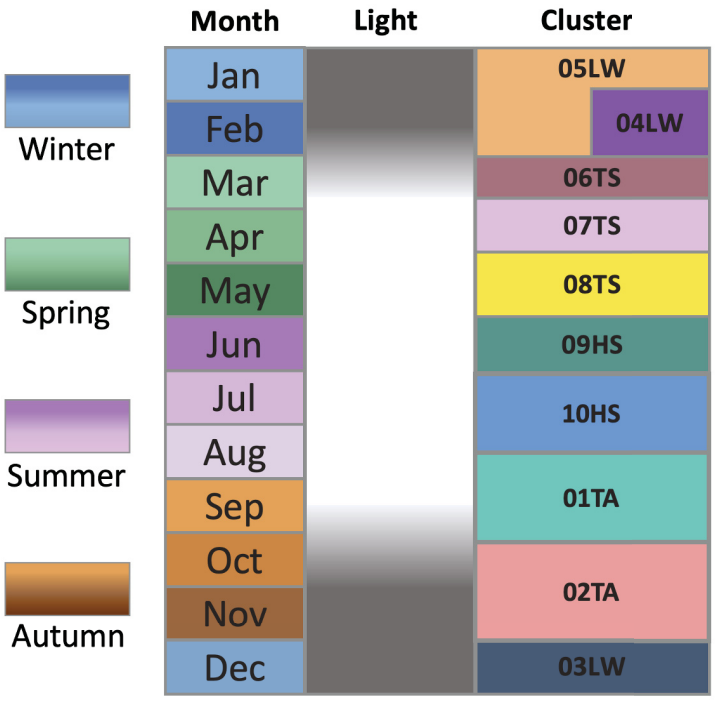

Figure 1. (Schematic) Cluster definition: Names are based on light availability (as defined in Oldenburg et al. (2024)), categorized in transition areas between dark and light (T), high light (H) and low light (L) phases based on PAR parameter (Figure 4) and the season spring (S), summer(S), autumn (A) and winter (W).

3 RESULTS

We examined a dataset of 1,019 eukaryotic ASVs and eight environmental parameters compiled over four years at mooring site F4 in the West Spitsbergen Current (WSC) in Fram Strait. The aim was to characterize species communities, analysing causal relationships between ASVs and to identify keystone and resilient species 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 species in a co-occurring community and to identify stable community states across different environmental conditions.

3.1 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 8,610 edges. In the following, we focus on this major connected component. The resulting undirected graph notably displays a clear seasonal cyclic pattern (Figure 2 A).

The network was partitioned using the Louvain community detection algorithm (Blondel et al., 2008), revealing ten discrete community clusters (Figure 2 B) labeled by the season in which the majority of clusters members had their maximum abundance (Table 1). To further group the clusters, we submerge each three month period to 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 season (09HS, 10HS).

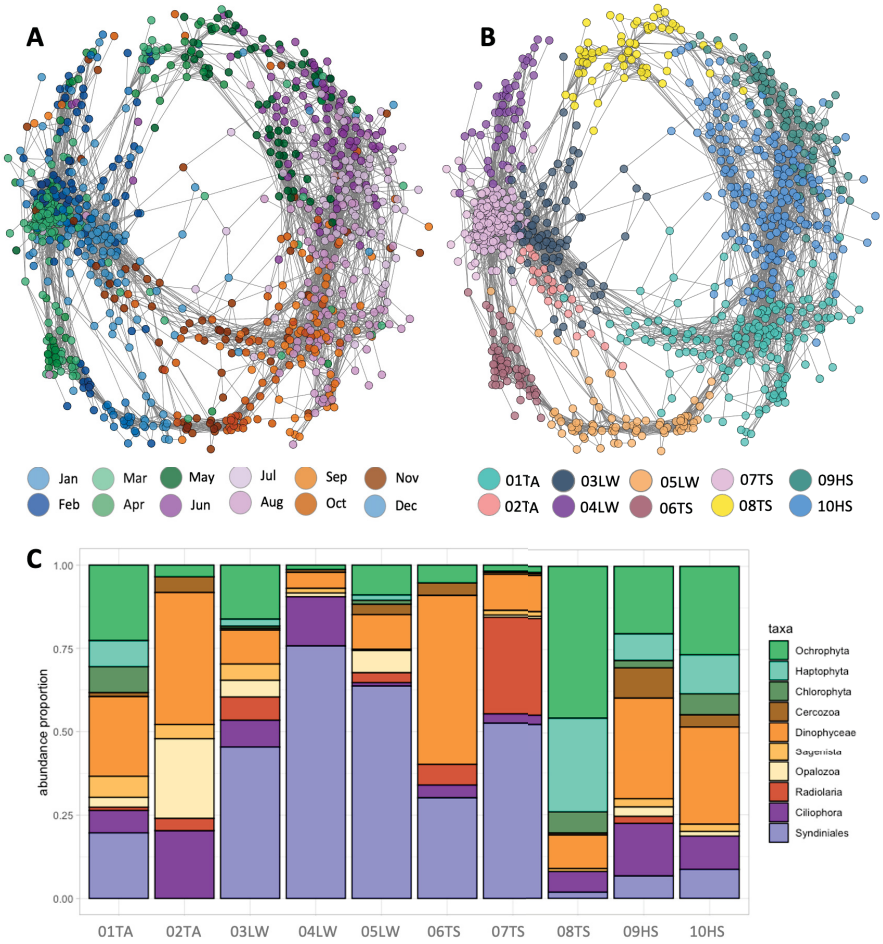

Figure 2. Co-Occurrence Networks and microeukaryotic composition at mooring F4 from 2016-08-01 to 2020-09-17.
Each network node represents an ASV, and each edge represents a similar temporal pattern of two ASVs. The edge weights correspond to the Pearson correlation coefficients determined from the comparison of the individual ASV temporal profiles. ASVs are connected if the coefficient is $r > 0.7, p < 0.05$. **A:** Node color reflect the month in which the ASV exhibit maximal abundance, calculated from the maximum abundance mode for each year ranging from January to December. **B:** In this representation, nodes are coloured based on the community membership that was determined by the Louvain community detection algorithm. **C:** The relative abundance of the top 10 taxonomic classes by cluster ('HS' high light summer , 'LW' low light winter, 'TS' corresponds to transition spring and 'TA' transition autumn). Colour shades illustrate the assignment to auto- (green), mixo- (orange) or heterotroph (purple) .

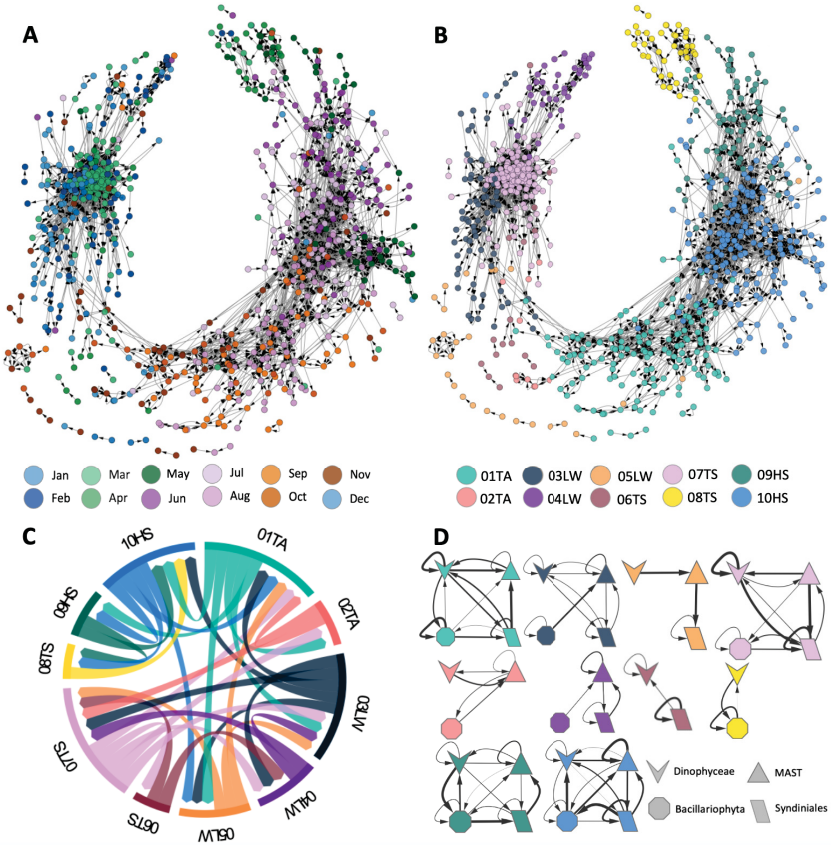

Figure 3. Convergence Cross Mapping Networks of microeukaryotes at mooring F4 from 2016-08-01 to 2020-09-17.
Each node in the CCM network represents an ASV, and each edge represents the causal influences. The edge weight corresponds to the Normalized Mutual Information determined from the comparison of the individual ASV and their predicted representation in the shadow manifold. ASVs are connected if the smoothed p-value of the weight is $p < 0.05$. **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 coloured 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 Normalised Mutual Information aggregated across the edges between the clusters, visually represented by thickness of the arrows corresponding to their respective values. Corresponding colour 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.

Table 1. Co-occurrence network clusters. 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.

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

3.2 Community composition

We explored the taxonomic community 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 - Results Figure S1). The beta biodiversity, measured by Bray-Curtis distance (Supplementary - Results Figure S2), between the winter and spring clusters (03LW, 04LW, 05LW and 06TS) is notably lower than the most other scores, except that between 01TA and 10HS. Cluster 02TA exhibits 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 - Results Figure S2).

In our study, we found distinct taxonomic compositions within various clusters. Photosynthetic organisms like Ochrophyta and Haptophyta dominate the light phases (Oldenburg et al., 2024). 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 wiht 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 (Figure 2 C).

3.3 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 3.3). The findings clearly show that

there is a stronger causal influence (higher NMI values) between co-occurring species compared to random or to unconnected nodes (Supplementary - Results Figure S8 and S9 and Table S1 and S2). The significance level was set at a nonparametric p-value of less than 0.05 calculated similar to Ma et al. (2016). Trimming edges with non-significant NMI values (Supplementary - Results Figure S10), produced a network graph consisting of 4,597 edges and 719 nodes, divided into 18 disconnected components, with the largest module encompassing 706 nodes with 4,572 edges(Fig. 3 A,B).

Of the total 12,648 edges eliminated during trimming, 18.16% represent connections between species reaching their peak abundance in March. This selective removal has a profound impact on the network structure, (Figure 3 A). In contrast to the co-occurrence network, the causal interaction network "breaks" during spring season, as demonstrated in Figure 3 A (months March and April). The winter cluster 05LW and the spring cluster 06TS collapse (see Figure 3 B), 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 species of Picozoa, Leegaardiella, Acantharea, Dinophyceae, MAST-1, and Syndiniales serve as essential hub nodes throughout the seasonal cycle in the network, highlighting their crucial function in maintaining network stability.

3.4 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 - Results). 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 (distance five to eight). Figure 3 C provides a visual and quantitative representation of these interactions. Subsequent analysis revealed a prevalence of connections at a network distance of two (91.5% of total connections), followed by distances of three (7.4% of total connections), and four (0.7% of total connections) (Supplementary - Results Figure S4, S5, S6).

For each seasonal cluster, we investigate in detail the mutual influence (NMI) of four taxonomic groups selected from the top ten classifications (Figure 2 D): Bacillariophyta, Syndiniales, Dinophyceae, and MAST (all MAST-X variants were classified under MAST). Bacillariophyta primarily comprises photo-autotrophic species (Mann et al., 2017), while Syndiniales include parasitic species, most of them characterized by their heterotrophic lifestyle (Suter et al., 2022). Dinophyceae are known for their diverse array of species and ecological roles, from symbionts to planktonic autotrophs (Lin et al., 2022). MASTs are heterotrophic 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 (Thaler and Lovejoy, 2014; Lin et al., 2022).

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 Figure 3 D). During the spring-summer and summer-fall transition, clusters 02TA and 06TS displayed fewer and weaker connections compared to other clusters (see Figure 3 B). 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. The start of the polar night (cluster 05LW), we detected the most substantial influence from the dinoflagellates (shown by the thickest arrow in Figure 3 D) to the pico-eukaryotic heterotrophic groups Syndinales and MAST, suggesting a crucial 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 zenith of the polar day (Cluster 10HS). Notably, Bacillariophyta (i.e. diatoms) showed the most robust connections during the polar day, owing to their role as predominat phototrophic biomass producer and form 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.

3.5 Community and Environment Interactions

For a more detailed understanding of which environmental conditions align with seasonal community clusters, we conducted a correlation analysis (Figure 4). 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 less deeper mixed layers. 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).

3.6 Energy Landscape Analysis determines stability of microbial communities

Energy landscape analysis assessed 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 Figure 5). 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 (Figure 5 A) the community displayed lower energy values than in other seasons; demonstrating high stability of the autumn community. This demonstrates that the autumn communities exhibit a high stability. For the winter cluster 03LW (Figure 5 B), 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 (Figure 5 C). Here, the interpolated energy landscape shows a broad and shallow minimum, and the energy values of the 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. The findings demonstrated that knowledge of the composition of winter communities does not allow for the prediction of the composition of spring communities, in conjunction with the observation of the CCM analysis and the gap between winter and spring clusters (Figure 3 B). Finally, the energy landscape in summer (Figure 5 D) 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.

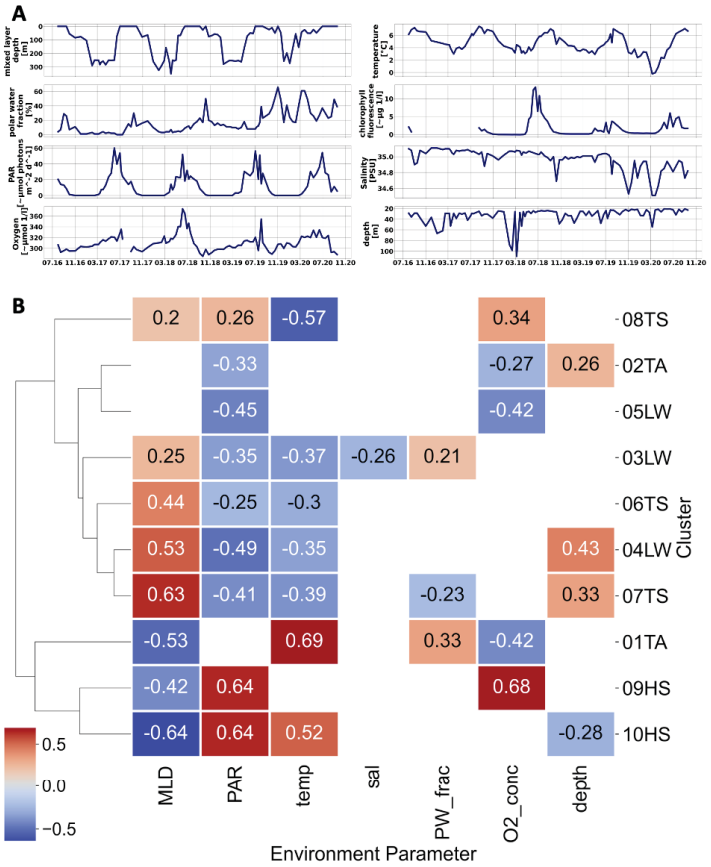

Figure 4. Environmental data and their correlation with Louvain clusters. A): Environmental data for F4 from 2016-08-01 to 2020-09-17. The x-axis represents the time period, while the y-axis indicates the following parameters: Mixed Layer Depth (MLD) [m], Temperature [°C], Chlorophyll Fluorescence [μ g/L], Polar Water Fraction [%], Photosynthetically Active Radiation (PAR) [μ mol photons/m^2/d], Salinity [Practical Salinity Units (PSU)], Oxygen Concentration [μ mol/L], Depth of measurement [m]. **B): Correlations between environmental parameters and seasonal Louvain clusters.** The displayed chart shows the environmental parameters from section a) in relation to seasonal clusters. These clusters are characterized by the cumulative relative abundance of ASVs. 'TA' denotes transition autumn, 'LW' represents low light winter, 'TS' transition spring and 'HS' high light summer. The colour 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.

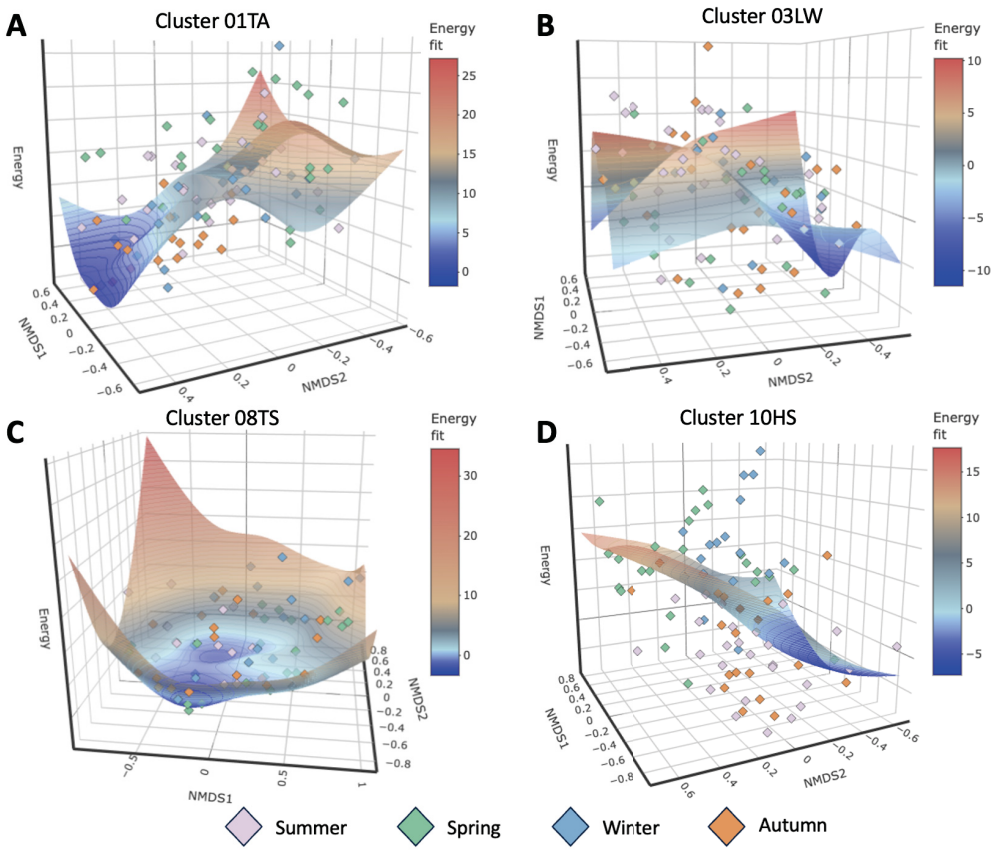

Figure 5. Energy landscapes depicting community structure dynamics. The plots display the reconstructed energy landscape on the NMDS surface for a cluster of each season. Environmental landscapes over the NMDS surface are reconstructed for each of the four example clusters. The z-axis displays the energy, 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.

3.7 Predicting keystone microeukaryotes in the Fram Strait

According to our definition (Section 2.5) 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 nine 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 blooms (Von Quillfeldt, 2000) that also serve as prey for microzooplankton (Cleary et al., 2016; Yang et al., 2015). 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 Eukaryota uc (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 3.6). The emergence of Chlorophyta during early autumn suggests a shift in primary production from Ochrophyta to Chlorophyta, including taxa that may prefer colder temperatures ((Tragin and Vaulot, 2018) and are better adapted to nutrient limitation (Maat et al., 2014).

Table 2. ASV identified as potential keystone species for clusters 10HS, 06TS, 03LW and 01TA. The taxonomic classifications, clusters, raw abundance, proportion of total raw abundance, cluster abundance, and proportion of cluster raw abundance are presented in summary form over the 4-year observation period. The column, significance, indicates if this ASV (Nodes) has at least on significant CCM connection measured in Normalized Mutual Information.

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

3.8 Predicting community stability in altered environments

One critical ecological question is how marine microbial communities change if they experience changing environmental conditions, such as those resulting from climate change. For a first approximation to the potential fate of herein defined clusters under changing conditions, we employed ELA to assess the stability of observed communities under environmental conditions different to those they usually experience. This is done by evaluating the ELA function (see Section 2.4) for given communities and environmental parameters. First, we test the reliability of this approach by evaluating the communities of the summer (10HS) and winter (03LW) clusters under "typical Atlantic" summer or winter days, defined by the averages environmental parameters during a three-month period (see Supplementary - Methods Table S0). As expected the energy values remain almost unchanged when evaluating the summer cluster communities in a typical summer day and the winter cluster communities in a typical winter day (see Supplementary - Results Figure S13). This is to be expected, because the ASVs in the summer clusters assume their most stable community configuration during the summer months, and likewise for the winter cluster. When evaluating the cluster communities under the opposite conditions, i.e. the summer cluster communities with environmental parameters representing a winter day and vice versa, the energy values are drastically changed (see Supplementary - Results Figure S13). Those communities of the summer cluster, which were still present in the winter, exhibit an increased stability (lower energy) than those present in summer. The Atlantic projection panel (Supplementary - Results Figure S13 C and E) display almost opposite oscillatory time courses. Placing communities of the winter cluster in summer conditions has a more differentiated effect. The communities found in winter 2018 and 2020 are highly unstable in this environment, while communities found in summers appear to exhibit an increased stability. This observation allows speculating that some ASVs of the winter cluster might even benefit from warmer summer conditions.

We next investigate a scenario that might result from future temperature increase and accelerated sea-ice melt. Ice-free summers in the Arctic will allow Atlantic water, and with this Atlantic microbial communities, to enter polar regions. We therefore explored the stability of communities from the summer (10HS) and winter (03LW) clusters changes if they experience typical Arctic conditions. For this, we defined "typical" Arctic summer and winter days by selecting extreme environmental parameter values from the central Arctic (see Supplementary - Methods Table S0). Subsequently, we evaluate the energy landscape functions as above with the respective environmental parameters. Interestingly, simulating communities of the summer cluster under Arctic summer or winter conditions results in similar energy values as placing these communities in Atlantic winter days. Specially, communities found in summer are highly unstable whereas communities present in winter show, as a tendency, increased stability (Supplementary - Results Figure S13 E,G,I - left panel, bottom three), indicating that summer cluster communities will face challenges to adapt to Arctic conditions. When comparing the ASVs, which are found in at least one community configuration predicted as stable, we observe that ASVs stable in Arctic environments are almost completely different from those stable in their natural (Altlantic) environment (Supplementary - Results Figure S14, Table S9 and S10). To estimate the importance of the ASVs identified to be stable under natural and Arctic conditions, respectively, we determine their average closeness centrality (Supplementary - Results Figure S15). Clearly, ASVs stable in their natural conditions (and in Atlantic conditions) display a significantly (p-value = $2\text{x}10^{-2}$) higher centrality than average, while those predicted to be stable under Arctic conditions show a reduced centrality. Likewise, the average NMI score for outgoing edges, indicating the mean influence of an ASV on other ecosystem members, is higher than average for ASVs stable under natural (and Atlantic) conditions (p-value = $4\text{x}10^{-2}$), but lower than average for ASVs predicted to be stable under Arctic conditions. The reduced connectivity and weaker influence on other ecosystem members suggests that species of the summer cluster, when exposed to Arctic conditions, will play a less important role in determining ecosystem dynamics and stability.

Simulating the effect of Arctic conditions on winter cluster communities reveals that neither communities present in summer nor winter display a pronounced stability. The only exception appear to be communities during spring 2020. Remarkably, this period was characterized by unusually low temperatures (Figure 4 A). In general the ASVs identified as members of stable

communities in the original conditions, Atlantic winter, or Arctic winter show a four times higher overlap compared with their summer conditions. These observations suggests that communities of the winter cluster might easily adapt to colder Arctic conditions, which is also supported by the observation that closeness centrality and averaged NMI scores (for outgoing edges) are almost unchanged for these three groups of organisms.

4 DISCUSSION

In this study, we developed a new method for investigating ecological time series data based on 18S metabarcoding derived abundance information, by combining three data analysis methods: Co-Occurrence Networks (CON) (Stephens et al., 2009), Convergent Cross Mapping (CCM) (Sugihara et al., 2012), and Energy Landscape Analysis (ELA) (Suzuki et al., 2021). Integrating these three methodological approaches, we aimed to to predict and characterize abundance of keystone microeukaryotes in the West Spitsbergen Current across different seasons, environmental parameters and in relation to other organisms. In addition, we examine how these keystone species are affected by changing environmental conditions, providing insights into potential responses to Arctic warming and Atlantification. We also investigate how different taxa groups affect other taxa groups, and how their effects vary with seasonal shifts and environmental factors.

Our co-occurrence network based on Fourier decomposition differs from previous methods that rely directly on the raw time series signals (Ma et al., 2016; Lima-Mendez et al., 2015; Ma et al., 2020). Our CON based on Fourier decomposition differs from previous methods that rely directly on the raw time series signals (Ma et al., 2016; Lima-Mendez et al., 2015; Ma et al., 2020). The resulting network accurately captured seasonal states and transitions, revealing community clusters that reflect the prevailing community structure (Dunne et al., 2002)): 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 an almost exclusively heterotrophic and parasitic taxa dominate in winter (04LW, 05LW, 07TS). The considerable difference of spring clusters to other seasonal clusters (Supplementary - Results Figure 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 (Jassey et al., 2015; Bruhn et al., 2021; Mitra et al., 2014). According to traditional ecological theory, keystone species are often defined as those with the most biomass (Kang and Fryxell, 1992; Sergeeva et al., 2018). The combination of CON, CCM and ELA allowed predicting keystone species, i.e. ASVs with strongest effects on the interaction network. We found both highly abundant (for example Fragilariopsis or Pseudo-nitzschia diatoms) as well as low-abundant keystone ASVs, suggesting that both common and rare members contribute to ecosystem stability CCM revealed that by far not all co-occurring ASVs actually influence each other (Fig. 3). A striking example is between clusters 06TS and 05LW, which were closely connected in the CON but not in the CCM network(Fig. 3). 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 predators (Zhao et al., 2022) 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. 3). We interpret this gap between winter and spring clusters as a 'winter reset' (Supplementary - Results Figure 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 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 fall to winter are significantly less than between other seasons (except winter to spring) Figure 3 B. All of these causal links are related to Syndiniales, which can be explained by their parasitic lifestyle, foraging upon mixotrophic species that are active during the transition autumn phase.

We compared our approach with two previous studies Cross Convergence Mapping (Ushio, 2022; Fujita et al., 2023). While our approach focuses on the specific interactions within and between clusters, Ushio et al. provides 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 CCM (Ushio, 2022). 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 size (Fujita et al., 2023), while our study utilized pairwise CCM on ASV time series signals within clusters to predict keystone species.

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 (Supplementary - Results Figure S12).

Keystone species represent the ecological roles played by network members in 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.

Energy landscape analysis was used to assess the stability of observed communities when exposed to environmental conditions different from their typical settings. Our results showed that communities maintained stable configurations during typical summer (for communities of the summer clusters) or winter (for winter clusters) days, indicating their adaptability to seasonal variations. However, significant changes in energy values occurred when communities were assessed under opposing conditions, suggesting diverse responses to environmental shifts. Subsequent simulations of typical Arctic conditions revealed interesting patterns within the microbial communities. The summer cluster communities showed a clearly decreased stability under Arctic conditions, in contrast to the winter cluster communities, which displayed a tendency towards increased stability. This suggests that winter communities may have a greater capacity to adapt to colder Arctic conditions compared to their summer counterparts (Supplementary - Results Figure S15). The Arctic projection notably harbored a higher number of unique species, indicating varied responses to environmental changes. Notably, the absence of shared keystone species among these different datasets in the summer cluster 10HS suggests a lower robustness of Atlantic ASVs to environmental shifts, with keystone species candidates exhibiting variability. Under Atlantic conditions, the closeness centrality of the summer cluster 10HS increased more compared to Arctic conditions. In the winter cluster community 03LW, closeness centrality is very similar in both projections (Supplementary - Results Figure S15).

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 and mechanistic insights into the mechanisms responsible for shaping and maintaining spatiotemporal heterogeneity in ecosystem composition (Suzuki et al., 2020).

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTION

WJvA, CB, MW and KM are responsible for the sampling design. KM contributed 18S metabarcoding data. MW processed raw metabarcoding reads into ASVs. WJvA contributed oceanographic data. EO devised the project, the main conceptual ideas and study outline. EO and RMK designed the model and the computational framework and analyzed the data. EO, RMK, OP and KM wrote the initial draft. RMK carried out the implementation. EO, RMK and OP interpreted the data and conceptualized the manuscript. OE and OP 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.

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 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.

FUNDING

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)(EO and OE). Additional funding came from the Helmholtz DataHub Information Infrastructure funds within project iLOVE (EO and RMK).

SUPPLEMENTARY - METHODS

SUPPLEMENTARY - RESULTS

DATA AVAILABILITY STATEMENT

The python codes will be available at `https://github.com/rakro101/otter.git` after acceptance. Raw metabarcoding reads will be accessible under ENA.

REFERENCES

Arrigo, K. R., van Dijken, G. L., and Strong, A. L. (2015). Environmental controls of marine productivity hot spots around a ntarctica. *Journal of Geophysical Research: Oceans* 120, 5545–5565

Assmy, P., Fernández-Méndez, M., Duarte, P., Meyer, A., Randelhoff, A., Mundy, C. J., et al. (2017). Leads in arctic pack ice enable early phytoplankton blooms below snow-covered sea ice. *Scientific reports* 7, 40850

Barber, D. G., Hop, H., Mundy, C. J., Else, B., Dmitrenko, I. A., Tremblay, J.-E., et al. (2015). Selected physical, biological and biogeochemical implications of a rapidly changing arctic marginal ice zone. *Progress in Oceanography* 139, 122–150

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* 2008, P10008

Bluhm, B. A., Gradinger, R., and Hopcroft, R. R. (2011). Arctic ocean diversity: synthesis. *Marine Biodiversity* 41, 1–4

Bluhm, B. A., Swadling, K. M., and Gradinger, R. (2017). Sea ice as a habitat for macrograzers. *Sea ice* , 394–414

Bruhn, C. S., Lundholm, N., Hansen, P. J., Wohlrab, S., and John, U. (2021). Transition from a mixotroph/heterotroph protist community during the dark winter to a photoautotrophic spring community in arctic surface waters. *PrePrint*

Cardozo-Mino, M. G., Salter, I., Nöthig, E.-M., Metfies, K., Ramondenc, S., Wekerle, C., et al. (2023). A decade of microbial community dynamics on sinking particles during high carbon export events in the eastern fram strait. *Frontiers in Marine Science* 10, 1173384

Cleary, A. C., Durbin, E. G., Rynearson, T. A., and Bailey, J. (2016). Feeding by pseudocalanus copepods in the bering sea: trophic linkages and a potential mechanism of niche partitioning. *Deep Sea Research Part II: Topical Studies in Oceanography* 134, 181–189

Clement Kinney, J., Maslowski, W., Osinski, R., Jin, M., Frants, M., Jeffery, N., et al. (2020). Hidden production: On the importance of pelagic phytoplankton blooms beneath arctic sea ice. *Journal of Geophysical Research: Oceans* 125, e2020JC016211

Dunne, J. A., Williams, R. J., and Martinez, N. D. (2002). Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology letters* 5, 558–567

Eppley, R. W. and Peterson, B. J. (1979). Particulate organic matter flux and planktonic new production in the deep ocean. *Nature* 282, 677–680

Fernández-Méndez, M., Katlein, C., Rabe, B., Nicolaus, M., Peeken, I., Bakker, K., et al. (2015). Photosynthetic production in the central arctic ocean during the record sea-ice minimum in 2012. *Biogeosciences* 12, 3525–3549

Fortier, M., Fortier, L., Michel, C., and Legendre, L. (2002). Climatic and biological forcing of the vertical flux of biogenic particles under seasonal arctic sea ice. *Marine Ecology Progress Series* 225, 1–16

Frey, K. (2017). Arctic ocean primary productivity. *In: Arctic Report Card 2017, NOAA, http://www. arctic. noaa. gov/Report-Card/Report-Card-2017/ArtMID/7798/ArticleID/701/Arctic-Ocean-Primary-Productivity*

Fujita, H., Ushio, M., Suzuki, K., Abe, M. S., Yamamichi, M., Iwayama, K., et al. (2023). Alternative stable states, nonlinear behavior, and predictability of microbiome dynamics. *Microbiome* 11, 63

Gosselin, M., Levasseur, M., Wheeler, P. A., Horner, R. A., and Booth, B. C. (1997). New measurements of phytoplankton and ice algal production in the arctic ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* 44, 1623–1644

Guimera, R. and Nunes Amaral, L. A. (2005). Functional cartography of complex metabolic networks. *nature* 433, 895–900

Hardge, K., Peeken, I., Neuhaus, S., Lange, B. A., Stock, A., Stoeck, T., et al. (2017). The importance of sea ice for exchange of habitat-specific protist communities in the central arctic ocean. *Journal of Marine Systems* 165, 124–138

Ives, A. R. and Carpenter, S. R. (2007). Stability and diversity of ecosystems. *science* 317, 58–62

Jassey, V. E., Signarbieux, C., Hättenschwiler, S., Bragazza, L., Buttler, A., Delarue, F., et al. (2015). An unexpected role for mixotrophs in the response of peatland carbon cycling to climate warming. *Scientific reports* 5, 16931

Javier, P. J. E. A., Liponhay, M. P., Dajac, C. V. G., and Monterola, C. P. (2022). Causal network inference in a dam system and its implications on feature selection for machine learning forecasting. *Physica A: Statistical Mechanics and its Applications* 604, 127893

Joyce, K. E., Laurienti, P. J., Burdette, J. H., and Hayasaka, S. (2010). A new measure of centrality for brain networks. *PloS one* 5, e12200

Kang, S.-H. and Fryxell, G. A. (1992). Fragilariopsis cylindrus (grunow) krieger: the most abundant diatom in water column assemblages of antarctic marginal ice-edge zones. *Polar Biology* 12, 609–627

Kraft, A., Nöthig, E.-M., Bauerfeind, E., Wildish, D. J., Pohle, G. W., Bathmann, U. V., et al. (2013). First evidence of reproductive success in a southern invader indicates possible community shifts among arctic zooplankton. *Marine Ecology Progress Series* 493, 291–296

Leu, E., Søreide, J., Hessen, D., Falk-Petersen, S., and Berge, J. (2011). Consequences of changing sea-ice cover for primary and secondary producers in the european arctic shelf seas: timing, quantity, and quality. *Progress in Oceanography* 90, 18–32

Lewis, K., Van Dijken, G., and Arrigo, K. R. (2020). Changes in phytoplankton concentration now drive increased arctic ocean primary production. *Science* 369, 198–202

Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., et al. (2015). Determinants of community structure in the global plankton interactome. *Science* 348, 1262073

Lin, I., Liu, W. T., Wu, C.-C., Wong, G. T., Hu, C., Chen, Z., et al. (2003). New evidence for enhanced ocean primary production triggered by tropical cyclone. *Geophysical Research Letters* 30

Lin, Y.-C., Chin, C.-P., Yang, J. W., Chiang, K.-P., Hsieh, C.-h., Gong, G.-C., et al. (2022). How communities of marine stramenopiles varied with environmental and biological variables in the subtropical northwestern pacific ocean. *Microbial ecology* , 1–13

Loreau, M. and De Mazancourt, C. (2013). Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecology letters* 16, 106–115

Ma, B., Wang, H., Dsouza, M., Lou, J., He, Y., Dai, Z., et al. (2016). Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern china. *The ISME journal* 10, 1891–1901

Ma, B., Wang, Y., Ye, S., Liu, S., Stirling, E., Gilbert, J. A., et al. (2020). Earth microbial co-occurrence network reveals interconnection pattern across microbiomes. *Microbiome* 8, 1–12

Maat, D. S., Crawfurd, K. J., Timmermans, K. R., and Brussaard, C. P. (2014). Elevated co2 and phosphate limitation favor micromonas pusilla through stimulated growth and reduced viral impact. *Applied and Environmental Microbiology* 80, 3119–3127

Mann, D. G., Crawford, R., Round, F., et al. (2017). Bacillariophyta. *Handbook of the Protists* 7, 205–266

Meredith, M., Sommerkorn, M., Cassota, S., Derksen, C., Ekaykin, A., Hollowed, A., et al. (2019). Polar regions. *Special report on ocean and cryosphere in a changing chapter Intergovernmental Panel on Climate Change (IPCC). In Chapter 3 Polar Issues*

Metfies, K., Bauerfeind, E., Wolf, C., Sprong, P., Frickenhaus, S., Kaleschke, L., et al. (2017). Protist communities in moored long-term sediment traps (fram strait, arctic)–preservation with mercury chloride allows for pcr-based molecular genetic analyses. *Frontiers in Marine Science* 4, 301

Mitra, A., Flynn, K. J., Burkholder, J. M., Berge, T., Calbet, A., Raven, J. A., et al. (2014). The role of mixotrophic protists in the biological carbon pump. *Biogeosciences* 11, 995–1005

Nöthig, E.-M., Ramondenc, S., Haas, A., Hehemann, L., Walter, A., Bracher, A., et al. (2020). Summertime chlorophyll a and particulate organic carbon standing stocks in surface waters of the fram strait and the arctic ocean (1991–2015). *Frontiers in Marine Science* 7, 350

Oldenburg, E., Popa, O., Wietz, M., von Appen, W.-J., Torres-Valdes, S., Bienhold, C., et al. (2024). Sea-ice melt determines seasonal phytoplankton dynamics and delimits the habitat of temperate Atlantic taxa as the Arctic Ocean atlantifies. *ISME Communications* 4, ycae027. doi:10.1093/ismeco/ycae027

Priest, T., von Appen, W.-J., Oldenburg, E., Popa, O., Torres-Valdés, S., Bienhold, C., et al. (2023). Atlantic water influx and sea-ice cover drive taxonomic and functional shifts in arctic marine bacterial communities. *The ISME Journal* 17, 1612–1625

Richter-Menge, J. A. and Farrell, S. L. (2013). Arctic sea ice conditions in spring 2009–2013 prior to melt. *Geophysical Research Letters* 40, 5888–5893

Sala, O. E., Stuart Chapin, F., Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., et al. (2000). Global biodiversity scenarios for the year 2100. *science* 287, 1770–1774

Sergeeva, V., Zhitina, L., Mosharov, S., Nedospasov, A., and Polukhin, A. (2018). Phytoplankton community structure in the polar front of the eastern barents sea at the end of the growth season. *Oceanology* 58, 700–709

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 13, 2498–2504

Soltwedel, T., Bauerfeind, E., Bergmann, M., Bracher, A., Budaeva, N., Busch, K., et al. (2016). Natural variability or anthropogenically-induced variation? insights from 15 years of

multidisciplinary observations at the arctic marine lter site hausgarten. *Ecological Indicators* 65, 89–102

Soltwedel, T., Bauerfeind, E., Bergmann, M., Budaeva, N., Hoste, E., Jaeckisch, N., et al. (2005). Hausgarten: multidisciplinary investigations at a deep-sea, long-term observatory in the arctic ocean. *Oceanography*

Spreen, G., Kaleschke, L., and Heygster, G. (2008). Sea ice remote sensing using amsr-e 89-ghz channels. *Journal of Geophysical Research: Oceans* 113

Stephens, C. R., Heau, J. G., González, C., Ibarra-Cerdeña, C. N., Sánchez-Cordero, V., and González-Salazar, C. (2009). Using biotic interaction networks for prediction in biodiversity and emerging diseases. *PLoS One* 4, e5725

Strass, V. H. and Nöthig, E. M. (1996). Seasonal shifts in ice edge phytoplankton blooms in the barents sea related to the water column stability. *Polar Biology* 16, 409–422

Sugihara, G., May, R., Ye, H., Hsieh, C.-h., Deyle, E., Fogarty, M., et al. (2012). Detecting causality in complex ecosystems. *science* 338, 496–500

Suter, E. A., Pachiadaki, M., Taylor, G. T., and Edgcomb, V. P. (2022). Eukaryotic parasites are integral to a productive microbial food web in oxygen-depleted waters. *Frontiers in Microbiology* 12, 764605

Suzuki, K., Nagaoka, S., Fukuda, S., and Masuya, H. (2020). Energy landscape analysis of ecological communities elucidates the phase space of community assembly dynamics. *Ecological Monographs*

Suzuki, K., Nakaoka, S., Fukuda, S., and Masuya, H. (2021). Energy landscape analysis elucidates the multistability of ecological communities across environmental gradients. *Ecological Monographs* 91, e01469

Thaler, M. and Lovejoy, C. (2014). Environmental selection of marine stramenopile clades in the arctic ocean and coastal waters. *Polar biology* 37, 347–357

Tragin, M. and Vaulot, D. (2018). Green microalgae in marine coastal waters: The ocean sampling day (osd) dataset. *Scientific Reports* 8, 14020

Ushio, M. (2022). Interaction capacity as a potential driver of community diversity. *Proceedings of the Royal Society B* 289, 20212690

van Leeuwe, M. A., Fenton, M., Davey, E., Rintala, J.-M., Jones, E. M., Meredith, M. P., et al. (2022). On the phenology and seeding potential of sea-ice microalgal species. *Elem Sci Anth* 10, 00029

Veilleux, B. (1979). An analysis of the predatory interaction between paramecium and didinium. *The Journal of Animal Ecology* , 787–803

von Appen, W.-J., Waite, A. M., Bergmann, M., Bienhold, C., Boebel, O., Bracher, A., et al. (2021). Sea-ice derived meltwater stratification slows the biological carbon pump: results from continuous observations. *Nature Communications* 12, 7309

von Appen, W.-J., Wekerle, C., Hehemann, L., Schourup-Kristensen, V., Konrad, C., and Iversen, M. H. (2018). Observations of a submesoscale cyclonic filament in the marginal ice zone. *Geophysical Research Letters* 45, 6141–6149

Von Quillfeldt, C. (2000). Common diatom species in arctic spring blooms: their distribution and abundance. *Botanica Marina*

Wietz, M., Bienhold, C., Metfies, K., Torres-Valdés, S., von Appen, W.-J., Salter, I., et al. (2021). The polar night shift: seasonal dynamics and drivers of arctic ocean microbiomes revealed by autonomous sampling. *ISME Communications* 1, 76

Wietz, M., Engel, A., Ramondenc, S., Niwano, M., von Appen, W.-J., Priest, T., et al. (2024). The arctic summer microbiome across fram strait: Depth, longitude, and substrate concentrations structure microbial diversity in the euphotic zone. *Environmental Microbiology*

Yang, E. J., Ha, H. K., and Kang, S.-H. (2015). Microzooplankton community structure and grazing impact on major phytoplankton in the chukchi sea and the western canada basin, arctic ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* 120, 91–102

Zhao, H., Zhang, Z., Nair, S., Zhao, J., Mou, S., Xu, K., et al. (2022). Vertically exported phytoplankton (¡ 20 μm) and their correlation network with bacterioplankton along a deep-sea seamount. *Frontiers in Marine Science* 9, 862494

Zhao, Q., Van den Brink, P. J., Xu, C., Wang, S., Clark, A. T., Karakoç, C., et al. (2023). Relationships of temperature and biodiversity with stability of natural aquatic food webs. *Nature Communications* 14, 3507

Chapter 6

Zooplankton: Automation of Zooplankton classification without internet access

In this chapter, we present our use case concerning zooplankton image classification: Deep LOKI. We first briefly introduce the topic "Image Classification using Machine Learning" and subsequently present our approach to automatic pre-sorting the zooplankton image using Deep Transfer Learning (DTL) and self-super-vised learning Figure 6.1.

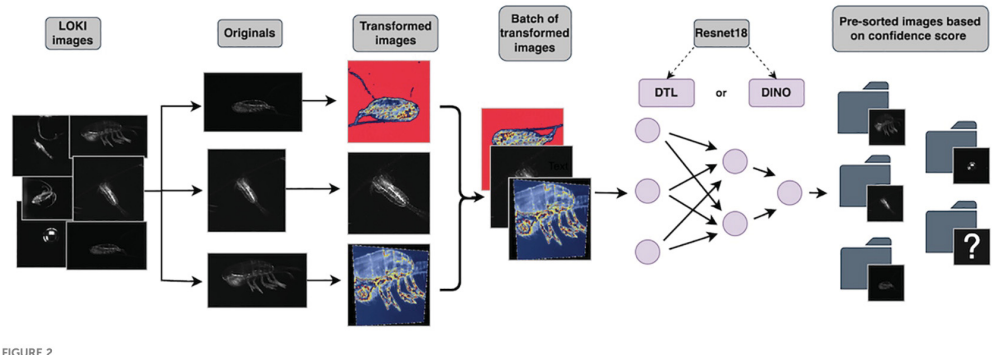

Figure 6.1: **DeepLOKI workflow in detail:** Images extracted from LOKI undergo augmentation through torchvision transform functions such as cropping, flipping, and auto-contrast. These augmented images are then inputted into one of the two variants of the ResNet18 neural network for classification (DTL, DINO). Our approach consists of two steps, first data training and classification and second data sorting in particular group after passing a confidence threshold, that we identify as classification likelihood. Images that fall below the threshold are moved to a folder labeled as unknown/unclear. The image is based on Oldenburg et al., 2023b.

6.1 DeepLOKI- A deep learning based approach to identify Zooplankton taxa on high-resolution images from the optical plankton recorder LOKI

In this section, we provide an overview of the contributions and impact of our paper (Oldenburg et al., 2023b):

Ellen Oldenburg, Raphael M. Kronberg, Oliver Ebenhöh, Barbara Nierhoff and Ovidiu Popa
"DeepLOKI- A deep learning based approach to identify Zooplankton taxa on high-resolution images from the optical plankton recorder LOKI"
In: *Frontiers in Marine Science, 2023, Volume 10*

Main Results in Simple Terms

Our research focuses on oceanic ecology, specifically on zooplankton, which are tiny organisms that play a crucial role in the marine food chain. These organisms are often distributed patchily throughout the ocean, making their study challenging using traditional methods such as net casts, which do not provide detailed information on their vertical distribution.
To address this issue, we developed a tool called DeepLOKI. This AI-based software uses machine learning, DeepLOKI can be used directly on board ships during research cruises, saving time and eliminating the need for internet connectivity.
DeepLOKI was tested on images from multiple research cruises and was found to accurately classify zooplankton taxa, including copepod developmental stages, with an average accuracy of 83.9%. This performance surpasses the commonly used method EcoTaxa by a factor of two.
Additionally, a user-friendly graphical interface for DeepLOKI was developed, streamlining the analysis process. Furthermore, DeepLOKI's capability to visualize the networks image representation aid in detecting irregularities in image parameters, thereby improving data quality control.
Our approach is flexible and can be applied to different imaging systems, making it a versatile tool for studying zooplankton in various oceanic environments. DeepLOKI provides researchers with a faster, more accurate, and more efficient way to study these vital organisms and their role in marine ecosystems.

Summary/Abstract

Zooplankton play a crucial role in the ocean's ecology, as they form a foundational component in the food chain by consuming phytoplankton or other zooplankton, supporting various marine species and influencing nutrient cycling. The vertical distribution of zooplankton in the ocean is patchy, and its relation to hydrographical conditions cannot be fully deciphered using traditional net casts due to the large depth intervals sampled. The Lightframe On-sight Keyspecies Investigation (LOKI) concentrates zooplankton with a net that leads to a flow-through chamber with a camera taking images. These high-resolution images allow for the

determination of zooplankton taxa, often even to genus or species level, and, in the case of copepods, developmental stages. Each cruise produces a substantial volume of images, ideally requiring onboard analysis, which presently consumes a significant amount of time and necessitates internet connectivity to access the EcoTaxa Web service. To enhance the analyses, we developed an AI-based software framework named DeepLOKI, utilizing Deep Transfer Learning (DTL) with a Convolution Neural Network Backbone. Our DeepLOKI can be applied directly on board. We trained and validated the model on pre-labeled images from four cruises, while images from a fifth cruise were used for testing. The best-performing model, utilizing the self-supervised pre-trained ResNet18 Backbone, achieved a notable average classification accuracy of 83.9%, surpassing the regularly and frequently used method EcoTaxa (default) in this field by a factor of two. In summary, we developed a tool for pre-sorting high-resolution black and white zooplankton images with high accuracy, which will simplify and quicken the final annotation process. In addition, we provide a user-friendly graphical interface for the DeepLOKI framework for efficient and concise processes leading up to the classification stage. Moreover, performing latent space analysis on the self-supervised pre-trained ResNet18 Backbone could prove advantageous in identifying anomalies such as deviations in image parameter settings. This, in turn, enhances the quality control of the data. Our methodology remains agnostic to the specific imaging end system used, such as LOKI, UVP, or ZooScan, as long as there is a sufficient amount of appropriately labeled data available to enable effective task performance by our algorithms.

Personal Contribution

EO devised the project, the main conceptual ideas, and the study outline. **EO** and RMK designed the model and the computational framework, analyzed the data, and wrote the initial draft. RMK carried out the implementation. **EO**, RMK, and OP interpreted the data, conceptualized, and drafted the manuscript. BN was responsible for data curation, funding acquisition, resources, validation, and contributed to writing – review and editing. OE contributed to funding acquisition and writing – review and editing. OP advised on data evaluation and data interpretation, and revised and finalized the manuscript. **All authors** contributed to improving the final manuscript.

Importance of the Research and Contribution to this Thesis

Comprehending the functioning of the delicate Arctic ecosystem requires an understanding of the structure and dynamics of its food web. Our research aims to gather comprehensive data on various components, including eukaryotes, bacteria, genomes, zooplankton, and environmental factors, to piece together a clearer picture.
Zooplankton, being a fundamental part of the food web, play a significant role in shaping Arctic marine ecosystems. Studying zooplankton distribution and abundance in the vast Arctic Ocean presents significant challenges. The research paper on DeepLOKI is essential for our research as it enables us to effectively identify zooplankton taxa from the vast hauls of images captured during research cruises.

The research paper on DeepLOKI is essential for our research as it enables us to effectively identify zooplankton taxa from the vast hauls of images captured during research cruises. Additionally, deep learning techniques can be used to classify zooplankton images, which allows for the alignment of image data with time-series data. This aligns with our research objective of integrating various data types to analyze ecosystem dynamics over time.
To summarise, our research addresses the fifth research question by implementing DeepLOKI, which provides a robust methodology for classifying zooplankton images. This classification facilitates the integration of image data into our time-series analysis, ultimately contributing to a deeper understanding of Arctic food web structure and community dynamics.

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 30 November 2023
DOI 10.3389/fmars.2023.1280510

OPEN ACCESS

EDITED BY
Carolin Müller,
Leibniz Centre for Tropical Marine Research (LG), Germany

REVIEWED BY
Tuomas Eerola,
LUT University, Finland
Katrin Linse,
British Antarctic Survey (BAS), United Kingdom

*CORRESPONDENCE
Ellen Oldenburg
Ellen.Oldenburg@hhu.de

†These authors have contributed equally to this work and share first authorship

RECEIVED 24 August 2023
ACCEPTED 06 November 2023
PUBLISHED 30 November 2023

CITATION
Oldenburg E, Kronberg RM, Niehoff B, Ebenhöh O and Popa O (2023) DeepLOKI- a deep learning based approach to identify zooplankton taxa on high-resolution images from the optical plankton recorder LOKI.
Front. Mar. Sci. 10:1280510.
doi: 10.3389/fmars.2023.1280510

DeepLOKI- a deep learning based approach to identify zooplankton taxa on high-resolution images from the optical plankton recorder LOKI

Ellen Oldenburg[1,2*†], Raphael M. Kronberg[3†], Barbara Niehoff[4], Oliver Ebenhöh[1,2] and Ovidiu Popa[1]

[1]Institute of Theoretical and Quantitative Biology, Heinrich Heine University, Düsseldorf, North Rhine-Westphalia, Germany, [2]Cluster of Excellence on Plant Sciences, Heinrich Heine University, Düsseldorf, North Rhine-Westphalia, Germany, [3]Mathematical Modelling of Biological Systems, Heinrich Heine University, Düsseldorf, North Rhine-Westphalia, Germany, [4]Polar Biological Oceanography, Alfred-Wegener- Institut Helmholtz Center for Polar and Marine Research, Bremerhaven, Germany

Zooplankton play a crucial role in the ocean's ecology, as they form a foundational component in the food chain by consuming phytoplankton or other zooplankton, supporting various marine species and influencing nutrient cycling. The vertical distribution of zooplankton in the ocean is patchy, and its relation to hydrographical conditions cannot be fully deciphered using traditional net casts due to the large depth intervals sampled. The Lightframe On-sight Keyspecies Investigation (LOKI) concentrates zooplankton with a net that leads to a flow-through chamber with a camera taking images. These high-resolution images allow for the determination of zooplankton taxa, often even to genus or species level, and, in the case of copepods, developmental stages. Each cruise produces a substantial volume of images, ideally requiring onboard analysis, which presently consumes a significant amount of time and necessitates internet connectivity to access the EcoTaxa Web service. To enhance the analyses, we developed an AI-based software framework named DeepLOKI, utilizing Deep Transfer Learning with a Convolution Neural Network Backbone. Our DeepLOKI can be applied directly on board. We trained and validated the model on pre-labeled images from four cruises, while images from a fifth cruise were used for testing. The best-performing model, utilizing the self-supervised pre-trained ResNet18 Backbone, achieved a notable average classification accuracy of 83.9%, surpassing the regularly and frequently used method EcoTaxa (default) in this field by a factor of two. In summary, we developed a tool for pre-sorting high-resolution black and white zooplankton images with high accuracy, which will simplify and quicken the final annotation process. In addition, we provide a user-friendly graphical interface for the DeepLOKI framework for efficient and concise processes leading up to the classification stage. Moreover, performing latent space analysis on the self-supervised pre-trained ResNet18 Backbone could prove advantageous in identifying anomalies such as deviations in image

10.3389/fmars.2023.1280510

parameter settings. This, in turn, enhances the quality control of the data. Our methodology remains agnostic to the specific imaging end system used, such as Loki, UVP, or ZooScan, as long as there is a sufficient amount of appropriately labeled data available to enable effective task performance by our algorithms.

KEYWORDS

computer vision, deep learning, DeepLOKI, keystone species, marine and zooplankton

1 Introduction

Imaging has become an important tool in marine zooplankton studies, in both the laboratory and the field, in the last decades (Gorsky et al., 1989; Schulz et al., 2010; Hauss et al., 2016; Kiko et al., 2020; Rubbens et al., 2023). To digitize and analyze preserved samples that have been collected by traditional net tows, the lab-based ZooScan system (Grosjean et al., 2004) has been developed. Also, several *in-situ* systems that continuously take images during deployment have been developed, allowing the study of the patchy distribution of particles and zooplankton organisms (Cowen and Guigand, 2008; Lertvilai, 2020). Among these, the underwater vision profiler (UVP, Hydroptic, France) is one of the most frequently used, being deployed worldwide (Picheral et al., 2010; Kiko et al., 2020; Picheral et al., 2022). The UVP takes images directly in the water column; therefore often the resolution is limited. However, it excels at capturing fragile organisms such as gelatinous zooplankton and particles that are typically destroyed by zooplankton nets. The LOKI (Lightframe-Onsight Key species Investigation; Isitec, Germany) system, in contrast, has been designed to collect high-resolution images by concentrating the zooplankton with a net that leads to a flow-through chamber with a digital camera (Schulz et al., 2010). LOKI captures images of genera or species and sometimes even developmental stages of rather hard-bodied organisms such as copepods, which often dominate zooplankton communities (Hirche et al., 2014; Orenstein et al., 2022). The device captures images continuously during vertical drops and records hydrographic parameters, including salinity, temperature, oxygen concentration and fluorescence. This enables a thorough analysis of the species distribution. For general *in-situ* zooplankton images, various approaches for digital classification have been employed (Rubbens et al., 2023). Initially, ensemble models were used (Schmid et al., 2016). With advancements in computing power, Deep Learning approaches emerged (LeCun et al., 2015), leveraging convolutional neural networks to process entire images and extract intricate patterns (Luo et al., 2018). Additionally, Transfer Learning, which involves the transfer of knowledge from large datasets to smaller ones (Yosinski et al., 2014), has been employed in Deep Learning algorithms to enhance their performance (Orenstein and Beijbom, 2017; Orenstein et al., 2022). For UVP and ZooScan images, the analysis software ZooProcess, a macro of ImageJ, and the web-based annotation tool EcoTaxa (Picheral et al., 2017) have been developed. In addition, there are numerous other methods (Bi et al., 2015; Bi et al., 2022; Yue et al., 2023). ZooProcess extracts numerical parameters from each image and automatically measures the size of each object (Grosjean et al., 2004; Picheral et al., 2017). EcoTaxa is then used to annotate the images, i.e., to sort the objects on the images into categories and label each image accordingly. EcoTaxa also allows to manually drag and drop images into the respective category; however, the application also provides automated annotation functionality through a Random Forest algorithm to predict the categories based on numerical image parameters. The algorithm must be trained with annotated images, and the better the training set, the more accurate the prediction is. For example, depending on the taxonomic resolution, over 80% of zooplankton images from the Fram Strait that were taken with ZooScan were correctly annotated (B. Niehoff, pers. obs.).

For the analysis of LOKI images, a software tool - the LOKI browser - was developed and provided together with LOKI hardware (Schulz et al., 2010). Similar to ZooProcess, this application generates numeric image descriptors and, similar to EcoTaxa, allows to sort the objects into categories. Unfortunately, the LOKI browser is outdated and lacks user-friendliness. For instance, the process of uploading more than 150,000 images from a single cruise requires manual handling of small batches of 2,000-5,000 images. Furthermore, the annotation procedure is inconvenient as it does not provide direct access to specific categories but requires traversing the entire taxonomic tree, which results in multiple clicks per image, especially in the case of species categorization. It also has to be noted that working on the EcoTaxa server requires a stable internet connection, which is not always a given during cruises to remote areas such as the Arctic. Therefore, the image data can typically only be processed after the cruise. In summary, the current workflow for LOKI images faces several issues: the time-consuming and upload-limited image pre-processing, the low prediction accuracy, and the dependence on internet access. Addressing these issues, we present an alternative workflow. (1) We developed two deep learning methods using the images as input instead of image descriptors and thus omitted the tedious upload to the LOKI browser, saving time and personnel. Aiming at a better prediction of the categories, we implemented a deep transfer learning (DTL) and a two-step self-supervised learning approach based on first pre-train self-supervised and then fine-tuning supervised, called self-distillation with no labels (DINO), which currently is one of the state of the art methods

10.3389/fmars.2023.1280510

(Chen et al., 2020; He et al., 2020; Caron et al., 2021). To ensure that only images with a high level of confidence in their label assignments are sorted into their respective categories, we implemented a confidence threshold, which is only used within the GUI as a parameter and not for training or evaluation. A similar approach for threshold implementation on CNNs was previously shown by (Kraft et al., 2022). By implementing this threshold, we can effectively exclude images for which the algorithm exhibits lower confidence in label assignments, thereby enabling a more precise categorization relying on confident predictions (on inference). As provided in EcoTaxa, we've included categories for artifacts such as bubbles, detritus, and unknown objects to allow for the categorization of these as well. (2) To allow for immediate analyses onboard, we adopted a small backbone model that is suitable for deployment on mobile laptops for both our methods, eliminating the need for constant internet connectivity. The objectives of this study thus were to enhance the classification accuracy in comparison to the existing workflow for zooplankton image analyses and to provide a user-friendly approach that can be readily applied on board. We utilized a dataset of 215,000 images from five cruises, which had been fully annotated at Alfred-Wegener-Institut (AWI) in advance, to perform a comprehensive evaluation of a reference dataset to assess and compare the efficiency of three distinct methods: the EcoTaxa workflow, as well as two innovative deep learning approaches, i.e., DTL and DINO.

2 Methods

2.1 Data

Please note: In this paper, we use *category* for the technical term *class* to avoid confusion with the biological term (taxa) classes. The technical term *class*, which we call category, refers to a collection of pictures containing similar motifs that are building a group after the classification procedure. In contrast, the taxonomical term class (for example, Ostracoda or Copepoda) refers to the rank of the organisms in an ancestral or hereditary hierarchy.

For the present study, we used images that were taken with the optical plankton recorder Lightframe On-Sight Keyspecies Investigation (LOKI) in Fram Strait during five expeditions of RV Polarstern (Table 1). Each of the 33 categories used in this study corresponds to a group of zooplankton fauna at several taxonomic levels (Supplementary Material DeepLOKI Section 2). LOKI consists of a net ($150\mu m$ mesh size) that concentrates the plankton during a vertical tow from a maximum of $1000m$ depth to the surface. The net leads to a flow-through chamber with a 6.1 MP camera (Prosilica GT 2750 with Sony ICX694 runs 19.8 frames per second at 6.1 MP resolution.) that takes images at a frame rate of max. $20sec^{-1}$. At the same time, sensors record depth, temperature, salinity, fluorescence, and oxygen concentration. The LOKI underwater computer extracts objects and stores the resulting images as well as sensor data on a hard drive (Schulz et al., 2010).

2.2 Workflow

Once the LOKI images are captured, they are stored on a hard drive. Depending on the classification tool being used, there are two distinct data pre-processing pipelines Figure 1. The current approach involves preparing the data for classification using EcoTaxa by importing the images into a specialized "LOKI browser" software, which calculates numeric features (Schulz et al., 2010). However, due to computational limitations and the need for internet access, the subsequent steps must be carried out after the cruise. This involves applying ZOOMIE software (Schmid et al., 2015) to exclude multiple images of the same organism, followed by loading the data into EcoTaxa for classification and storage Figure 1. In contrast, our proposed pipeline involves only one step. The DeepLOKI tool can directly classify raw images without any preparation steps like, for example, feature extraction by the LOKI browser. As a final stage at the end of both pipelines, a human carries out a final assessment of the pre-sorted images. During this process, any necessary label corrections and duplicate removals are made Figure 1. We display the current and the proposed workflow (Figures 1, 2).

2.3 Data acquisition and ground truth

For the classification procedure, we used the images from all five cruises after applying ZOOMIE to reduce the number of duplicates.

To develop the new pipeline, we used a data set of 194,479 (train/val) images (Supplementary Material DeepLOKI Table S3) that, in addition, underwent parameterization using the LOKI

TABLE 1 Overview of LOKI images from five RV Polarstern cruises.

cruise	#n images	#n categories	collection	usage	study area	#LOKI deploys*
PS99.2	20683	31	June/July 2016	test	Fram Strait	4
PS106.2	42462	33	July 2017	train/val	Fram Strait	20
PS107	121628	32	July/Aug. 2017	train/val	Fram Strait	17
PS114	7199	25	July 2018	train/val	Fram Strait	1
PS122	23190	30	Nov.2019-Sept.2020	train/val	Fram Strait	19

The table presents the number of images used for this study (#n), the number of categories into which the images were sorted, the months of image collection (obtained from object cruise), and the usage for either training and validation (train/val) or testing the application. The study area is always Fram strait and the number of LOKI deploys* is based on the meta data file (technical counting).

10.3389/fmars.2023.1280510

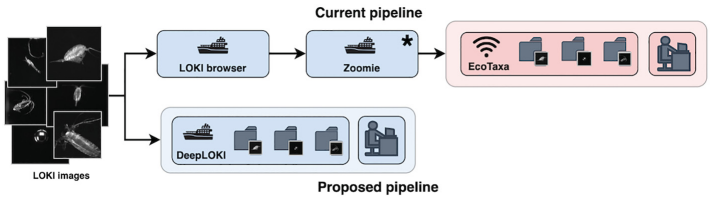

FIGURE 1
Overview of the current workflow (top) and the DeepLOKI pipeline (bottom) to categorize objects on LOKI images. Current pipeline: images are loaded into the LOKI Browser, then processed with ZOOMIE to remove double takes, and loaded to the EcoTaxa website for automatic classification based on numeric image parameters fed to random-forest algorithms, and finally for a quality check by a scientist. The DeepLOKI pipeline consists of the automatic classification of images based on one of the two variants of the ResNet18 neural network (DTL or DINO), and a final check by a scientist as well. Blue boxes: step can be performed on the ship; red box: step requires internet access. Optional steps are indicated by *.

browser. All objects on these LOKI images were previously identified by zooplankton scientists, often through manual efforts in order to reach the lowest possible taxonomic level. In our study, however, we aimed at a less detailed distinction of taxa, and thus, we combined categories to higher taxonomic levels

(see (Supplementary Material DeepLOKI Table S3), Figure 3). For instance, we grouped the several morphologically very similar developmental stages of specific copepod species, as determined by marine biologists, into a single species category. We, however, solely utilized the images themselves and not their associated parameters. Additionally, we analyzed a dataset for double takes using ZOOMIE and re-categorized the images with EcoTaxa, including both automatic predictions and final evaluations by a scientist. To ensure accurate model development, 80% of these images were randomly chosen for training purposes. The remaining 20% of the data was reserved exclusively for internal model validation, ensuring its reliability. To evaluate the model's effectiveness, images from the fifth cruise (Supplementary Material DeepLOKI Table S3) that had not been used for pipeline development were used, providing an objective measure of its performance. The dataset encompassed various quantities of images within its categories, with the smallest category, Foraminifera, comprising n = 121 images, while the largest category, (Copepoda_Calanus), contained n =43,620 images (Supplementary Material DeepLOKI Table S3).

2.4 The baseline: EcoTaxa

The classification pipeline via the LOKI browser and EcoTaxa was used as a baseline to compare the performance of DeepLOKI. To ensure comparability between results obtained by our approach and those via the EcoTaxa workflow, we extracted the categories that have been distinguished by the scientists in EcoTaxa but grouped, for example, development stages of species at a higher taxonomic level. The EcoTaxa classifier was trained using up to 5,000 images per category as a maximum. The default for EcoTaxa is a Random Forest classifier, and although it is possible to upload

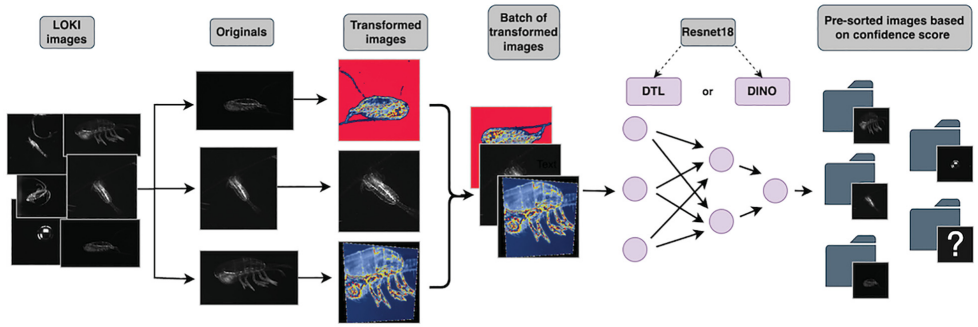

FIGURE 2
DeepLOKI workflow in detail: Images extracted from LOKI undergo augmentation through torchvision transform functions such as cropping, flipping, and auto-contrast. These augmented images are then inputted into one of the two variants of the ResNet18 neural network for classification (DTL, DINO). Our approach consists of two steps, first data training and classification and second data sorting in particular group after passing a confidence threshold, that we identify as classification likelihood. Images that fall below the threshold are moved to a folder labeled as unknown/unclear.

10.3389/fmars.2023.1280510

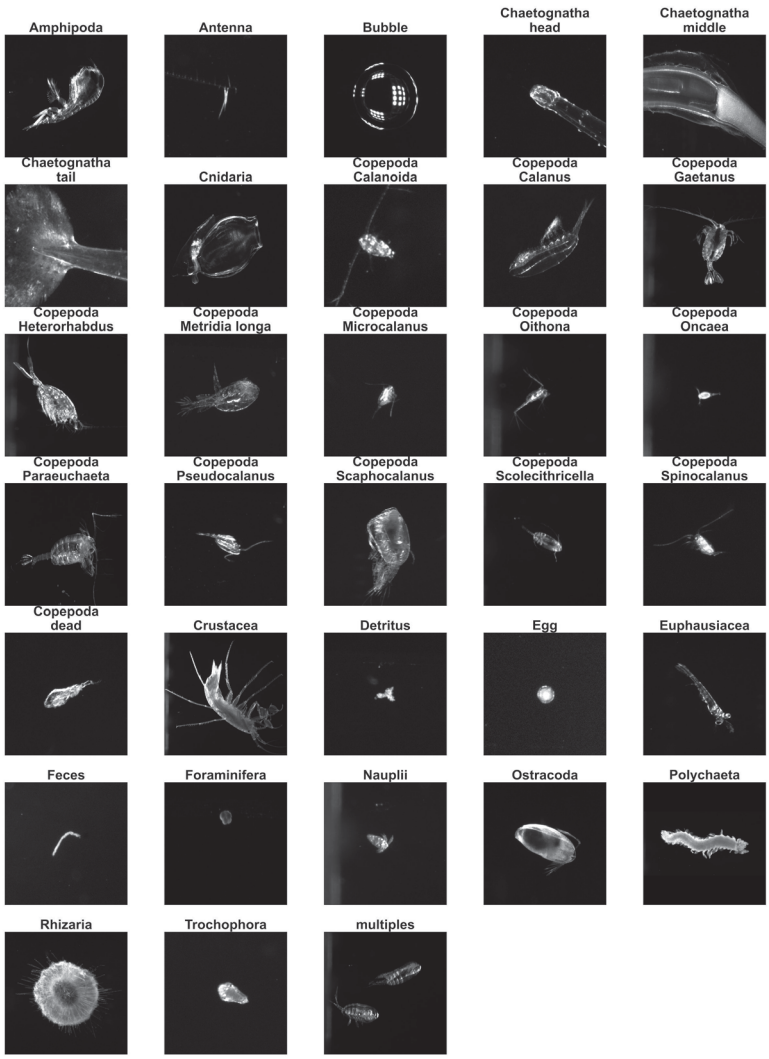

FIGURE 3
Examples for the 33 categories that were utilized in this study. Each image corresponds to zooplankton fauna at different taxonomic levels or other objects (antennae, bubbles, feces, multiples, detritus), arranged alphabetically according to their name.

other model implementations (https://github.com/ecotaxa/ecotaxa_ML_template/blob/main/4.train_classifier.py), the images have only been processed with the standard algorithm. The Random Forrest performance is constrained due to imposed parameters (optimized for UPV or ZooScan images) or training data restrictions due to EcoTaxa. To evaluate the performance of our classification tool, DeepLOKI, we first evaluated the accuracy of EcoTaxa's current classifier. Therefore, we established a new project on the EcoTaxa platform and utilized its integrated training and classification services, using 122,693 image training examples. The exact configuration can be found in the Supplementary Material EcoTaxa Setup.

2.5 Metrics

We've compiled the following metrics into a detailed table (Supplementary Material DeepLOKI Table S4) that displays

Precision, Recall, F1-Score, and sample size for every class. The last three rows of the table indicate accuracy, followed by the columns macro and weighted averages. Additionally, we've created a visual depiction of the confusion matrix by plotting the human-labeled ground truth against the predictions generated by our algorithm. To evaluate and compare our models, we used the following five metrics, here defined for the binary case. Consider a scenario in which the dataset consists of samples that fall into one of two distinct categories. Each sample can be assigned to either to these categories, resulting in a binary classification problem. The scores of the metrics are in an interval from 0 to 1, with higher scores indicating better performance.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1} - \text{score} = \frac{2\text{TP}}{2\text{TP} + \text{FN} + \text{FP}}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}}$$

In the context of multicategory classification, where the dataset contains more than two possible categories (technical term: classes) for each sample, we employed the One-vs-All approach. This approach involved designating one class as the Positive (P) category, which served as the target category for calculating specific metrics such as Precision. Conversely, all other categories were treated as Negative (N) categories.

By adopting this approach, we were able to compute separate metrics for True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN), allowing us to assess the performance of the model for each category individually.

2.6 Deep learning

For our Deep Learning approach, we used two variants of the ResNet18 neural network and compared them. The first approach involved Deep Transfer Learning, where we used a pre-trained ResNet18 model on ImageNet (Tan et al., 2018; Kronberg, 2022; Kronberg et al., 2022). This approach was selected due to its well-documented effectiveness in diverse areas of biology and medicine, with a specific emphasis on image analysis. Another factor that influenced our decision was the efficiency of fine-tuning, allowing us to train the model on hardware configurations, such as a Macbook, without requiring extensive computational resources. (Using Visual Transformers instead of simple ResNet18 could boost the accuracy by a bit, as tested in CNNs vs ViTs by the authors for (Yue et al., 2023) For our second method, we chose the DINO approach (Caron et al., 2021) primarily because it integrates one of the most advanced self-supervised pre-training techniques available in the field of computer vision. To ensure compatibility with our GPU infrastructure, we made slight adjustments to the DINO method.

2.6.1 Image augmentation

Here, we describe the pre-processing steps applied to the images extracted from the LOKI dataset for the Training. The primary image augmentation techniques utilized are as follows:

1. **Random Resized Crop**: Images are randomly cropped and resized to a fixed size of 300 pixels while maintaining an aspect ratio within the range of 0.8 to 1.0. This resizing process ensures that the model receives input images of varying scales, improving its robustness to different object sizes.
2. **Random Rotation**: We apply random rotations to the images, introducing variability in the orientation of objects. This augmentation technique helps the model learn to recognize objects from various angles and perspectives, with degrees of rotation up to 15 degrees.
3. **Random Horizontal Flip**: Images are subjected to random horizontal flips. This operation allows the model to learn features that may appear differently when mirrored horizontally, aiding in better generalization.
4. **Center Crop**: After the aforementioned augmentations, we perform a center crop on the images, resulting in a final image size of 224x224 pixels. This cropping operation ensures that the model focuses on the central region of the image, which often contains the most relevant information.
5. **Normalization**: Normalization is applied to the pixel values of the images. We subtract the mean values [0.485, 0.456, 0.406] and divide by the standard deviations [0.229, 0.224, 0.225] for each color channel. This step helps standardize the input data, making it suitable for neural network training.
6. **Random Autocontrast**: Autocontrast is applied randomly with a probability of 25%. This technique enhances image contrast, which can be beneficial for improving the model's ability to distinguish between objects with subtle variations in lighting and contrast.
7. **Random Perspective**: Images undergo random perspective transformations with a distortion scale of 0.25 and a probability of 25%. This augmentation introduces geometric distortions, simulating variations that may occur in real-world scenarios.
8. **Random Adjust Sharpness**: Random sharpness adjustments are applied with a sharpness factor of 4 and a probability of 25%. This operation can help the model focus on fine details and edges within the images.

These augmentation techniques collectively contribute to a more diverse and informative dataset, enabling our model to better generalize and recognize objects under various conditions and orientations.

2.6.2 Deep transfer learning

To implement our DTL approach architecture (Tan et al., 2018; Kronberg, 2022; Kronberg et al., 2022) (Supplementary Material

10.3389/fmars.2023.1280510

DeepLOKI Section 4), we used a pre-processing pipeline that involved resizing the images (Height, Width, Color channel) = (300, 300, 3) and crop to the input size of (224, 224, 3), as well as normalization, and various augmentations to add robustness to the training. We fine-tuned the ResNet18 neural network (He et al., 2016) as previously described (Werner et al., 2021) and adapted it for our purposes. Specifically, we exchanged the fully connected layer(matching to our number of classes) of the ResNet18 model (Figure 4) and fine-tuned the full network. Adam (Kingma and Ba, 2014) (Supplementary Material DeepLOKI Def. Adam) was used as the optimizer for this deep transfer learning approach. The network was trained with a batch size of 1536 and trained for 3 out of 20 epochs due to early stopping on the images of 80% of the samples from the dataset using the biologists' labeled images as ground truth. The predicted probability for each image to contain each of the labels of our 33 classes was used as the objective/loss function (Supplementary Material DeepLOKI Def. Cross-entropy loss) in training. We used an initial learning rate of 0.0001.

To assess the performance of our trained model, we conducted an evaluation on the remaining 20% of the dataset, which was not encountered by the algorithm during the training phase. By comparing the results obtained from the model with the ground truth, we were able to gauge its effectiveness. To ensure fair and unbiased comparison among different algorithmic approaches, we incorporated a reference dataset, specifically the Cruise PS99.2, throughout the study. This reference dataset served as a standardized benchmark for evaluating the performance of our model alongside other approaches.

2.6.3 Self-supervised pre-training and then supervised fine-tuning (DINO)

For the self-supervised pre-training stage, we utilized all available images in our dataset, disregarding any label information (Noroozi et al., 2018). The purpose was to train the model to learn a latent space representation, as described in (Caron et al., 2021). To maintain consistency with our DTL approach, to be able to compare the ImageNet pre-trained ResNet18 with the pre-training using self-super-vised learning method, we decided to employ the same ResNet18 architecture as the backbone for the self-supervised pre-training. In the original paper even a ResNet50 was used. To reduce training and inference computing costs, we decided to downscale to ResNet18. For the DINO approach, we used a 450 epoch for pre-training and trained on 8 x A100 GPUs with 64 workers and a batch size of 512. The full parameter setup can be found in our GitHub Repo using the lightly Python package. Consequently, during the fine-tuning stage, we added the fully connected layers accordingly to our classification problem and then fine-tuned all layers using our training dataset Figure 4. This fine-tuning process was performed for and trained for 12 out of 20 epoch due to early stopping (on the validation accuracy with patience of 2) while keeping all parameters identical to the previous approach.

2.6.3.1 Visualization of the latent space after self-supervised pre-training

The purpose of this approach is to learn a condensed representation of the classes in a lower-dimensional vector space. Put simply; it aims to create a representation where images of the same class are closer to each other while images of different classes are farther apart. In general, there are two commonly used methods for visualizing classification results based on a latent space: UMAP and t-SNE. UMAP tends to preserve more of the overall structure of the data, while t-SNE focuses on highlighting the local structure (van der Maaten and Hinton, 2008; McInnes et al., 2018). We employed UMAP (Uniform Manifold Approximation and Projection) as another dimensionality reduction technique to visualize the results from the classification algorithms. UMAP provides an alternative perspective on the relationships among the data points in a lower-dimensional space.

2.7 Graphic User Interface

A web-based Graphical User Interface (GUI) has been developed to streamline the use of DeepLOKI, our powerful deep

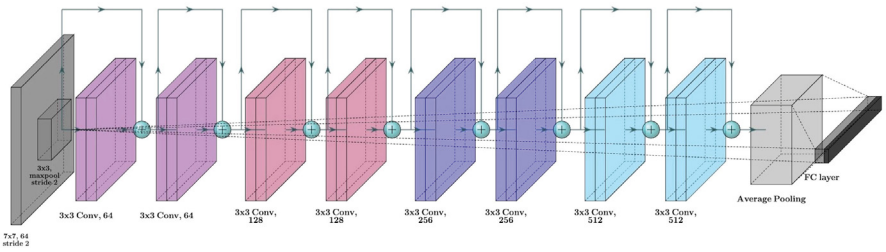

FIGURE 4
Architecture of the implemented neuronal network. Convolutional layers (Conv.): These layers use filters (k x k, Conv. f where k is the kernel size and f the number filters) to extract features from the input image. ResNet18 has a total of 17 convolutional layers. Batch normalization layers: These layers normalize the output of the convolutional layers, which helps to improve the stability and performance of the network. Pooling layers (Max/Average Pooling): These layers reduce the spatial dimensions of the feature maps generated by the convolutional layers, which helps to decrease the number of parameters in the network and reduce overfitting. Fully connected layers (FC layers): These layers connect all the neurons in one layer to all the neurons in the next layer. In ResNet18, there is one fully connected layer at the end of the network, which is used for classification (He et al., 2016).

learning framework for taxa group recognition in zooplankton images. Through the incorporation of the Streamlit Python library, a seamless and intuitive interface has been created, thereby simplifying the process of image classification. Supplementary Tools have also been integrated, offering a user-friendly interface for labeling new images, thus enabling further training (Supplementary Material DeepLOKI) Figure 1.

2.8 Hard and software

Training and Validation were performed on an Nvidia A100 (Nvidia Corp., Santa Clara, CA, USA) and on Apple M1 MAX with 32 GB (Apple, USA), depending on the computational power needed, for example, self-supervised pre-training was performed on a Hyper performing cluster with Nvidia A100. On the Macbook Pro (Apple, USA) we used:

Python VERSION:3.10.5
pyTorch VERSION:13.1.3

On the cluster we used cluster specifics versions of the software:

Python VERSION:3.10.5
pyTorch VERSION:13.1.3
CUDNN VERSION:1107)

3 Results

This section describes and compares the results of the image analyses based on three methods: EcoTaxa (web link required), ResNet18 with DTL (autonomous), and ResNet18 with DINO (autonomous). For the comparisons, we use the PS99.2 data set (20,683 images). AS evaluation metrics, we deliberately decided to use the F1-score and, in addition, the classification accuracy to highlight the different aspects of the "quality of label assignment".

3.1 EcoTaxa default classifier performance - study baseline

The automatic EcoTaxa categorization process generated a dataset where each data point (image) was assigned its respective categorization. These categories include non-living objects (detritus, bubbles, feces), high-level taxonomic groups (Crustacea (not further identified), Amphipoda, Calanoida (not further identified) Euphausiacea, and Ostracoda; Polychaeta; Cnidaria, Rhizaria (not further identified) and Foraminifera), copepod genera (Calanus, Geatanus, Heterorhabudus, Metridia, Microcalanus, Oithona, Oncaea, Paraeuchaeta, Pseudocalanus, Scaphocalanus, Scolethricella, Spinocalanus), dead copepods, early life stages (eggs, nauplii, and trochophora larvae), specific parts of organisms (antennae as well as heads, middle parts and tails of chaetognaths) and "multiples" with images with more than one object. We compared these labels to the scientists' final annotations and produced a confusion matrix and classification report (Figure 5). The EcoTaxa algorithm achieved an accuracy of 44.4% applied to a PS99.2 dataset of 20,683 images. Examining the F1-scores, which consider both precision and recall, we found an overall score of 44% for all categories Supplementary Material DeepLOKI Table S5). Among the categories, the highest-performing categories were Detritus, Chaetognata middle, and Bubble, with a score of 70 -73% (3,363, 184 and 209 images). Categories such as Copepoda_Heterorhabdus, Copepoda_Scolecithricella, Copepoda_Spinocalanus, Copepoda_dead, Crustacea, Euphausiacea, Foraminifera, Trochophora had the lowest scores of 0%, however, these categories contained few images (i.e. 148, 22, 10, 0,28, 4, 5,0 images, respectively) (Figure 5, Supplementary Material DeepLOKI Table S5). It should be noted that zero image counts correspond to misclassified images when the category was present in the training set but absent in the test set. In summary, 17% of the total images, representing only 2 out of 33 categories, attained an F1-score of 70% or higher.

3.2 DeepLOKI: ResNet18 - DTL classifier

As part of our DeepLOKI framework, we initially employed a Transfer Learning approach, utilizing a fine-tuned ResNet18 model pre-trained on the ImageNet dataset (Yang et al., 2020). The ImageNet dataset consists of 1000 categories, encompassing various objects such as animals, cars, and airplanes. By employing this technique, we achieved an overall accuracy of 83.1% on our test dataset. Using a zooplankton test dataset PS99.2 containing 20,683 images, the ResNet18 model achieved a weighted average F1-score of 82.4% for all categories (Supplementary Material DeepLOKI Table S6). Notably, the categories Ostracoda, Bubble, and Copepoda_Calanus exhibited the highest F1- scores, reaching 92-95% (748, 209, and 3,614 images, respectively). Conversely, the categories Copepoda_Scolecithricella Copepoda_Spinocalanus, Copepoda_dead, Crustacea, Euphausiacea, Foraminifera, Trochophora again had the lowest scores of 0%. However, these categories contained few images (Figure 6, as summarized in Supplementary Material DeepLOKI Table S6).

In summary, 13 out of 33 categories, comprising 86% of all images, achieved an F1-score of 70% or higher. As with the EcoTaxa results, we generated a confusion matrix from the categorization outcomes. Our findings revealed that images belonging to the groups prefixed with Copepoda exhibited a high level of confusion. This confusion was evident from the higher values observed in the non-diagonal elements of the matrix. Specifically, we observed a distinct pattern of confusion within a block encompassing various species and sub-species of Copepoda, as indicated by the framed region, notably in the category Copepoda_Calanoida: here, 318 images were misclassified as Copepoda_Mircocalanus, 95 as Copepoda_Metridia_longa, 104 as Copepoda_Scapohocalanus and some smaller number of images to other that have in their assigned category name Copepoda Figure 6.

10.3389/fmars.2023.1280510

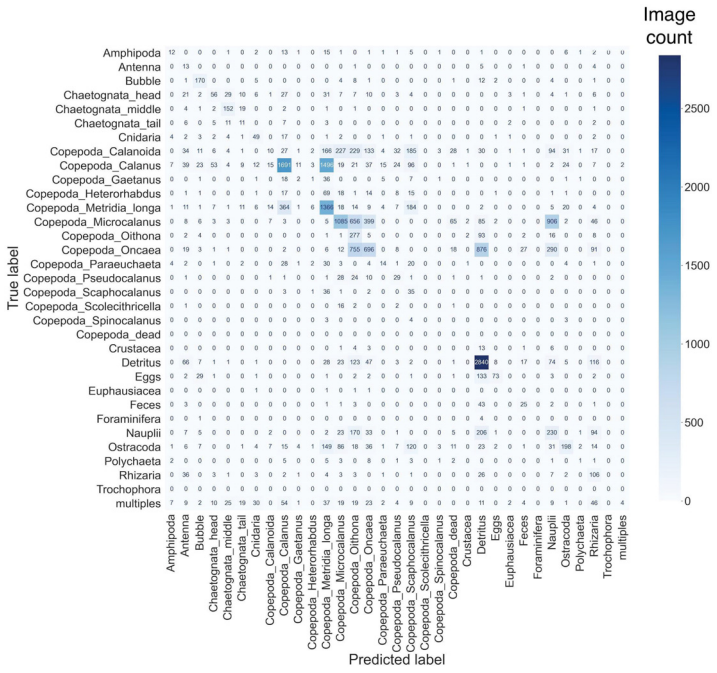

FIGURE 5
Confusion matrix of categories predicted by EcoTaxa versus final categorization by a scientist: The y-axis shows the evaluated labels and the x-axis the predicted labels of PS99.2. The color gradient indicates the number of images.

3.3 DeepLOKI: ResNet18 - DINO classifier

Using state-of-the-art methods involving self-supervised learning, specifically a slightly down-scaled student-teacher approach (Caron et al., 2021), followed by fine-tuning on our labeled data, we achieved an overall accuracy of 83.9% on the PS 99.2 dataset of 20,683 images. This performance surpassed both the EcoTaxa workflow and the DTL approach. Notably, the categories Ostracoda, Bubble, and Copepoda_Calanus demonstrated the highest accuracy rates (measured as F1-score), ranging from 92% to 95% (748, 209, and 3,614 images, respectively Supplementary Material DeepLOKI Table S7). Conversely, the poorest-performing categories were Copepoda_Scolecithricella Copepoda_Spinocalanus, Copepoda_dead, Crustacea, Euphausiacea, Foraminifera, Trochophora each with a 0% F1-score, (Figure 7, data summarized in Supplementary Material DeepLOKI Table S7). In summary, 12 out of 33 categories, covering 86.5% of all images, achieved an F1-score of 70% or higher.

The confusion matrix analysis for this categorization of zooplankton images revealed a consistent pattern. Similar to DTL analysis, images from the Copepoda groups exhibited higher non-diagonal values, indicating confusion.

This pattern was particularly noticeable within a framed block representing various Copepoda species and sub-species. Remarkably similar to the DTL approach is the high confusion in the category Copepoda Calanoida: here, 243 images are misclassified as Copepoda_Mircocalanus, 69 as Copepoda_Metridia longa, 54 as Copepoda_Scaphocalanus and similar patterns for the smaller categories including Copepoda in their names Figure 7.

3.3.1 Data projection on latent space

To visualize the classification of the self-supervised pre-trained ResNet18 (DINO) approach, we used UMAP (Figure 8) as a dimension reduction approach to project the data on a 2D space. Each dot represents an image, and the distance to each other reflects their similarity. Dots that are close to each other reflect the same or similar image content. The coloring was done based on the pre-defined categories.

These are Copepoda_Microcalanus, Copepoda_Calanus, Detritus, Copepoda_Oncaea, Copepoda_Calanoida, Copepoda_Metridia_longa, multiples, Ostracoda, Nauplii and Copepoda_Oithona. The UMAP plot has revealed the presence of a mirror symmetry at the x = 5 coordinate, effectively dividing the plot into two distinct regions. Notably, comparable clusters can be observed on both the left and right sides of this axis. This finding suggests a symmetrical organization of data points with similar characteristics in both regions. The two taxa Copepoda_Calanus and Copepoda_Metridia_longa exhibited close

10.3389/fmars.2023.1280510

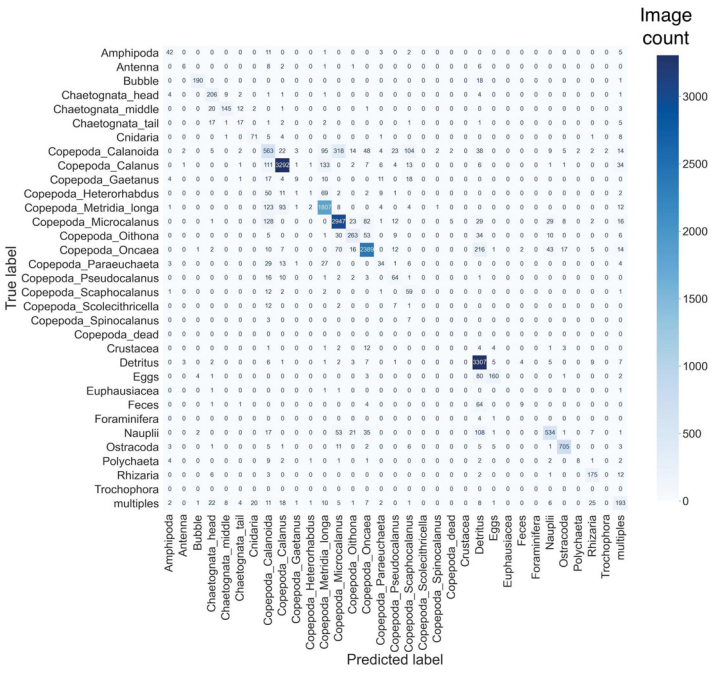

FIGURE 6
Confusion matrix of categories predicted by DTL ResNet18 -Classifier versus final categorization by a scientist: The y-axis presents the evaluated labels and the x-axis the predicted labels of PS99.2. The color gradient indicates the number of images.

proximity to each other in the two-dimensional latent space plot, mirroring their similar morphology. It is worth mentioning that the groups Copepoda_Oncaea, Copepoda_Microcalanus, Nauplii, and Detritus exhibited a tendency to cluster closely together within a specific region of the plot. The categories Copepoda_Calanoida and Ostracoda show a similar representation in the latent space.

Upon conducting a deeper analysis of the metadata, particularly focusing on the stations, we discovered that the initial group of clusters, the left clusters (comprising all clusters with x umap ¡ 5), was linked to three specific locations. Interestingly, these three locations were distinct from the other locations in the dataset. Notably, the stations were not part of the pre-training or training processes of the DINO ResNet18 model, yet their consideration offered meaningful insights into the separation of these clusters based on distinct geographical associations.

3.4 Comparison of the default EcoTaxa classifier, with the two DeepLOKI classifiers: DTL ResNet18 and Dino ResNet18

Our proposed DINO classifier demonstrated an accuracy of 83.9% in predicting categories for objects on LOKI images.

Furthermore, our classifier reached higher F1-Scores for 26 out of 33 categories, which accounted for 93,9% of all samples. These categories excluded Copepoda_Scolecithricella, Copepoda_Spinocalanus, Copepoda_dead, Crustacea, Euphausiacea, Foraminifera, Trochophora and Feces, where only the category Feces was Eco Taxa more accurate with a delta of 0.16. For the remaining categories both classifiers reached 0% F1-Score (Supplementary Material DeepLOKI Table S4, Figure 9).

Overall, the newly developed DeepLOKI classifier proved to be superior for LOKI images when compared to the standard random forest classifier employed in EcoTaxa that had been designed for UVP and ZooScan images Figure 9.

3.5 The GUI interface

We have developed a graphical user interface (GUI) for pre-sorting purposes and for manual labeling, which includes a user-friendly image viewer. The GUI was built using Streamlit, a framework that enables easy deployment of web applications. One of the key advantages of our GUI is that it can be accessed directly through a web browser (platform independent), without the need for an internet connection (Figure 10, Supplementary Material DeepLOKI Figure S1).

10.3389/fmars.2023.1280510

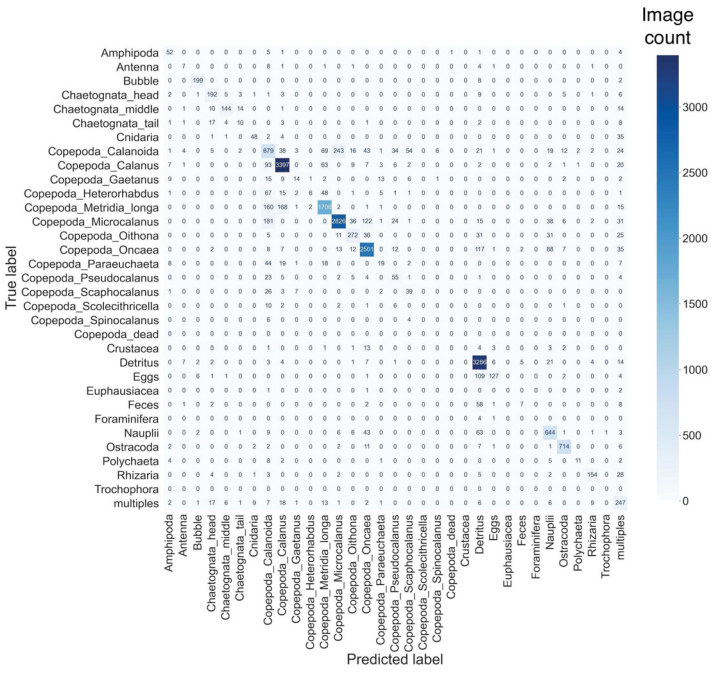

FIGURE 7
Confusion matrix of categories predicted by Dino DTL ResNet18 -Classifier versus final categorization by a scientist: The y-axis presents the evaluated labels and the x-axis the predicted labels of PS99.2. The color gradient indicates the number of images.

3.6 Time consumption and benchmark

The pre-training with 450 epochs took less than two days on the 8 - A100 GPU DGX system. The inference of the Haul containing about 1,400 images only took 14.8202 seconds.

4 Discussion

This study introduces the framework DeepLOKI, designed for categorizing objects on high-resolution black-and-white images obtained from the LOKI system (Schulz et al., 2010). Our approach incorporates advanced deep learning techniques, specifically convolutional neural networks (CNNs), utilizing both deep transfer learning and self-supervised learning methods. In terms of accuracy and F1-score, DeepLOKI outperforms the currently used web-based tool EcoTaxa that has been developed for images collected by other camera systems, specifically the UVP and the ZooScan (Grosjean et al., 2004; Picheral et al., 2017). To improve the usability of DeepLOKI, we have developed a user-friendly graphical user interface (GUI) that simplifies interaction with the framework. This GUI streamlines the classifying of zooplankton images, making this process easy to use. An advantage of DeepLOKI is that the application can be executed on a portable computer, making it suitable for implementation during cruises as it does not rely on internet access, which is often limited in remote regions such as the Arctic. Our DeepLOKI pipeline thus allows for fast image classification immediately after a cast and prompt evaluation of the zooplankton community in almost real-time. In this study, we utilized a training set of approximately 194,479 images collected during four cruises. We achieved a classification accuracy of 83.9% when classifying images from an independent fifth cruise, including 20,683 images. Both our proposed approaches are considered to be supervised, because the fine-tuning is based on labeled data. This demonstrates the effectiveness and robustness of DeepLOKI in accurately categorizing zooplankton images from the same area, i.e., Fram Strait. This approach significantly improves the efficiency of the research process and empowers researchers to make more informed decisions on sampling during a cruise.

Deep transfer learning has been employed to classify plankton data successfully(Orenstein and Beijbom, 2017; Cheng et al., 2019; Lumini and Nanni, 2019). Utilizing openly accessible datasets of categorized plankton images can facilitate the development of such designs (Sosik et al., 2015; Elineau et al., 2018; González et al., 2019; Cornils et al., 2022). Instead of just relying on transfer learning, the focus has now shifted to a two-step process called self-supervised learning. In this approach, a backbone is first pre-trained using self-supervised data to obtain a latent space representation of the images. Once the backbone has been pre-trained and its weights learned, it can then be fine-tuned

10.3389/fmars.2023.1280510

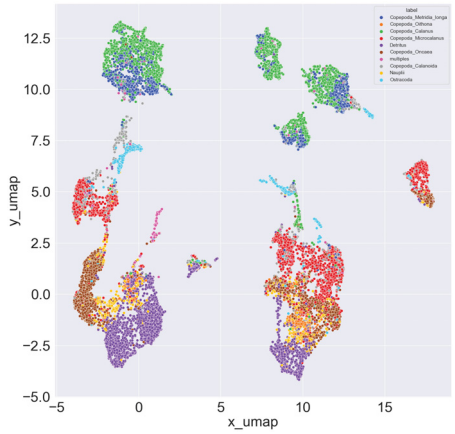

FIGURE 8
Learned latent-space Embedding using UMAP: The UMAP plot of the 512-dimensional latent space representation is displayed, which was obtained from self-supervised pre-training with DINO. This plot displays the ten categories, which accounted for 95% of the data. The UMAP plot provides an alternative visualization of the clustering of these species in the latent space. The color corresponds to the images categories as validated by the scientists.

for a classification task using annotated images. Numerous such methods exist, for example, Momentum contrast (MoCo), Self-distillation with no labels (DINO),and Simple Framework for Contrastive Learning (SimCLR): (Chen et al., 2020; He et al., 2020; Caron et al., 2021). We applied the DINO method for our work but did not evaluate the other approaches. It's worth noting that these alternative methods incur additional computational overhead compared to simply utilizing the pre-trained ImageNet models. Our framework is designed to be modular. We have presented two potential classifiers in this paper, but it can be extended with more classifiers, which should be implemented in PyTorch. Thus, we abstain from comparing our approach to others since both our classifiers are based on a relatively simple backbone architecture. This architecture may be replaced or expanded, given sufficient resources. Our framework facilitates the training and inference of these methods. To make appropriate methods comparisons, standardized benchmark datasets should be employed. Our study utilized exclusive proprietary data due to the specialized nature of our LOKI dataset. Our main goal is to simplify workflow and integrate the self-supervised learning technique Dino in LOKI images, along with presenting a detailed representation of the latent space.

To enhance the initial sorting process and the precision of image categorization, we incorporated a confidence threshold for the neural network. This threshold enables us to exercise greater caution when assigning images to specific category folders. As a consequence, an image that may have previously been classified with uncertainty is now allocated to the "unknown" folder. This approach minimizes the risk of mislabeling images. However, it is important to note that implementing this method may result in a trade-off with recall, which refers to the ability to detect true positives. While our focus on improving classification precision may lead to a decrease in recall, as we may detect fewer true positive cases, the overall result is a more reliable classification system with lower numbers of misclassifications. We don't tackle the problem of "Previously unseen classes and unknown particles" (Eerola et al., 2023); this could be future work and is out of scope for this study.

For the automatic removal of duplicate images, certain algorithms that do not rely on image parameters could be implemented. However, it is beyond the scope of this study.

To recap, our approach optimizes the allocation of resources by automatically categorizing images and allows researchers to focus on more complex cases, i.e., images that are marked as unknown by the algorithm. Our approach also enables the identification and separation of artifacts, such as dead tissue, eggs, and bubbles. This not only enhances the overall accuracy of the classification process but also saves valuable time by streamlining the analysis.

During our analysis, we noticed some mislabeled images within the annotated image data, particularly in categories that comprise copepod genera (i.e., Copepoda_Metridia, Copepoda_Pseudocalanus). This highlights the need for additional datasets to improve the performance of neural networks in zooplankton classification. Accurate identification of the calanoids is especially critical in deep learning-based analyses of LOKI images, as this is the specific area where the algorithms encountered the highest level of confusion. Focusing on improving the network's ability to correctly identify these species is crucial for enhancing the overall effectiveness of the analysis. Furthermore, our study successfully demonstrated the classification of species with a relatively low number of images in the training data. This indicates the potential of our approach to effectively classify zooplankton species even when limited training data is available. For instance, we achieved an F1-Score of 0.83 for Chaetognata_middle, despite having only 1063 training/validation available for training and 184 test images for testing. However, our algorithm encountered difficulties when dealing with smaller categories, such as Foraminifera, with a training dataset size of only 121 examples.

We observed the trend that for categories with less than 1000 training images, the F1-Score was below 50% with the exception of Cnidaria, which are morphologically very distinct from all other categories. The accuracy of neural networks, in general, (Kavzoglu, 2009; LeCun et al., 2015; Krizhevsky et al., 2017) tends to increase with data set size. This is because a larger dataset provides more diverse examples, allowing the network to learn more robust representations of the underlying patterns. Based on more data, the network can also better capture the underlying distribution of the problem space and improve its generalization capabilities. However, it's worth noting that the relationship between data size and accuracy is not always linear and can reach a plateau or even decrease with excessively large datasets (Echle et al., 2020). By augmenting (i.e., rotating the data) the training data, we can provide an algorithm with a more comprehensive and representative set of examples, enabling it to better understand and classify groups that have limited representation in the current dataset.

Our analysis of the latent space embedding of zooplankton images obtained through self-supervised pre-training using ResNet18 revealed an insightful observation. By examining the

10.3389/fmars.2023.1280510

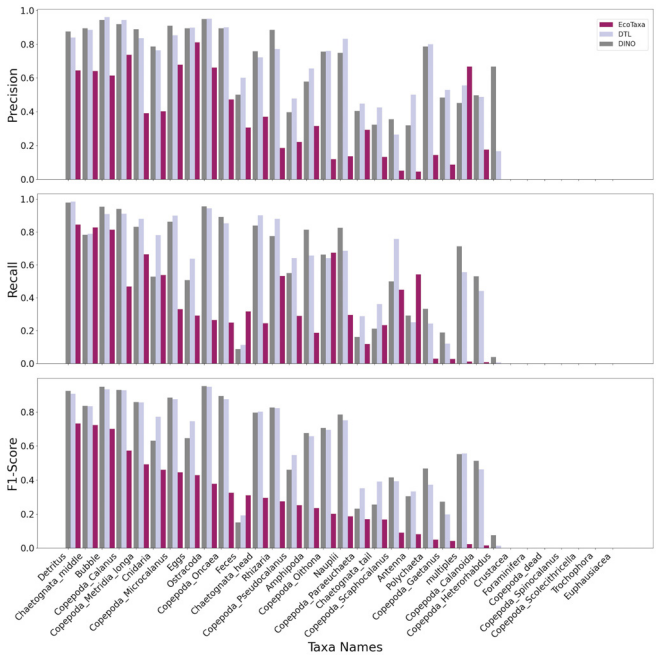

FIGURE 9
Precision, Recall and F1-Score on the hold out test dataset the images of Cruise P99.2 for the classes. The y-axis indicates the scores and the x-axis indicates the label for each of the 33 classes.

metadata, specifically station information, we discovered a distinct clustering pattern that was not influenced by the pre- or training process of the ResNet18 model. These clusters were primarily associated with three specific geographic locations (longitude, latitude), which were separated from all other locations in the dataset. When we tested the image parameters extracted from the LOKI browser, we found significant differences in most descriptors, indicating that the images from different geographic locations indeed differed in their optical properties (For example, see Supplementary Material DeepLOKI Table S1 and Supplementary Material DeepLOKI Table S2 that display the variations in image parameters and image descriptors for Copepoda_Calanus). At present, these differences cannot be explained by, for example, zooplankton population dynamics, and more detailed analyses would be necessary to make use of this result. Deep-learning-supported annotation has some requirements that the researchers must meet. For example, the DINO approach that we implemented needs access to sufficient computing power. We recommend re-pre-training the DINO ResNet18 when a sufficient number of new data (i.e. the next 200,000 images) is available through additional cruises. If these images are annotated, DINO and DTL should undergo retraining utilizing the finetuning techniques. To mitigate such resource limitations, we employed a small ResNet18 model as the backbone, enabling inference (pre-sorting of the images) on portable computers and mitigating all resource limitations. Our data clearly illustrate the ability of various algorithmic approaches to differentiate among the 33 categories present in our dataset. Nevertheless, during the training phase, we observed that the ResNet18 model with student-teacher self-supervised pre-training exhibited strong performance for the specific task at hand. The current method performed successfully and promises to enhance LOKI image analyses. It has to be noted that LOKI is currently used only in few working groups, however, with the technical improvements of cameras, the resolution of images from other optical systems used in marine studies may increases considerably. DeepLOKI does not require high-resolution images only; it resizes (convert down in resolution) the image down to 224 x224 pixel.

Therefore, our pipeline can serve as a valuable principle for future applications. Its successful performance and ability to handle the challenges specific to high-resolution images that present great morphological details could make it a promising framework that can be adapted and extended for further image-based research and analysis in marine environments.

5 Conclusion

Our study presents a novel and effective approach for zooplankton image classification using a self-supervised pre-trained ResNet18 model in conjunction with the LOKI system. The results

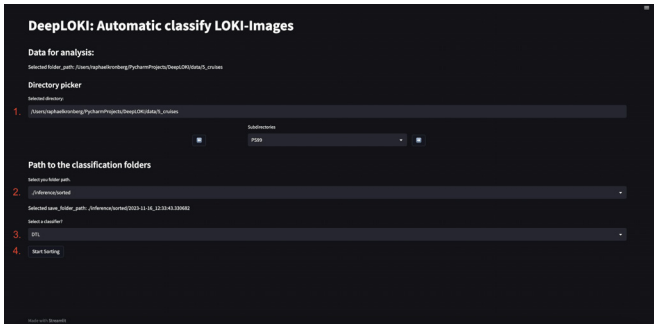

FIGURE 10
Graphical User Interface (GUI) of the DeepLOKI tool - Sorting:1. A directory picker to select the folder with the image, that should be pre-sorted. 2. A second input field to determine the location of class folders and the unknown folder. 3. A drop down menu to select the classifier. Folder directly from the Loki System can be process in a few minutes. 4. The Start Sorting Button for starting the sorting process.

clearly demonstrate the advantages of this approach over traditional methods, such as random forest classification, especially when a large training dataset is available. Compared to the default random forest available in EcoTaxa, the ResNet18 classification performance was two times higher. The DeepLOKI framework offers streamlined processes that save time and minimize potential errors, eliminating the need for calculating image parameters and providing a more efficient pipeline from the LOKI system to the classified images. Deep learning-based classification not only improves accuracy but also expedites the estimation of organism density at each station. By considering the potential for miss-classification, density is inferred based on the number of organisms sorted into specific folders.

Despite the promising results, there is still room for further improvement by including additional data samples to enhance the model's performance.

One of the advantages of our deep learning framework is its versatility and flexibility. Besides handling just images from a single device like the LOKI system, the framework can be easily adapted to different image data, for example, phytoplankton obtained from multiple devices. By leveraging the power of deep learning, our framework effectively extracts meaningful features and patterns to classify images, making it a valuable tool for researchers working with diverse datasets in various domains.

In conclusion, our study demonstrates the efficacy of utilizing a self-supervised pre-trained ResNet18 model in combination with the LOKI system for zooplankton image classification. The DeepLOKI framework offers improved performance, efficiency, and adaptability, making it a promising approach for advancing zooplankton research and facilitating accurate analysis across different ecosystems.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The data used for the model training, validation and testing are part of other research projects, not all of these dataset are fully annotated by now. The datasets to test the GUI can be found in the Sciebo https://uni-duesseldorf.sciebo.de/s/okWh4728VwnCBGp. The codes available at https://gitlab.com/qtb-hhu/qtb-sda/DeepLOKI. Requests to access these datasets should be directed to Barbara Niehoff Barbara.Niehoff@awi.de; Ellen Oldenburg Ellen.Oldenburg@hhu.de.

Author contributions

EO: Data curation, Formal analysis, Methodology, Project administration, Supervision, Visualization, Writing – original draft, Conceptualization. RMK: Conceptualization, Formal analysis, Methodology, Software, Visualization, Writing – original draft. BN: Data curation, Funding acquisition, Resources, Validation, Writing – review & editing. OE: Funding acquisition, Writing – review & editing. OP: Conceptualization, Resources, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was further 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)(EO and OE). The image analyses were supported by the BMBF funded projects QUARCS (grant number 03F0777A) and MAZE (03V01666).

Acknowledgments

Computational infrastructure and support were provided by the Centre for Information and Media Technology at Heinrich Heine University Düsseldorf. We thank Nicole Hildebrandt, Benjamin

Meyer and Dong-yung Kim for evaluating the automatic LOKI image classification. We also thank the scientists, the captains and crews of RV Polarstern for their excellent support at sea during the, expeditions included for this study. Ship time for RV Polarstern was provided under grants AWI_PS99, AWI_PS106, AWI_PS107_05, AWI_PS114_01, AWI_PS122 of RV Polarstern. We are grateful to both reviewers for their valuable input and time dedicated to improving our manuscript. Their insights have been instrumental in enhancing its quality.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2023.1280510/full#supplementary-material

References

Bi, H., Guo, Z., Benfield, M. C., Fan, C., Ford, M., Shahrestani, S., et al. (2015). A semi-automated image analysis procedure for in *situ* plankton imaging systems. *PloS One* 10, e0127121. doi: 10.1371/journal.pone.0127121

Bi, H., Song, J., Zhao, J., Liu, H., Cheng, X., Wang, L., et al. (2022). Temporal characteristics of plankton indicators in coastal waters: high-frequency data from planktonscope. *J. Sea Res.* 189, 102283. doi: 10.1016/j.seares.2022.102283

Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P., et al. (2021). "Emerging properties in self-supervised vision transformers," *2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada*. 9650–9660. doi: 10.1109/ICCV48922.2021.00951

Chen, T., Kornblith, S., Norouzi, M., and Hinton, G. (2020). "A simple framework for contrastive learning of visual representations," in *International conference on machine learning (PMLR)*, vol 119 of Proceedings of Machine Learning Research 1597–1607. https://proceedings.mlr.press/v119/chen20j.html

Cheng, K., Cheng, X., Wang, Y., Bi, H., and Benfield, M. C. (2019). Enhanced convolutional neural network for plankton identification and enumeration. *PloS One* 14, e0219570. doi: 10.1371/journal.pone.0219570

Cornils, A., Thomisch, K., Hase, J., Hildebrandt, N., Auel, H., and Niehoff, B. (2022). Testing the usefulness of optical data for zooplankton long-term monitoring: Taxonomic composition, abundance, biomass, and size spectra from zooscan image analysis. *Limnol. Oceanogr.: Methods* 20, 428–450. doi: 10.1002/lom3.10495

Cowen, R. K., and Guigand, C. M. (2008). *In situ* ichthyoplankton imaging system (isiis): system design and preliminary results. *Limnol. Oceanogr.: Methods* 6, 126–132. doi: 10.4319/lom.2008.6.126

Echle, A., Grabsch, H. I., Quirke, P., van den Brandt, P. A., West, N. P., Hutchins, G. G., et al. (2020). Clinical-grade detection of microsatellite instability in colorectal tumors by deep learning. *Gastroenterology* 159, 1406–1416. doi: 10.1053/j.gastro.2020.06.021

Eerola, T., Batrakhanov, D., Barazandeh, N. V., Kraft, K., Haraguchi, L., Lensu, L., et al. (2023). Survey of automatic plankton image recognition: Challenges, existing solutions and future perspectives. *arXiv preprint arXiv:2305.11739*. doi: 10.48550/arXiv.2305.11739

Elineau, A., Desnos, C., Jalabert, L., Olivier, M., Romagnan, J.-B., Brandao, M., et al. (2018). Zooscannet: plankton images captured with the zooscan. *SEANOE*.

González, P., Castaño, A., Peacock, E. E., Díez, J., Del Coz, J. J., and Sosik, H. M. (2019). Automatic plankton quantification using deep features. *J. Plankton Res.* 41, 449–463. doi: 10.1093/plankt/fbz023

Gorsky, G., Guilbert, P., and Valenta, E. (1989). The autonomous image analyzer–enumeration, measurement and identification of marine phytoplankton. *Mar. Ecol. Prog. Ser.* 58, 133–142. doi: 10.3354/meps058133

Grosjean, P., Picheral, M., Warembourg, C., and Gorsky, G. (2004). Enumeration, measurement, and identification of net zooplankton samples using the zooscan digital imaging system. *ICES J. Mar. Sci.* 61, 518–525. doi: 10.1016/j.icesjms.2004.03.012

Hauss, H., Christiansen, S., Schütte, F., Kiko, R., Edvam Lima, M., Rodrigues, E., et al. (2016). Dead zone or oasis in the open ocean? zooplankton distribution and migration in low-oxygen modewater eddies. *Biogeosciences* 13, 1977–1989. doi: 10.5194/bg-13-1977-2016

He, K., Fan, H., Wu, Y., Xie, S., and Girshick, R. (2020). "Momentum contrast for unsupervised visual representation learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 9729–9738. doi: 10.1109/CVPR42600.2020.00975

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778. doi: 10.1109/CVPR.2016.90

Hirche, H., Barz, K., Ayon, P., and Schulz, J. (2014). High resolution vertical distribution of the copepod calanus Chilensis in relation to the shallow oxygen minimum zone off northern Peru using loki, a new plankton imaging system. *Deep Sea Res. Part I: Oceanogr. Res. Papers* 88, 63–73. doi: 10.1016/j.dsr.2014.03.001

Kavzoglu, T. (2009). Increasing the accuracy of neural network classification using refined training data. *Environ. Model. Softw.* 24, 850–858. doi: 10.1016/j.envsoft.2008.11.012

Kiko, R., Brandt, P., Christiansen, S., Faustmann, J., Kriest, I., Rodrigues, E., et al. (2020). Zooplanktonmediated fluxes in the eastern tropical north atlantic. *Front. Mar. Sci.* 7, 358. doi: 10.3389/fmars.2020.00358

Kingma, D. P., and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. doi: 10.48550/arXiv.1412.6980

Kraft, K., Velhonoja, O., Eerola, T., Suikkanen, S., Tamminen, T., Haraguchi, L., et al. (2022). Towards operational phytoplankton recognition with automated high-throughput imaging, near-real-time data processing, and convolutional neural networks. *Front. Mar. Sci.* 9, 867695. doi: 10.3389/fmars.2022.867695

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Commun. ACM* 60, 84–90. doi: 10.1145/3065386

Kronberg, R. M. (2022). Applications of Supervised Deep (Transfer) Learning for Medical Image Classification. Ph.D. thesis (HHU).

Kronberg, R. M., Haeberle, L., Pfaus, M., Xu, H. C., Krings, K. S., Schlensog, M., et al. (2022). Communicator-driven data preprocessing improves deep transfer learning of histopathological prediction of pancreatic ductal adenocarcinoma. *Cancers* 14, 1964. doi: 10.3390/cancers14081964

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539

Lertvilai, P. (2020). The in *situ* plankton assemblage explorer (ipax): An inexpensive underwater imaging system for zooplankton study. *Methods Ecol. Evol.* 11, 1042–1048. doi: 10.1111/2041-210X.13441

Lumini, A., and Nanni, L. (2019). Deep learning and transfer learning features for plankton classification. *Ecol. Inf.* 51, 33–43. doi: 10.1016/j.ecoinf.2019.02.007

Luo, J. Y., Irisson, J.-O., Graham, B., Guigand, C., Sarafraz, A., Mader, C., et al. (2018). Automated plankton image analysis using convolutional neural networks. *Limnol. Oceanogr.: Methods* 16, 814–827. doi: 10.1002/lom3.10285

McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*. doi: 10.48550/arXiv.1802.03426

Noroozi, M., Vinjimoor, A., Favaro, P., and Pirsiavash, H. (2018). "Boosting self-supervised learning via knowledge transfer," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 9359–9367. doi: 10.1109/CVPR.2018.00975

Orenstein, E. C., Ayata, S.-D., Maps, F., Becker, É. C., Benedetti, F., Biard, T., et al. (2022). Machine learning techniques to characterize functional traits of plankton from image data. *Limnol. Oceanogr.* 67, 1647–1669. doi: 10.1002/lno.12101

Orenstein, E. C., and Beijbom, O. (2017). "Transfer learning and deep feature extraction for planktonic image data sets," in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV) (IEEE)*. 1082–1088. doi: 10.1109/WACV.2017.125

Picheral, M., Catalano, C., Brousseau, D., Claustre, H., Coppola, L., Leymarie, E., et al. (2022). The underwater vision profiler 6: an imaging sensor of particle size spectra and plankton, for autonomous and cabled platforms. *Limnol. Oceanogr.: Methods* 20, 115–129. doi: 10.1002/lom3.10475

Picheral, M., Colin, S., and Irisson, J. (2017). *Ecotaxa, a tool for the taxonomic classification of images*.

Picheral, M., Guidi, L., Stemmann, L., Karl, D. M., Iddaoud, G., and Gorsky, G. (2010). The underwater vision profiler 5: An advanced instrument for high spatial resolution studies of particle size spectra and zooplankton. *Limnol. Oceanogr.: Methods* 8, 462–473. doi: 10.4319/lom.2010.8.462

Rubbens, P., Brodie, S., Cordier, T., Destro Barcellos, D., Devos, P., Fernandes-Salvador, J. A., et al. (2023). Machine learning in marine ecology: an overview of techniques and applications. *ICES J. Mar. Sci.* 80 (7), 1829–1853. doi: 10.1093/icesjms/fsad100

Schmid, M. S., Aubry, C., Grigor, J., and Fortier, L. (2015). *Zoomie v1. 0 (zooplankton multiple image exclusion)*.

Schmid, M. S., Aubry, C., Grigor, J., and Fortier, L. (2016). The loki underwater imaging system and an automatic identification model for the detection of zooplankton taxa in the arctic ocean. *Methods Oceanogr.* 15, 129–160. doi: 10.1016/j.mio.2016.03.003

Schulz, J., Barz, K., Ayon, P., Luedtke, A., Zielinski, O., Mengedoht, D., et al. (2010). Imaging of plankton specimens with the lightframe on-sight keyspecies investigation (loki) system. *J. Eur. Optical Society Rapid Publications* 5. doi: 10.2971/jeos.2010.10017s

Sosik, H. M., Peacock, E. E., and Brownlee, E. F. (2015) *Annotated plankton images data set for developing and evaluating classification methods*. Available at: http://darchive.mblwhoilibrary.org/handle/1912/7341.

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., and Liu, C. (2018). "A survey on deep transfer learning," in *International conference on artificial neural networks (Springer)*. 270–279. doi: 10.1007/978-3-030-01424-7_27

van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-sne. *J. Mach. Learn. Res.* 9, 2579–2605.

Werner, J., Kronberg, R. M., Stachura, P., Ostermann, P. N., Müller, L., Schaal, H., et al. (2021). Deep transfer learning approach for automatic recognition of drug toxicity and inhibition of sars-cov-2. *Viruses* 13, 610. doi: 10.3390/v13040610

Yang, K., Qinami, K., Fei-Fei, L., Deng, J., and Russakovsky, O. (2020). "Towards fairer datasets: Filtering and balancing the distribution of the people subtree in the imagenet hierarchy," in *Conference on Fairness, Accountability, and Transparency*. 547–558. doi: 10.1145/3351095.3375709

Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How transferable are features in deep neural networks? *Adv. Neural Inf. Process. Syst.* 27.

Yue, J., Chen, Z., Long, Y., Cheng, K., Bi, H., and Cheng, X. (2023). Towards efficient deep learning system for *in-situ* plankton image recognition. *Front. Mar. Sci.* 10, 1186343. doi: 10.3389/fmars.2023.1186343

Chapter 7

Conclusion

7.1 Key results

The objective of this thesis is to elucidate the complex dynamics of Arctic Ocean ecosystems, with a specific focus on the Fram Strait. To achieve this, a multifaceted approach combining mathematical models, descriptive statistics, and artificial intelligence (Chapters 4 to 6) was employed. This research aimed to describe the impacts of environmental changes, particularly Atlantification, on various microbial and planktonic communities (Chapters 2 to 4) .
The study, comprising five distinct papers, provides a deeper understanding of the impact of these changes on the composition, structure, and functionality of Arctic marine ecosystems. The successful utilization of diverse data sources, including 16S and 18S amplicon data, gene sequences, metagenome data, and various environmental parameters from Remote Access Samplers (RAS) and the CTD, has been instrumental in achieving this deeper perspective (Chapter 2). Furthermore, zooplankton images were captured using the Lightframe On-sight Keyspecies Investigation (LOKI) device. This comprehensive suite of data sources enabled a detailed analysis of microbial and zooplankton communities in the Arctic Ocean.
This thesis employs a range of sophisticated methodologies to analyze extensive datasets collected from the Arctic Ocean. Descriptive statistics and predictive modeling based on statistical physics provided insights into the behavior of selected taxa and sub-communities under varying environmental conditions (Chapters 3 to 5). The collection of high-resolution temporal data via autonomous samplers and in situ sensors enabled the study of microbial dynamics (Chapter 2). Furthermore, a novel framework was developed to identify co-occurrence patterns, interaction networks, and keystone species within microbial communities (Chapter 5). Building upon these findings, the final study explored the application of computer vision and deep learning techniques for automated zooplankton image classification (Chapter 6). This data can be integrated into a comprehensive time-series analysis.
This thesis has highlighted the multifaceted impacts of environmental change on the Arctic Ocean ecosystem, with a particular focus on the consequences of sea ice retreat and Atlantification, due to their specific adaptations in life cycle, ecology, and physiology. Our findings indicate that specialized and rare taxa may face significant challenges under these changing conditions (Chapter 2). Consequently, further large-scale studies are necessary to identify these specialized taxa and clarify the impact of climate change on their abundance and diversity. In addition, to biogeographical studies, ecophysiological studies are essential for assessing the responses of Arctic specialists to environmental changes in comparison to widespread Arctic

generalists. As a consequence of climate change, the melting of sea ice is occurring at an earlier stage (Meredith et al., 2019), while its formation is occurring at a later point in the year. This has an impact on the abundance of sea ice communities and the subsequent spring bloom (Chapters 2 and 5).

The primary findings of our first study (Chapter 2) demonstrated that environmental conditions, such as sea ice melt, exert a significant influence on Arctic phytoplankton species. Polar pelagic and ice-associated taxa were more prevalent in Atlantic water-influenced regions, while temperate taxa struggled in colder waters. This trend suggests a substantial shift in phytoplankton composition as the Arctic warms, which could potentially trigger cascading effects on food web dynamics and nutrient cycling.

The second paper (Chapter 3) builds on this foundation by examining bacterial communities in the East Greenland Current (EGC). It highlights the impact of Atlantic water influx and changing sea-ice cover on these communities. A comprehensive four-year analysis revealed significant alterations in microbial composition, structure, and functionality. The influx of warmer Atlantic water and the reduction in sea ice cover led to changes in the relative abundance of different bacterial taxa, with potential implications for ecosystem processes such as carbon cycling and primary production. This study emphasises the necessity of elucidating microbial responses to environmental alterations in order to anticipate the future state of the Arctic Ocean at EGC.

In our third paper (Chapter 4), we investigate the dynamics of the bacterial microbiome at the West Spitzbergen Current (WSC). We identified five distinct seasonal modules within the prokaryotic microbiome, each linked to specific microeukaryotic populations and environmental conditions. These modules represent unique ecological states that recur annually, thereby illustrating the pronounced seasonal dynamics of Arctic microbial communities. Our findings indicate that environmental factors and microbial community structure are intricately linked, with different selection pressures exerted across the various seasonal states. The thesis highlight the intricate link between environmental factors and microbial community structure, contributing to a deeper understanding of the temporal patterns regulating microbial diversity and functionality in the Arctic pelagic ecosystem.

The fourth paper (Chapter 5) presents a novel framework for identifying keystone species through the analysis of co-occurrence patterns, interaction networks, and stability analysis within microbial communities. The framework was tested on a four-year dataset from the WSC, and it identified keystone species, including certain microeukaryotes, which are crucial for the Arctic food web's structure and resilience. Furthermore, this framework allowed for the modelling of community responses to environmental shifts, thereby providing insights into the mechanisms of ecosystem stability and potential future impacts. The applicability of this framework extends beyond microbes; future studies could incorporate micro-scale organisms such as viruses and phages, as well as macroscopic zooplankton, in order to gain a more holistic understanding of the ecosystem.

The fifth paper (Chapter 6) on zooplankton classification introduced DeepLOKI, a deep learning-based method for classifying zooplankton images. This innovation addressed the challenge of rapidly processing large volumes of image data collected during research cruises, which

would otherwise have been impractical to process manually. The automation of zooplankton taxa identification enabled the integration of image data into our time-series analyses. This approach facilitated the tracking of changes in zooplankton distribution and abundance with greater efficiency. The results demonstrated that the incorporation of image-based data significantly enhances the capacity to monitor and understand zooplankton dynamics in relation to environmental changes. The frame is not restricted to LOKI images, the underlying networks can be train and then applied to various image sources, given proper annotated data.
Overall, this thesis provides a comprehensive analysis of Arctic Ocean ecosystems, utilizing advanced methodologies to reveal the impacts of environmental changes on microbial and planktonic communities. The findings underscore the critical need for continued research to understand and mitigate the effects of climate change on these vital ecosystems.

7.2 Future Work

The objective of researchers is to achieve a more comprehensive understanding of the Arctic food web. This will necessitate the development of a robust ecosystem model with respect to the initial conditions and the incorporation of additional communities and data types. The results provide a foundation for future research investigating the combined dynamics of bacteria, archaea, and microeukaryotes in the Arctic food web.
Our research will be expanded to include other key players in the Arctic ecosystem, such as phages and zooplankton. Phages, which regulate bacterial populations and biogeochemical cycles, are crucial for understanding the dynamics of this unique and ever-changing environment (Heinrichs et al., 2023). Further arctic expeditions are planned to collect phage samples with the objective of expanding the current pipeline and further investigating the interactions between organisms.
The incorporation of additional data types, such as metagenomes and metatranscriptomes, into our amplicon (16S and 18S) data sets will facilitate the generation of functional insights in our analyses. The metagenomes of the Arctic Ocean represent the collective genetic material obtained directly from environmental samples. They provide insights into the diversity and functional potential of microbial communities, as demonstrated in Chapter 4. However, it is necessary to expand the combined analysis. Metatranscriptomes provide a comprehensive view of RNA transcripts present at a given time, illuminating active genes and metabolic pathways and enabling a dynamic understanding of how these communities respond to environmental changes (Pearson et al., 2015).
In the medium to long term, our objective is to develop a comprehensive ecosystem model that encompasses several spatial, temporal, and biological scales. This model will be used to model community dynamics over the duration of a research cruise, such as the MOSAiC expedition (Mock et al., 2022), where the data varies in time and space. The resulting model will facilitate the study of complex interactions within the Arctic food web and their impact on biogeochemical nutrient cycles. The utilization of novel technologies, including DeepLOKI, is intended to facilitate a greater comprehension of species abundance and ecosystem functionality.

In this context, the feasibility of developing mechanistic models capable of making predictive statements about the future development of ecosystems is of particular interest. One example of such models is ecosystem models based on the principles of generalized Lotka-Volterra (gLV) theory (Lotka, 1920; Volterra, 1927). These models mathematically represent few species interactions, typically between predators and prey, and capture population oscillations and equilibrium states over time (Malcai et al., 2002). One challenge in this field is that we often lack knowledge about the specific interactions between species. These interactions can change in response to environmental factors, and it is not always straightforward to map them. For instance, a species may engage in multiple different roles of interactions with other species, depending on the species involved.

Therefore, implementing this approach presents several immediate challenges. What are the actual difficulties of such a predictive dynamic model, and what resources are needed to overcome them? The first issue is the exponential growth of the parameters relative to the action parameters. Experimental methods to determine these parameters include the isolation of communities, which proves challenging in real ecosystems due to the significant number of influencing parameters. Another issue is the number and selection of the parameters due to the presence of correlation and dependencies. Despite these limitations, no single model can currently be used to model the full complexity of the Arctic food web.

Despite the availability of all requisite data, the quality of the forecast may not be guaranteed. The model may lack sufficient robustness with respect to the initial conditions. This will enable reliable predictions and necessitating simplification. It is of critical importance to determine the optimal level of complexity that will maximize the model's predictive power (Burnham and Anderson, 2002; MacKay, 2003). One might inquire whether it is expedient to construct a model comprising 200 species. A smaller number of functional guilds may be a superior solution, analogous to the clustering approach described in Chapter 5.

The multitude of variables that affect an organism's behavior renders a comprehensive examination impractical. An overly granular analysis of too many details can result in the obscuring of the overall picture. By aggregating or integrating smaller units into larger entities, patterns in the data can become more discernible. For instance, as an alternative to modeling dynamics based on individual species, it may prove more fruitful to focus on higher taxonomic levels such as family or genus, particularly since a substantial number of species remain poorly understood. This prompts the question of which new approaches are required.

The application of LV models necessitate the measurement of absolute densities, whereas the typical datasets only provide estimates of relative abundances. To address these challenges, various approaches have been developed. One approach is to derive a new nonlinear dynamical system for microbial dynamics, termed the compositional Lotka-Volterra (cLV) model. This unifies approaches using generalized Lotka-Volterra (gLV) equations from community ecology and compositional data analysis (Joseph et al., 2020).

Goldford et al., 2018 paper provides an illustration of the complexity of systems and the potential for simplification and reduction. Promising approaches have already been developed for another microbial system. Very diverse and taxonomically rich samples were collected to analyze the microbiome of soils and plant leaves. Between 110 and 1290 exact sequence vari-

ants were sequenced. The study investigated which species survived on a minimal medium with glucose, hypothesizing that only one species should survive on a limited resource utilising consumer research models. Despite the simplicity of the consumer-resource model, they may be capable of capturing many of the observed qualitative features in the experiment, as well as the more subtle aspects, including the existence of temporal blooms. It is noteworthy that approximately five to seventeen species exhibited remarkable resilience, demonstrating the robustness of the microbial community at the family level despite the observed variations within individual species. This observation indicates that a detailed examination is unnecessary to recognize basic patterns. Another study (Marsland III et al., 2019) introduces an adapted version of MacArthur's Consumer Resource Model, which has been tailored for microbial ecosystems. The ecological complexity of these communities has made it challenging to develop a universal predictive model that accounts for the variability of these features in response to environmental conditions. To address this, the authors developed a versatile simulation framework and explored a wide range of parameters and modelling assumptions in order to uncover common patterns. This model incorporates energetic, stochastic colonization, and the exchange and consumption of metabolites, focusing on large, diverse ecosystems with numerous species and resources. To evaluate the community structure and resource stability, the authors conducted simulations of ecosystems comprising 200 species and 100 resources, with species exhibiting an average of 10 resource preferences. Parameters and consumer preferences were randomly generated to simulate diverse community dynamics under varying conditions. The findings revealed two fundamentally different types of community structures that align with observed biodiversity patterns and offer new insights into their stability and functionality. This demonstrated which parameters indicated overall energy flow, with the original nutrient-derived energy determining the number of coexisting species. The methods presented are examples of further analytical approaches applicable to ocean microbiomes. The subsequent stage entails examining and simplifying natural communities under laboratory conditions with the objective of identifying comparable patterns in different environments. This represents a preliminary step towards an understanding of the complexity of the Arctic Ocean ecosystem. It is assumed that ecosystems have various niches whose exact functions are random, yet their roles are analogous across cases. General data analyses eventually encounter limitations in identifying actual interaction principles, necessitating further comprehensive experiments of reduced complexity in controlled, reproducible environments. To address this, further comprehensive experiments of reduced complexity in controlled, reproducible environments are required. Consequently, it is necessary to grow these samples in a controlled environment in order to apply them to different conditions. Conversely, field work in the Arctic is also of paramount importance, as evidenced by the many previous expeditions to the region. Developing hypotheses about ecosystem interactions requires an empirical testing capacity, which is difficult to achieve in the ecosystem itself.

An illustrative example of experiments of reduced complexity in controlled, reproducible environments is the use of small-scale bioreactors and medium- to large-scale planktotrons (Gall et al., 2017; Mustaffa et al., 2020). Planktotrons represent a distinctive experimental apparatus used to study the behavior of planktonic organisms in a controlled environment. These tanks

are constructed with high transparency, allowing for precise regulation of the water's physical and chemical conditions. Researchers can investigate the specific effects of varying factors on plankton, such as temperature, light, or nutrient availability. Planktotrons facilitate long-term observation of changes, which can often be challenging in natural waters. They contribute to the understanding of plankton population development and ecosystem reactions to environmental changes. These controlled experiments enhance the reproducibility of scientific results and aid in comprehending the functions and evolution of marine ecosystems in response to diverse environmental stimuli.

In conclusion, these studies and the methodologies employed underscore the necessity of developing robust, mechanistic models to predict ecosystem dynamics in the Arctic. By utilizing controlled experimental systems such as planktotrons and integrating comprehensive data analyses, we can simplify and understand complex marine ecosystems. These efforts will not only enhance our predictive capabilities but also deepen our understanding of the intricate interactions within these ecosystems.

Personal Publications

Peer-Reviewed Journal Articles

Thomas Mock et al. (Oct. 2022). "Multiomics in the central Arctic Ocean for benchmarking biodiversity change". In: *PLOS Biology* 20.10, pp. 1–6. DOI: 10.1371/journal.pbio.3001835. URL: https://doi.org/10.1371/journal.pbio.3001835

Ellen Oldenburg et al. (2023b). "DeepLOKI-a deep learning based approach to identify zooplankton taxa on high-resolution images from the optical plankton recorder LOKI". in: *Frontiers in Marine Science* 10, p. 1280510

Taylor Priest et al. (2023). "Atlantic water influx and sea-ice cover drive taxonomic and functional shifts in Arctic marine bacterial communities". In: *The ISME Journal* 17.10, pp. 1612–1625

Ellen Oldenburg et al. (2024b). "Sea-ice melt determines seasonal phytoplankton dynamics and delimits the habitat of temperate Atlantic taxa as the Arctic Ocean atlantifies". In: *ISME communications* 4.1, ycae027

Peer-Reviewed Journal Reviews

Ovidiu Popa et al. (2020b). "From sequence to information". In: *Philosophical Transactions of the Royal Society B* 375.1814, p. 20190448

Veronica Estrada et al. (2022). "Mapping the long rocky road to effective spinal cord injury therapy: a meta-review of pre-clinical and clinical research". In: *Journal of Neurotrauma* 39.9-10, pp. 591–612

Manuscript Submitted for Publication

Allison Fong et al. (2023). "Overview of the MOSAiC expedition: Ecosystem". In: *EarthArXiv*, pp. 2023–01

Ellen Oldenburg et al. (2024a). "Beyond blooms: A novel time series analysis framework predicts seasonal keystone species and sheds light on Arctic pelagic ecosystem stability". In: *bioRxiv*, pp. 2024–03

Taylor Priest et al. (2024). "Seasonal recurrence and modular assembly of an Arctic pelagic marine microbiome". In: *bioRxiv*, pp. 2024–05

Abbreviations

AI Artificial intelligence. 23, 24, 28, 29, 118, 119

ASV Amplicon Sequence Variant. 51, 67, 69, 91, 185

AW Atlantic Water. 4

CAO Central Arctic Ocean. 4–6

CCM convergent cross mapping. 22

cLV compositional Lotka-Volterra. 142

CMIP5 Coupled Model Intercomparison Project. 3

CO_2 Carbon Dioxide. 4, 13–15, 183

CTD conductivity-temperature-depth rosette. 139

DINO self-**di**stillation with **no** labels. 117, 186

DL Deep learning. 23, 24

DNA Deoxyribonucleic acid. 19

DTL Deep Transfer Learning. 117, 119, 186

EGC East Greenland Current. 1, 2, 26, 50, 51, 140

ELA Energy landscape analysis. 23

EPS extracellular polymeric substances. 8, 9

FYI First year ice. 5, 7

gLV generalized Lotka-Volterra. 142

IPCC Intergovernmental Panel on Climate Change. 3

LOKI Lightframe On-sight Keyspecies Investigation. 28, 117–119, 139, 141, 186

LTER Long-Term Ecological Research. 1

LV Lotka-Volterra. 142

ML Machine learning. 23

MOSAiC Multidisciplinary drifting Observatory for the Study of Arctic Climate. 2, 26, 50, 51, 141, 183

MYI Multi year ice. 5, 7

PML Polar Mixed Layer. 6

psu Practical Salinity Units. 5, 6

PW Polar Water. 2

RAS Remote Access Samplers. 139

RNA Ribonucleic acid. 19, 20, 141

UVP Underwater vision profiler. 119

WSC West Spitzbergen Current. 1, 2, 140

Bibliography

Aagaard, Knut, JH Swift, and EC Carmack (1985). “Thermohaline circulation in the Arctic Mediterranean seas”. In: *Journal of Geophysical Research: Oceans* 90.C3, pp. 4833–4846 (cit. on p. 6).

Admiraal, W and LAH Venekamp (1986). “Significance of tintinnid grazing during blooms of Phaeocystis pouchetii (Haptophyceae) in Dutch coastal waters”. In: *Netherlands Journal of Sea Research* 20.1, pp. 61–66 (cit. on p. 11).

Alexeev, Vladimir A and Craig H Jackson (2013). “Polar amplification: is atmospheric heat transport important?” In: *Climate dynamics* 41, pp. 533–547 (cit. on p. 15).

Alonso-Sáez, Laura, Olga Sánchez, Josep M Gasol, Vanessa Balagué, and Carlos Pedrós-Alio (2008). “Winter-to-summer changes in the composition and single-cell activity of near-surface Arctic prokaryotes”. In: *Environmental Microbiology* 10.9, pp. 2444–2454 (cit. on p. 7).

Amann, Rudolf I, Wolfgang Ludwig, and Karl-Heinz Schleifer (1995). “Phylogenetic identification and in situ detection of individual microbial cells without cultivation”. In: *Microbiological reviews* 59.1, pp. 143–169 (cit. on p. 19).

Andersson, Tom R, J Scott Hosking, Marıéa Pérez-Ortiz, Brooks Paige, Andrew Elliott, Chris Russell, Stephen Law, Daniel C Jones, Jeremy Wilkinson, Tony Phillips, et al. (2021). “Seasonal Arctic sea ice forecasting with probabilistic deep learning”. In: *Nature communications* 12.1, p. 5124 (cit. on p. 24).

Aparıécio, Sara (2023). “Earth observations of Melt Ponds on Sea Ice”. In: *The Cryosphere Discussions* 2023, pp. 1–65 (cit. on p. 9).

Appen, Wilken-Jon von, Anya M Waite, Melanie Bergmann, Christina Bienhold, Olaf Boebel, Astrid Bracher, Boris Cisewski, Jonas Hagemann, Mario Hoppema, Morten H Iversen, et al. (2021). “Sea-ice derived meltwater stratification slows the biological carbon pump: results from continuous observations”. In: *Nature Communications* 12.1, p. 7309 (cit. on pp. 6, 8, 18).

Ardyna, Mathieu, CJ Mundy, Nicolas Mayot, Lisa C Matthes, Laurent Oziel, Christopher Horvat, Eva Leu, Philipp Assmy, Victoria Hill, Patricia A Matrai, et al. (2020). “Under-ice phytoplankton blooms: Shedding light on the “invisible” part of Arctic primary production”. In: *Frontiers in Marine Science* 7, p. 608032 (cit. on p. 10).

Aristegui, Javier, Josep M Gasol, Carlos M Duarte, and Gerhard J Herndld (2009). “Microbial oceanography of the dark ocean’s pelagic realm”. In: *Limnology and Oceanography* 54.5, pp. 1501–1529 (cit. on p. 7).

Arrigo, Kevin R (2014). "Sea ice ecosystems". In: *Annual review of marine science* 6, pp. 439–467 (cit. on pp. 3, 5, 8, 10).

– (2017). "Sea ice as a habitat for primary producers". In: *Sea ice*, pp. 352–369 (cit. on p. 3).

Arrigo, Kevin R, Gert van Dijken, and Sudeshna Pabi (2008). "Impact of a shrinking Arctic ice cover on marine primary production". In: *Geophysical Research Letters* 35.19 (cit. on pp. 5, 8).

Arrigo, Kevin R and Gert L van Dijken (2015). "Continued increases in Arctic Ocean primary production". In: *Progress in oceanography* 136, pp. 60–70 (cit. on p. 18).

Arrigo, Kevin R, Donald K Perovich, Robert S Pickart, Zachary W Brown, Gert L Van Dijken, Kate E Lowry, Matthew M Mills, Molly A Palmer, William M Balch, Frank Bahr, et al. (2012). "Massive phytoplankton blooms under Arctic sea ice". In: *Science* 336.6087, pp. 1408–1408 (cit. on pp. 10, 18).

Assmy, Philipp, Jens K Ehn, Mar Fernández-Méndez, Haakon Hop, Christian Katlein, Arild Sundfjord, Katrin Bluhm, Malin Daase, Anja Engel, Agneta Fransson, et al. (2013). "Floating ice-algal aggregates below melting Arctic sea ice". In: *PLoS One* 8.10, e76599 (cit. on p. 8).

Assmy, Philipp, Mar Fernández-Méndez, Pedro Duarte, Amelie Meyer, Achim Randelhoff, Christopher J Mundy, Lasse M Olsen, Hanna M Kauko, Allison Bailey, Melissa Chierici, et al. (2017). "Leads in Arctic pack ice enable early phytoplankton blooms below snow-covered sea ice". In: *Scientific reports* 7.1, p. 40850 (cit. on p. 15).

Bamber, Jonathan (2022). "The Arctic is Warming Nearly Four Times Faster Than the Rest of the World." In: *CounterPunch* (cit. on p. 3).

Barber, David G, Haakon Hop, Christopher J Mundy, Brent Else, Igor A Dmitrenko, Jean-Eric Tremblay, Jens K Ehn, Philipp Assmy, Malin Daase, Lauren M Candlish, et al. (2015). "Selected physical, biological and biogeochemical implications of a rapidly changing Arctic Marginal Ice Zone". In: *Progress in Oceanography* 139, pp. 122–150 (cit. on p. 18).

Barberán, Albert, Scott T Bates, Emilio O Casamayor, and Noah Fierer (2012). "Using network analysis to explore co-occurrence patterns in soil microbial communities". In: *The ISME journal* 6.2, pp. 343–351 (cit. on p. 22).

Bates, NR, MI Orchowska, R Garley, and JT Mathis (2013). "Summertime calcium carbonate undersaturation in shelf waters of the western Arctic Ocean–how biological processes exacerbate the impact of ocean acidification". In: *Biogeosciences* 10.8, pp. 5281–5309 (cit. on p. 11).

Bathmann, Ulrich V, Rolf Peinert, Thomas T Noji, and Bodo V Bodungen (1990). "Pelagic origin and fate of sedimenting particles in the Norwegian Sea". In: *Progress in Oceanography* 24.1-4, pp. 117–125 (cit. on p. 10).

Bauerfeind, Eduard, BV Bodungen, K Arndt, and Wolfgang Koeve (1994). "Particle flux, and composition of sedimenting matter, in the Greenland Sea". In: *Journal of marine systems* 5.6, pp. 411–423 (cit. on p. 10).

Bauerfeind, Eduard, Eva-Maria Nöthig, Agnieszka Beszczynska, Kirsten Fahl, Lars Kaleschke, Kathrin Kreker, Michael Klages, Thomas Soltwedel, Christiane Lorenzen, and Jan Wegner (2009). "Particle sedimentation patterns in the eastern Fram Strait during 2000–2005: Results from the Arctic long-term observatory HAUSGARTEN". In: *Deep Sea Research Part I: Oceanographic Research Papers* 56.9, pp. 1471–1487 (cit. on p. 10).

Baumas, Chloé and Mina Bizic (2024). "A focus on different types of organic particles and their significance in the open ocean carbon cycle". In: *Progress in Oceanography*, p. 103233 (cit. on p. 11).

Bayer-Giraldi, Maddalena, Ilka Weikusat, Hüseyin Besir, and Gerhard Dieckmann (2011). "Characterization of an antifreeze protein from the polar diatom Fragilariopsis cylindrus and its relevance in sea ice". In: *Cryobiology* 63.3, pp. 210–219 (cit. on p. 8).

Behrenfeld, Michael J, Kimberly H Halsey, Emmanuel Boss, Lee Karp-Boss, Allen J Milligan, and Graham Peers (2021). "Thoughts on the evolution and ecological niche of diatoms". In: *Ecological Monographs* 91.3, e01457 (cit. on p. 10).

Belan, Boris D, Gérard Ancellet, Irina S Andreeva, Pavel N Antokhin, Viktoria G Arshinova, Mikhail Y Arshinov, Yurii S Balin, Vladimir E Barsuk, Sergei B Belan, Dmitry G Chernov, et al. (2022). "Integrated airborne investigation of the air composition over the Russian sector of the Arctic". In: *Atmospheric Measurement Techniques* 15.13, pp. 3941–3967 (cit. on p. 15).

Bellouquid, Abdelghani and Marcello Delitala (2006). *Mathematical modeling of complex biological systems*. Springer (cit. on p. 23).

Berg, Gabriele, Daria Rybakova, Doreen Fischer, Tomislav Cernava, Marie-Christine Champomier Vergès, Trevor Charles, Xiaoyulong Chen, Luca Cocolin, Kellye Eversole, Gema Herrero Corral, et al. (2020). "Microbiome definition re-visited: old concepts and new challenges". In: *Microbiome* 8, pp. 1–22 (cit. on p. 19).

Berge, Jørgen, Malin Daase, Paul E Renaud, William G Ambrose, Gerald Darnis, Kim S Last, Eva Leu, Jonathan H Cohen, Geir Johnsen, Mark A Moline, et al. (2015a). "Unexpected levels of biological activity during the polar night offer new perspectives on a warming Arctic". In: *Current Biology* 25.19, pp. 2555–2561 (cit. on p. 25).

Berge, Jørgen, Paul E Renaud, Gerald Darnis, Finlo Cottier, Kim Last, Tove M Gabrielsen, Geir Johnsen, Lena Seuthe, Jan Marcin Weslawski, Eva Leu, et al. (2015b). "In the dark: a review of ecosystem processes during the Arctic polar night". In: *Progress in Oceanography* 139, pp. 258–271 (cit. on p. 5).

Berry, David and Stefanie Widder (2014). "Deciphering microbial interactions and detecting keystone species with co-occurrence networks". In: *Frontiers in microbiology* 5, p. 90985 (cit. on p. 22).

Beszczynska-Möller, Agnieszka, Eberhard Fahrbach, Ursula Schauer, and Edmond Hansen (2012). "Variability in Atlantic water temperature and transport at the entrance to the Arctic Ocean, 1997–2010". In: *ICES Journal of Marine Science* 69.5, pp. 852–863 (cit. on p. 4).

Bluhm, Bodil A and Rolf Gradinger (2008). "Regional variability in food availability for Arctic marine mammals". In: *Ecological Applications* 18.sp2, S77–S96 (cit. on p. 3).

Bluhm, Bodil A, Haakon Hop, Mikko Vihtakari, Rolf Gradinger, Katrin Iken, Igor A Melnikov, and Janne E Søreide (2018). "Sea ice meiofauna distribution on local to pan-Arctic scales". In: *Ecology and evolution* 8.4, pp. 2350–2364 (cit. on p. 8).

Boetius, Antje, Sebastian Albrecht, Karel Bakker, Christina Bienhold, Janine Felden, Mar Fernández-Méndez, Stefan Hendricks, Christian Katlein, Catherine Lalande, Thomas Krumpen, et al. (2013). "Export of algal biomass from the melting Arctic sea ice". In: *Science* 339.6126, pp. 1430–1432 (cit. on pp. 8, 10, 19).

Boeuf, Dominique, Florian Humily, and Christian Jeanthon (2014). "Diversity of Arctic pelagic Bacteria with an emphasis on photoheterotrophs: a review". In: *Biogeosciences* 11.12, pp. 3309–3322 (cit. on pp. 6, 7).

Botterell, Zara LR, Penelope K Lindeque, Richard C Thompson, and Nicola J Beaumont (2023). "An assessment of the ecosystem services of marine zooplankton and the key threats to their provision". In: *Ecosystem Services* 63, p. 101542 (cit. on p. 13).

Bowman, Jeff S (2015). "The relationship between sea ice bacterial community structure and biogeochemistry: A synthesis of current knowledge and known unknowns". In: *Elementa* 3, p. 000072 (cit. on pp. 9, 10).

Bowman, Jeff S, Simon Rasmussen, Nikolaj Blom, Jody W Deming, Søren Rysgaard, and Thomas Sicheritz-Ponten (2012). "Microbial community structure of Arctic multiyear sea ice and surface seawater by 454 sequencing of the 16S RNA gene". In: *The ISME journal* 6.1, pp. 11–20 (cit. on pp. 6, 9).

Bowman, John (2013). "Sea-ice microbial communities". In: (cit. on p. 8).

Box, George EP, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung (2015). *Time series analysis: forecasting and control*. John Wiley & Sons (cit. on p. 21).

Box, Jason E, William T Colgan, Torben Røjle Christensen, Niels Martin Schmidt, Magnus Lund, Frans-Jan W Parmentier, Ross Brown, Uma S Bhatt, Eugénie S Euskirchen, Vladimir E Romanovsky, et al. (2019). "Key indicators of Arctic climate change: 1971–2017". In: *Environmental Research Letters* 14.4, p. 045010 (cit. on p. 15).

Brown, Christopher W and James A Yoder (1994). "Coccolithophorid blooms in the global ocean". In: *Journal of Geophysical Research: Oceans* 99.C4, pp. 7467–7482 (cit. on p. 11).

Brussaard, CPD, AAM Noordeloos, H Witte, MCJ Collenteur, K Schulz, Andrea Ludwig, and Ulf Riebesell (2013). "Arctic microbial community dynamics influenced by elevated CO 2 levels". In: *Biogeosciences* 10.2, pp. 719–731 (cit. on p. 21).

Bunse, Carina and Jarone Pinhassi (2017). "Marine bacterioplankton seasonal succession dynamics". In: *Trends in microbiology* 25.6, pp. 494–505 (cit. on p. 7).

Burnham, Kenneth P and David R Anderson (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer (cit. on p. 142).

Burns, Bethany L, Daniel D Rhoads, and Anisha Misra (2023). "The use of machine learning for image analysis artificial intelligence in clinical microbiology". In: *Journal of clinical microbiology* 61.9, e02336–21 (cit. on p. 24).

Buškus, Kazimieras, Evaldas Vaičiukynas, Antanas Verikas, Saulė Medelytė, and Sergej Olenin (2021). "Prototype framework for deep learning driven semantic segmentation of Arctic seabed imagery". In: *Arctic science: Arctic change 2020 conference: book of abstracts*. Vol. 7. 1. Canadian science publishing (cit. on p. 24).

Buttigieg, Pier Luigi, Eduard Fadeev, Christina Bienhold, Laura Hehemann, Pierre Offre, and Antje Boetius (2018). "Marine microbes in 4D—using time series observation to assess the dynamics of the ocean microbiome and its links to ocean health". In: *Current opinion in microbiology* 43, pp. 169–185 (cit. on p. 22).

Cai, Shuxiang, Chuanxiang Wu, Wenguang Yang, Wenfeng Liang, Haibo Yu, and Lianqing Liu (2020). "Recent advance in surface modification for regulating cell adhesion and behaviors". In: *Nanotechnology Reviews* 9.1, pp. 971–989 (cit. on p. 8).

Callaghan, Terry V, Lars Olof Björn, Yuri Chernov, Terry Chapin, Torben R Christensen, Brian Huntley, Rolf A Ims, Margareta Johansson, Dyanna Jolly, Sven Jonasson, et al. (2004). "Effects on the function of arctic ecosystems in the short-and long-term perspectives". In: *AMBIO: a Journal of the Human Environment* 33.7, pp. 448–458 (cit. on p. 21).

Cao, Shunan, Weipeng Zhang, Wei Ding, Meng Wang, Shen Fan, Bo Yang, Andrew Mcminn, Min Wang, Bin-bin Xie, Qi-Long Qin, et al. (2020). "Structure and function of the Arctic and Antarctic marine microbiota as revealed by metagenomics". In: *Microbiome* 8, pp. 1–12 (cit. on p. 19).

Cao, Yong, Lingen Bian, and Jinping Zhao (2018). "Impacts of changes in sea ice and heat flux on Arctic warming". In: *Atmospheric and Climate Sciences* 9.1, pp. 84–99 (cit. on p. 5).

Carmack, Edward C, Michiyo Yamamoto-Kawai, Thomas WN Haine, Sheldon Bacon, Bodil A Bluhm, Camille Lique, Humfrey Melling, Igor V Polyakov, Fiamma Straneo, M-L Timmermans, et al. (2016). "Freshwater and its role in the Arctic Marine System: Sources,

disposition, storage, export, and physical and biogeochemical consequences in the Arctic and global oceans". In: *Journal of Geophysical Research: Biogeosciences* 121.3, pp. 675–717 (cit. on pp. 1, 6, 15).

Caron, David A, Peter D Countway, Adriane C Jones, Diane Y Kim, and Astrid Schnetzer (2012). "Marine protistan diversity". In: *Annual review of marine science* 4, pp. 467–493 (cit. on p. 11).

Cavalieri, Donald J and Claire L Parkinson (2012). "Arctic sea ice variability and trends, 1979–2010". In: *The Cryosphere* 6.4, pp. 881–889 (cit. on p. 15).

Cavan, Emma L, Stephanie A Henson, Anna Belcher, and Richard Sanders (2017). "Role of zooplankton in determining the efficiency of the biological carbon pump". In: *Biogeosciences* 14.1, pp. 177–186 (cit. on p. 13).

Chatfield, Chris and Haipeng Xing (2019). *The analysis of time series: an introduction with R*. Chapman and hall/CRC (cit. on p. 21).

Chen, Siwen, Kehan Li, Hongpeng Fu, Ying Cheng Wu, and Yiyi Huang (2023). "Sea ice extent prediction with machine learning methods and subregional analysis in the Arctic". In: *Atmosphere* 14.6, p. 1023 (cit. on p. 23).

Chen, Zhuo, Jun Sun, Ting Gu, Guicheng Zhang, and Yuqiu Wei (2021). "Nutrient ratios driven by vertical stratification regulate phytoplankton community structure in the oligotrophic western Pacific Ocean". In: *Ocean Science* 17.6, pp. 1775–1789 (cit. on p. 18).

Cheng, Kaichang, Xuemin Cheng, Yuqi Wang, Hongsheng Bi, and Mark C Benfield (2019). "Enhanced convolutional neural network for plankton identification and enumeration". In: *PLoS One* 14.7, e0219570 (cit. on p. 24).

Cohen, Jonathan H, Jørgen Berge, Mark A Moline, Geir Johnsen, and Artur P Zolich (2020). "Light in the polar night". In: *POLAR NIGHT marine ecology: life and light in the dead of night*, pp. 37–66 (cit. on p. 5).

Collins, R Eric, Shelly D Carpenter, and Jody W Deming (2008). "Spatial heterogeneity and temporal dynamics of particles, bacteria, and pEPS in Arctic winter sea ice". In: *Journal of Marine Systems* 74.3-4, pp. 902–917 (cit. on p. 9).

Collins, R Eric and Jody W Deming (2011). "Abundant dissolved genetic material in Arctic sea ice Part I: Extracellular DNA". In: *Polar biology* 34, pp. 1819–1830 (cit. on p. 8).

Collins, R Eric, Gabrielle Rocap, and Jody W Deming (2010). "Persistence of bacterial and archaeal communities in sea ice through an Arctic winter". In: *Environmental microbiology* 12.7, pp. 1828–1841 (cit. on p. 9).

Comeau, André M, William KW Li, Jean-Éric Tremblay, Eddy C Carmack, and Connie Lovejoy (2011). "Arctic Ocean microbial community structure before and after the 2007 record sea ice minimum". In: *PloS one* 6.11, e27492 (cit. on pp. 9, 18, 21).

Comiso, Josefino C (2003). "Warming trends in the Arctic from clear sky satellite observations". In: *Journal of Climate* 16.21, pp. 3498–3510 (cit. on p. 5).

Comiso, Josefino C, Claire L Parkinson, Robert Gersten, and Larry Stock (2008). "Accelerated decline in the Arctic sea ice cover". In: *Geophysical research letters* 35.1 (cit. on p. 5).

Costa, Ohana YA, Jos M Raaijmakers, and Eiko E Kuramae (2018). "Microbial extracellular polymeric substances: ecological function and impact on soil aggregation". In: *Frontiers in microbiology* 9, p. 337094 (cit. on p. 8).

Cowie, Rebecca OM, Gareth J Williams, Elizabeth W Maas, K Matt Voyles, and Ken G Ryan (2014). "Antarctic sea-ice microbial communities show distinct patterns of zonation in response to algal-derived substrates". In: *Aquatic Microbial Ecology* 73.2, pp. 123–134 (cit. on p. 9).

Crowley, Thomas J (2000). "Causes of climate change over the past 1000 years". In: *Science* 289.5477, pp. 270–277 (cit. on p. 4).

Crump, Byron C, George W Kling, Michele Bahr, and John E Hobbie (2003). "Bacterioplankton community shifts in an arctic lake correlate with seasonal changes in organic matter source". In: *Applied and Environmental Microbiology* 69.4, pp. 2253–2268 (cit. on p. 21).

Dai, Aiguo, Dehai Luo, Mirong Song, and Jiping Liu (2019). "Arctic amplification is caused by sea-ice loss under increasing CO2". In: *Nature communications* 10.1, p. 121 (cit. on p. 15).

Deal, Clara, Meibing Jin, Scott Elliott, Elizabeth Hunke, Mathew Maltrud, and Nicole Jeffery (2011). "Large-scale modeling of primary production and ice algal biomass within arctic sea ice in 1992". In: *Journal of Geophysical Research: Oceans* 116.C7 (cit. on p. 8).

Decho, Alan W and Tony Gutierrez (2017). "Microbial extracellular polymeric substances (EPSs) in ocean systems". In: *Frontiers in microbiology* 8, p. 265214 (cit. on pp. 8, 9).

DeLong, Edward F and Norman R Pace (2001). "Environmental diversity of bacteria and archaea". In: *Systematic biology* 50.4, pp. 470–478 (cit. on p. 19).

Deming, Jody W (2010). "Sea ice bacteria and viruses 2nd". In: *Sea ice*, pp. 247–282 (cit. on p. 9).

Deming, JW and H Eicken (2007). *Life in Ice. Planets and Life: The Emerging Science of Astrobiology. JA Baross and WTS III* (cit. on p. 9).

Deming, JW, C Gerday, and N Glansdorff (2007). *Physiology and Biochemistry of Extremophiles* (cit. on pp. 8, 9).

Deslippe, Julie R, Martin Hartmann, Suzanne W Simard, and William W Mohn (2012). "Long-term warming alters the composition of Arctic soil microbial communities". In: *FEMS microbiology ecology* 82.2, pp. 303–315 (cit. on p. 21).

Diebold, Francis X and Glenn D Rudebusch (2022). "Probability assessments of an ice-free Arctic: Comparing statistical and climate model projections". In: *Journal of Econometrics* 231.2, pp. 520–534 (cit. on p. 3).

Dieckmann, Gerhard S. and Hartmut H. Hellmer (n.d.). "The Importance of Sea Ice: An Overview". In: *Sea Ice*. John Wiley & Sons, Ltd. ISBN: 9781444317145. DOI: https://doi.org/10.1002/9781444317145. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781444317145. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781444317145 (cit. on p. 8).

Dittmar, Thorsten and Gerhard Kattner (2003). "The biogeochemistry of the river and shelf ecosystem of the Arctic Ocean: a review". In: *Marine chemistry* 83.3-4, pp. 103–120 (cit. on p. 1).

Doherty, Stacey Jarvis, Robyn A Barbato, A Stuart Grandy, W Kelley Thomas, Sylvain Monteux, Ellen Dorrepaal, Margareta Johansson, and Jessica G Ernakovich (2020). "The transition from stochastic to deterministic bacterial community assembly during permafrost thaw succession". In: *Frontiers in Microbiology* 11, p. 596589 (cit. on p. 22).

Driscoll, Simon, Alessio Bozzo, Lesley J Gray, Alan Robock, and Georgiy Stenchikov (2012). "Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions". In: *Journal of Geophysical Research: Atmospheres* 117.D17 (cit. on p. 3).

Ducklow, Hugh W, Scott C Doney, and Deborah K Steinberg (2009). "Contributions of long-term research and time-series observations to marine ecology and biogeochemistry". In: *Annual Review of Marine Science* 1, pp. 279–302 (cit. on p. 21).

Dupont, Frédéric (2012). "Impact of sea-ice biology on overall primary production in a biophysical model of the pan-Arctic Ocean". In: *Journal of Geophysical Research: Oceans* 117.C8 (cit. on p. 18).

Eerola, Tuomas, Daniel Batrakhanov, Nastaran Vatankhah Barazandeh, Kaisa Kraft, Lumi Haraguchi, Lasse Lensu, Sanna Suikkanen, Jukka Seppälä, Timo Tamminen, and Heikki Kälviäinen (2024). "Survey of automatic plankton image recognition: Challenges, existing solutions and future perspectives". In: *Artificial Intelligence Review* 57.5, p. 114 (cit. on p. 24).

Eikrem, Wenche, Linda K Medlin, Jorijntje Henderiks, Sebastian Rokitta, Björn Rost, Ian Probert, Jahn Throndsen, and Bente Edvardsen (2016). "Haptophyta". In: (cit. on p. 11).

Eilers, Heike, Jakob Pernthaler, Frank Oliver Glöckner, and Rudolf Amann (2000). "Culturability and in situ abundance of pelagic bacteria from the North Sea". In: *Applied and environmental microbiology* 66.7, pp. 3044–3051 (cit. on p. 19).

Ellertsen, Hans Chr (1993). "Spring blooms and stratification". In: *Nature* 363.6424, pp. 24–24 (cit. on pp. 5, 10).

Eme, Laura and Daniel Tamarit (2024). "Microbial Diversity and Open Questions about the Deep Tree of Life". In: *Genome Biology and Evolution* 16.4, evae053 (cit. on p. 19).

Eronen-Rasimus, Eeva, Anne-Mari Luhtanen, Janne-Markus Rintala, Bruno Delille, Gerhard Dieckmann, Antti Karkman, and Jean-Louis Tison (2017). "An active bacterial community linked to high chl-a concentrations in Antarctic winter-pack ice and evidence for the development of an anaerobic sea-ice bacterial community". In: *The ISME Journal* 11.10, pp. 2345–2355 (cit. on p. 9).

Eronen-Rasimus, Eeva, Christina Lyra, Janne-Markus Rintala, Klaus Jürgens, Vilma Ikonen, and Hermanni Kaartokallio (2015). "Ice formation and growth shape bacterial community structure in Baltic Sea drift ice". In: *FEMS microbiology ecology* 91.2, pp. 1–13 (cit. on p. 9).

Eronen-Rasimus, Eeva, Jonna Piiparinen, Antti Karkman, Christina Lyra, Sebastian Gerland, and Hermanni Kaartokallio (2016). "Bacterial communities in Arctic first-year drift ice during the winter/spring transition". In: *Environmental microbiology reports* 8.4, pp. 527–535 (cit. on p. 9).

Estrada, Veronica, Ellen Oldenburg, Ovidiu Popa, and Hans Werner Müller (2022). "Mapping the long rocky road to effective spinal cord injury therapy: a meta-review of pre-clinical and clinical research". In: *Journal of Neurotrauma* 39.9-10, pp. 591–612 (cit. on p. 145).

Ewert, Marcela and Jody W Deming (2013). "Sea ice microorganisms: Environmental constraints and extracellular responses". In: *Biology* 2.2, pp. 603–628 (cit. on p. 8).

Feng, Shi, Shane M Powell, Richard Wilson, and John P Bowman (2014). "Extensive gene acquisition in the extremely psychrophilic bacterial species Psychroflexus torquis and the link to sea-ice ecosystem specialism". In: *Genome biology and evolution* 6.1, pp. 133–148 (cit. on pp. 8, 9).

Fernández-Méndez, Mar, Christian Katlein, Benjamin Rabe, Marcel Nicolaus, Ilka Peeken, Karel Bakker, Hauke Flores, and Antje Boetius (2015). "Photosynthetic production in the central Arctic Ocean during the record sea-ice minimum in 2012". In: *Biogeosciences* 12.11, pp. 3525–3549 (cit. on p. 10).

Fernández-Méndez, Mar, Frank Wenzhöfer, Ilka Peeken, Heidi L Sørensen, Ronnie N Glud, and Antje Boetius (2014). "Composition, buoyancy regulation and fate of ice algal aggregates in the Central Arctic Ocean". In: *PLoS One* 9.9, e107452 (cit. on p. 8).

Fetterer, Florence and Norbert Untersteiner (1998). "Observations of melt ponds on Arctic sea ice". In: *Journal of Geophysical Research: Oceans* 103.C11, pp. 24821–24835 (cit. on p. 5).

Fong, Allison, Clara Hoppe, Nicole Aberle, Carin J Ashjian, Philipp Assmy, Youcheng Bai, Dorothee CE Bakker, Deborah Bozzato, Jacqueline Stefels, and Maria van Leeuwe (2023). "Overview of the MOSAiC expedition: Ecosystem". In: *EarthArXiv*, pp. 2023–01 (cit. on p. 145).

Foulon, Elodie, Fabrice Not, Fabienne Jalabert, Thierry Cariou, Ramon Massana, and Nathalie Simon (2008). "Ecological niche partitioning in the picoplanktonic green alga Micromonas pusilla: evidence from environmental surveys using phylogenetic probes". In: *Environmental microbiology* 10.9, pp. 2433–2443 (cit. on p. 11).

Franz, Jasmin (2012). "Effects of changing N: P supply on phytoplankton in the eastern tropical Pacific and Atlantic Ocean". PhD thesis. Christian-Albrechts-Universität Kiel (cit. on p. 18).

Fujita, Hiroaki, Masayuki Ushio, Kenta Suzuki, Masato S Abe, Masato Yamamichi, Koji Iwayama, Alberto Canarini, Ibuki Hayashi, Keitaro Fukushima, Shinji Fukuda, et al. (2023). "Alternative stable states, nonlinear behavior, and predictability of microbiome dynamics". In: *Microbiome* 11.1, p. 63 (cit. on p. 23).

Fulford, Richard S, Sheila JJ Heymans, and Wei Wu (2020). "Mathematical modeling for ecosystem-based management (EBM) and ecosystem goods and services (EGS) assessment". In: *Ecosystem-based management, ecosystem services and aquatic biodiversity: Theory, tools and applications*, pp. 275–289 (cit. on p. 23).

Gall, Andrea, Udo Uebel, Uwe Ebensen, Helmut Hillebrand, Sandra Meier, Gabriel Singer, Alexander Wacker, and Maren Striebel (2017). "Planktotrons: A novel indoor mesocosm facility for aquatic biodiversity and food web research". In: *Limnology and Oceanography: Methods* 15.7, pp. 663–677 (cit. on p. 143).

Gascard, Jean-Claude, Jean Festy, Hervé le Goff, Matthieu Weber, Burghard Bruemmer, Michael Offermann, Martin Doble, Peter Wadhams, René Forsberg, Susan Hanson, et al. (2008). "Exploring Arctic transpolar drift during dramatic sea ice retreat". In: *Eos, Transactions American Geophysical Union* 89.3, pp. 21–22 (cit. on p. 26).

Gasparini, Stéphane, Marie Hermande Daro, Elvire Antajan, Michele Tackx, Véronique Rousseau, J-Y Parent, and Christiane Lancelot (2000). "Mesozooplankton grazing during the Phaeocystis globosa bloom in the southern bight of the North Sea". In: *Journal of Sea Research* 43.3-4, pp. 345–356 (cit. on p. 11).

Golden, KM, SF Ackley, and VI Lytle (1998). "The percolation phase transition in sea ice". In: *Science* 282.5397, pp. 2238–2241 (cit. on p. 5).

Goldford, Joshua E, Nanxi Lu, Djordje Bajić, Sylvie Estrela, Mikhail Tikhonov, Alicia Sanchez-Gorostiaga, Daniel Segrè, Pankaj Mehta, and Alvaro Sanchez (2018). "Emergent simplicity in microbial community assembly". In: *Science* 361.6401, pp. 469–474 (cit. on p. 142).

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016). *Deep learning*. MIT press (cit. on p. 24).

Gosink, John J, Roar L Irgens, and James T Staley (1993). "Vertical distribution of bacteria in arctic sea ice". In: *FEMS microbiology ecology* 11.2, pp. 85–90 (cit. on p. 8).

Gosselin, Michel, Maurice Levasseur, Patricia A Wheeler, Rita A Horner, and Beatrice C Booth (1997). "New measurements of phytoplankton and ice algal production in the Arctic Ocean". In: *Deep Sea Research Part II: Topical Studies in Oceanography* 44.8, pp. 1623–1644 (cit. on p. 10).

Gradinger, R and J Ikävalko (1998). "Organism incorporation into newly forming Arctic sea ice in the Greenland Sea". In: *Journal of Plankton Research* 20.5, pp. 871–886 (cit. on p. 8).

Gradinger, Rolf (2009). "Sea-ice algae: Major contributors to primary production and algal biomass in the Chukchi and Beaufort Seas during May/June 2002". In: *Deep Sea Research Part II: Topical Studies in Oceanography* 56.17, pp. 1201–1212 (cit. on p. 10).

Grebmeier, Jacqueline M, James E Overland, Sue E Moore, Ed V Farley, Eddy C Carmack, Lee W Cooper, Karen E Frey, John H Helle, Fiona A McLaughlin, and S Lyn McNutt (2006). "A major ecosystem shift in the northern Bering Sea". In: *Science* 311.5766, pp. 1461–1464 (cit. on p. 18).

Gustafson Jr, Daniel E, Diane K Stoecker, Matthew D Johnson, William F Van Heukelem, and Kerri Sneider (2000). "Cryptophyte algae are robbed of their organelles by the marine ciliate Mesodinium rubrum". In: *Nature* 405.6790, pp. 1049–1052 (cit. on p. 11).

Halvorsen, MH, LH Smedsrud, R Zhang, and K Kloster (2015). "Fram Strait spring ice export and September Arctic sea ice". In: *The Cryosphere Discussions* 9.4, pp. 4205–4235 (cit. on p. 5).

Hamdan, Leila J, Richard B Coffin, Masoumeh Sikaroodi, Jens Greinert, Tina Treude, and Patrick M Gillevet (2013). "Ocean currents shape the microbiome of Arctic marine sediments". In: *The ISME journal* 7.4, pp. 685–696 (cit. on p. 6).

Han, Dukki, Ilnam Kang, Ho Kyung Ha, Hyun Cheol Kim, Ok-Sun Kim, Bang Yong Lee, Jang-Cheon Cho, Hor-Gil Hur, and Yoo Kyung Lee (2014). "Bacterial communities of surface mixed layer in the Pacific sector of the western Arctic Ocean during sea-ice melting". In: *PloS one* 9.1, e86887 (cit. on p. 9).

Hartmann, Dennis L, Albert MG Klein Tank, Matilde Rusticucci, Lisa V Alexander, Stefan Brönnimann, Yassine Abdul Rahman Charabi, Frank J Dentener, Edward J Dlugokencky, David R Easterling, Alexey Kaplan, et al. (2013). "Observations: atmosphere and surface".

In: *Climate change 2013 the physical science basis: Working group I contribution to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, pp. 159–254 (cit. on p. 4).

Hatam, Ido, Rhianna Charchuk, Benjamin Lange, Justin Beckers, Christian Haas, and Brian Lanoil (2014). "Distinct bacterial assemblages reside at different depths in Arctic multiyear sea ice". In: *FEMS microbiology ecology* 90.1, pp. 115–125 (cit. on p. 9).

Hatam, Ido, Benjamin Lange, Justin Beckers, Christian Haas, and Brian Lanoil (2016). "Bacterial communities from Arctic seasonal sea ice are more compositionally variable than those from multi-year sea ice". In: *The ISME Journal* 10.10, pp. 2543–2552 (cit. on p. 9).

Heinrichs, Mara Elena, Goncalo J Piedade, Ovidiu Popa, Pacifica Sommers, Gareth Trubl, Julia Weissenbach, and Janina Rahlff (2023). "Breaking the Ice: A Review of Phages in Polar Ecosystems". In: *Bacteriophages: Methods and Protocols*, pp. 31–71 (cit. on p. 141).

Heinz, Jacob, Janosch Schirmack, Alessandro Airo, Samuel P Kounaves, and Dirk Schulze-Makuch (2018). "Enhanced microbial survivability in subzero brines". In: *Astrobiology* 18.9, pp. 1171–1180 (cit. on p. 9).

Holt, Jason, Corinna Schrum, Heather Cannaby, Ute Daewel, Icarus Allen, Yuri Artioli, Laurent Bopp, Momme Butenschon, Bettina A Fach, James Harle, et al. (2016). "Potential impacts of climate change on the primary production of regional seas: A comparative analysis of five European seas". In: *Progress in Oceanography* 140, pp. 91–115 (cit. on p. 6).

Howes, Ella L, Fortunat Joos, C Mark Eakin, and Jean-Pierre Gattuso (2015). "An updated synthesis of the observed and projected impacts of climate change on the chemical, physical and biological processes in the oceans". In: *Frontiers in Marine Science* 2, p. 36 (cit. on p. 4).

Hugenholtz, Philip and Norman R Pace (1996). "Identifying microbial diversity in the natural environment: a molecular phylogenetic approach". In: *Trends in biotechnology* 14.6, pp. 190–197 (cit. on p. 19).

Huse, Susan M, Les Dethlefsen, Julie A Huber, David Mark Welch, David A Relman, and Mitchell L Sogin (2008). "Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing". In: *PLoS genetics* 4.11, e1000255 (cit. on p. 20).

Ibarbalz, Federico M, Mariéa Victoria Pérez, Eva LM Figuerola, and Leonardo Erijman (2014). "The bias associated with amplicon sequencing does not affect the quantitative assessment of bacterial community dynamics". In: *PLoS One* 9.6, e99722 (cit. on p. 20).

Jacob, M. (2014). *Influence of Global Change on Microbial Communities in Arctic Sediments*. Accessed March 18, 2018. URL: `https://elib.suub.uni-bremen.de/edocs/00103916-1.pdf` (cit. on p. 19).

Jahn, Alexandra, Marika M Holland, and Jennifer E Kay (2024). "Projections of an ice-free Arctic Ocean". In: *Nature Reviews Earth & Environment*, pp. 1–13 (cit. on p. 3).

Jahnke, Richard A and George A Jackson (1992). "The spatial distribution of sea floor oxygen consumption in the Atlantic and Pacific Oceans". In: *Deep-sea food chains and the global carbon cycle*. Springer, pp. 295–307 (cit. on p. 5).

Jain, Ankita and Varsha Tailor (2020). "Emerging Trends of Biotechnology in Marine Bioprospecting: A New Vision". In: *Marine Niche: Applications in Pharmaceutical Sciences: Translational Research*, pp. 1–36 (cit. on p. 20).

Jakobsson, Martin (2002). "Hypsometry and volume of the Arctic Ocean and its constituent seas". In: *Geochemistry, Geophysics, Geosystems* 3.5, pp. 1–18 (cit. on p. 1).

Jakobsson, Martin, Arthur Grantz, Yngve Kristoffersen, and Ron Macnab (2003). "Physiographic provinces of the Arctic Ocean seafloor". In: *Geological Society of America Bulletin* 115.12, pp. 1443–1455 (cit. on p. 1).

Jakobsson, Martin, Arthur Grantz, Yngve Kristoffersen, Ron Macnab, Robie W MacDonald, Egil Sakshaug, Ruediger Stein, and Wilfried Jokat (2004). "The Arctic Ocean: Boundary conditions and background information". In: *The organic carbon cycle in the Arctic Ocean*. Springer, pp. 1–32 (cit. on pp. 1, 6).

Jakobsson, Martin, Olafur Ingolfsson, Antony J Long, and Robert F Spielhagen (2014). "The dynamic Arctic". In: *Quaternary Science Reviews* 92, pp. 1–8 (cit. on p. 1).

Javier, Prince Joseph Erneszer A, Marissa P Liponhay, Carlo Vincienzo G Dajac, and Christopher P Monterola (2022). "Causal network inference in a dam system and its implications on feature selection for machine learning forecasting". In: *Physica A: Statistical Mechanics and its Applications* 604, p. 127893 (cit. on p. 22).

Jean-Michel, Lellouche, Greiner Eric, Bourdallé-Badie Romain, Garric Gilles, Melet Angélique, Drévillon Marie, Bricaud Clément, Hamon Mathieu, Le Galloudec Olivier, Regnier Charly, et al. (2021). "The Copernicus global 1/12 oceanic and sea ice GLORYS12 reanalysis". In: *Frontiers in Earth Science* 9, p. 698876 (cit. on p. 15).

Jensen, Frank and Benni Winding Hansen (2000). "Ciliates and heterotrophic dinoflagellates in the marginal ice zone of the central Barents Sea during spring". In: *Journal of the Marine Biological Association of the United Kingdom* 80.1, pp. 45–54 (cit. on p. 11).

Ji, Zhiwei, Ke Yan, Wenyang Li, Haigen Hu, Xiaoliang Zhu, et al. (2017). "Mathematical and computational modeling in complex biological systems". In: *BioMed research international* 2017 (cit. on p. 23).

Jones, E Peter (2001). "Circulation in the arctic ocean". In: *Polar Research* 20.2, pp. 139–146 (cit. on p. 6).

Jose, Jonath (Aug. 2022). "INTRODUCTION TO TIME SERIES ANALYSIS AND ITS APPLICATIONS". In: URL: https://www.researchgate.net/publication/362389180_INTRODUCTION_TO_TIME_SERIES_ANALYSIS_AND_ITS_APPLICATIONS (cit. on p. 21).

Joseph, Tyler A, Liat Shenhav, Joao B Xavier, Eran Halperin, and Itsik Pe'er (2020). "Compositional Lotka-Volterra describes microbial dynamics in the simplex". In: *PLoS computational biology* 16.5, e1007917 (cit. on p. 142).

Junge, Karen, Brent Christner, and James T Staley (2011). "Diversity of psychrophilic bacteria from sea ice-and glacial ice communities". In: *Extremophiles handbook* (cit. on p. 9).

Junge, Karen, Hajo Eicken, and Jody W Deming (2004). "Bacterial activity at- 2 to- 20 C in Arctic wintertime sea ice". In: *Applied and Environmental Microbiology* 70.1, pp. 550–557 (cit. on p. 9).

Kahru, M, V Brotas, M Manzano, and BG Mitchell (2010). "Changes in the Timing of Phytoplankton Blooms Related to Diminished Ice Cover in the Arctic". In: *AGU Fall Meeting Abstracts*. Vol. 2010, C52B–05 (cit. on p. 18).

Kanzow, T. (2021). "(2023): The Expedition PS131 of the Research Vessel POLARSTERN to the Fram Strait in 2022 / H". In: (cit. on pp. 2, 183).

Karam, Salar, Céline Heuzé, Mario Hoppmann, and Laura de Steur (2024). "Continued warming of deep waters in Fram Strait". In: *EGUsphere* 2024, pp. 1–23 (cit. on p. 4).

Katelaris, Constance H and Paul J Beggs (2018). "Climate change: allergens and allergic diseases". In: *Internal medicine journal* 48.2, pp. 129–134 (cit. on p. 4).

Katlein, Christian, Mar Fernández-Méndez, Frank Wenzhöfer, and Marcel Nicolaus (2015). "Distribution of algal aggregates under summer sea ice in the Central Arctic". In: *Polar biology* 38, pp. 719–731 (cit. on p. 8).

Khilyuk, Leonid (2003). "Global warming: are we confusing cause and effect?" In: *Energy Sources* 25.4, pp. 357–370 (cit. on p. 4).

Kim, Yeon-Hee, Seung-Ki Min, Nathan P Gillett, Dirk Notz, and Elizaveta Malinina (2023). "Observationally-constrained projections of an ice-free Arctic even under a low emission scenario". In: *Nature Communications* 14.1, p. 3139 (cit. on pp. 3, 5).

Kinnard, Christophe, Christian M Zdanowicz, David A Fisher, Elisabeth Isaksson, Anne de Vernal, and Lonnie G Thompson (2011). "Reconstructed changes in Arctic sea ice over the past 1,450 years". In: *Nature* 479.7374, pp. 509–512 (cit. on p. 3).

Kirchman, David L, Matthew T Cottrell, and Connie Lovejoy (2010). "The structure of bacterial communities in the western Arctic Ocean as revealed by pyrosequencing of 16S rRNA genes". In: *Environmental microbiology* 12.5, pp. 1132–1143 (cit. on p. 7).

Kirtman, Ben, Scott B Power, Akintayo John Adedoyin, George J Boer, Roxana Bojariu, Ines Camilloni, Francisco Doblas-Reyes, Arlene M Fiore, Masahide Kimoto, Gerald Meehl, et al. (2013). "Near-term climate change: projections and predictability". In: (cit. on p. 5).

Kohlbach, Doreen (2017). "The role of sea ice algae-produced carbon in Arctic and Antarctic food webs: dependency of polar life on a threatened food source". PhD thesis. Staats-und Universitätsbibliothek Hamburg Carl von Ossietzky (cit. on pp. 10, 18).

Korhonen, Meri, Bert Rudels, M Marnela, A Wisotzki, and J Zhao (2013). "Time and space variability of freshwater content, heat content and seasonal ice melt in the Arctic Ocean from 1991 to 2011". In: *Ocean Science* 9.6, pp. 1015–1055 (cit. on pp. 3, 18).

Kottmeier, Steven T and Cornelius W Sullivan (1990). "Bacterial biomass and production in pack ice of Antarctic marginal ice edge zones". In: *Deep Sea Research Part A. Oceanographic Research Papers* 37.8, pp. 1311–1330 (cit. on p. 9).

Kraemer, Susanne A, Arthi Ramachandran, Vera E Onana, William KW Li, and David A Walsh (2024). "A multiyear time series (2004–2012) of bacterial and archaeal community dynamics in a changing Arctic Ocean". In: *ISME communications* 4.1, ycad004 (cit. on p. 21).

Krembs, C, R Gradinger, and M Spindler (2000). "Implications of brine channel geometry and surface area for the interaction of sympagic organisms in Arctic sea ice". In: *Journal of Experimental Marine Biology and Ecology* 243.1, pp. 55–80 (cit. on p. 8).

Krembs, Christopher, Hajo Eicken, and Jody W Deming (2011). "Exopolymer alteration of physical properties of sea ice and implications for ice habitability and biogeochemistry in a warmer Arctic". In: *Proceedings of the National Academy of Sciences* 108.9, pp. 3653–3658 (cit. on pp. 5, 9).

Krembs, Christopher, Hajo Eicken, Karen Junge, and JW Deming (2002). "High concentrations of exopolymeric substances in Arctic winter sea ice: implications for the polar ocean carbon cycle and cryoprotection of diatoms". In: *Deep Sea Research Part I: Oceanographic Research Papers* 49.12, pp. 2163–2181 (cit. on p. 8).

Krueger, Oliver and Jin-Song Von Storch (2011). "A simple empirical model for decadal climate prediction". In: *Journal of climate* 24.4, pp. 1276–1283 (cit. on p. 21).

Krumpen, T, R Gerdes, C Haas, S Hendricks, A Herber, V Selyuzhenok, L Smedsrud, and G Spreen (2015). "Recent summer sea ice thickness surveys in the Fram Strait and associated volume fluxes." In: *Cryosphere Discussions* 9.5 (cit. on p. 5).

Kudryavtseva, Elena, Marina Kravchishina, Larisa Pautova, Igor Rusanov, Dmitry Glukhovets, Alexander Shchuka, Ivan Zamyatin, Nadezhda Torgunova, Anna Chultsova, Nadezhda Politova, et al. (2023). "Sea Ice as a Factor of Primary Production in the European Arctic:

Phytoplankton Size Classes and Carbon Fluxes". In: *Journal of Marine Science and Engineering* 11.11, p. 2131 (cit. on p. 10).

Kumar, Rajnish, Garima Yadav, Mohammed Kuddus, Ghulam Md Ashraf, and Rachana Singh (2023). "Unlocking the microbial studies through computational approaches: how far have we reached?" In: *Environmental Science and Pollution Research* 30.17, pp. 48929–48947 (cit. on p. 24).

Kwok, Ron (2007). "Near zero replenishment of the Arctic multiyear sea ice cover at the end of 2005 summer". In: *Geophysical Research Letters* 34.5 (cit. on p. 15).

Kwok, Rothrock and DA Rothrock (2009). "Decline in Arctic sea ice thickness from submarine and ICESat records: 1958–2008". In: *Geophysical Research Letters* 36.15 (cit. on pp. 5, 7).

Kyewalyanga, Margareth (2016). "Phytoplankton primary production". In: *The regional state of the coast report: Western Indian Ocean. Nairobi: UNEP-Nairobi Convention*, pp. 213–230 (cit. on p. 13).

Lane, D. J., B. Pace, G. J. Olsen, D. A. Stahl, M. L. Sogin, and N. R. Pace (1985). "Rapid Determination of 16S Ribosomal RNA Sequences for Phylogenetic Analyses". In: *Proc. Natl. Acad. Sci. U. S. A.* 82. Accessed July 4, 2018, pp. 6955–6959. URL: `http://www.ncbi.nlm.nih.gov/pubmed/2413450` (cit. on p. 19).

Laney, Samuel R, Richard A Krishfield, John M Toole, Terence R Hammar, Carin J Ashjian, and Mary-Louise Timmermans (2014). "Assessing algal biomass and bio-optical distributions in perennially ice-covered polar ocean ecosystems". In: *Polar Science* 8.2, pp. 73–85 (cit. on p. 6).

Larose, Catherine, Aurélien Dommergue, and Timothy M Vogel (2013). "The dynamic arctic snow pack: an unexplored environment for microbial diversity and activity". In: *Biology* 2.1, pp. 317–330 (cit. on p. 9).

Lawton, John H (1988). "More time means more variation". In: *Nature* 334.6183, pp. 563–563 (cit. on p. 21).

Laxon, Seymour W, Katharine A Giles, Andy L Ridout, Duncan J Wingham, Rosemary Willatt, Robert Cullen, Ron Kwok, Axel Schweiger, Jinlun Zhang, Christian Haas, et al. (2013). "CryoSat-2 estimates of Arctic sea ice thickness and volume". In: *Geophysical Research Letters* 40.4, pp. 732–737 (cit. on pp. 5, 15).

Le Hir, G, Yoram Teitler, F Fluteau, Yannick Donnadieu, and P Philippot (2014). "The faint young Sun problem revisited with a 3-D climate–carbon model–Part 1". In: *Climate of the Past* 10.2, pp. 697–713 (cit. on p. 5).

Lean, Judith L (2017). "Sun-climate connections". In: *Oxford Research Encyclopedia of Climate Science* (cit. on p. 13).

Lebrato, Mario, Markus Pahlow, Jessica R Frost, Marie Küter, Pedro de Jesus Mendes, Juan-Carlos Molinero, and Andreas Oschlies (2019). "Sinking of gelatinous zooplankton biomass increases deep carbon transfer efficiency globally". In: *Global Biogeochemical Cycles* 33.12, pp. 1764–1783 (cit. on p. 13).

Lee, Kyung Ha, Hae Jin Jeong, Tae Young Jang, An Suk Lim, Nam Seon Kang, Ju-Hyoung Kim, Kwang Young Kim, Ki-Tae Park, and Kitack Lee (2014). "Feeding by the newly described mixotrophic dinoflagellate Gymnodinium smaydae: feeding mechanism, prey species, and effect of prey concentration". In: *Journal of experimental marine biology and ecology* 459, pp. 114–125 (cit. on p. 11).

Legendre, Louis, Stephen F Ackley, Gerhard S Dieckmann, Bjørn Gulliksen, Rita Horner, Takao Hoshiai, Igor A Melnikov, William S Reeburgh, Michael Spindler, and Cornelius W Sullivan (1992). "Ecology of sea ice biota: 2. Global significance". In: *Polar biology* 12, pp. 429–444 (cit. on p. 8).

Lellouche, Jean-Michel, Eric Greiner, Olivier Le Galloudec, Gilles Garric, Charly Regnier, Marie Drevillon, Mounir Benkiran, Charles-Emmanuel Testut, Romain Bourdalle-Badie, Florent Gasparin, et al. (2018). "Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1/12° high-resolution system". In: *Ocean Science* 14.5, pp. 1093–1126 (cit. on p. 15).

Li, Ming, Yuren Chen, Fan Zhang, Yang Song, Patricia M Glibert, and Diane K Stoecker (2022). "A three-dimensional mixotrophic model of Karlodinium veneficum blooms for a eutrophic estuary". In: *Harmful Algae* 113, p. 102203 (cit. on p. 11).

Li, William KW, Fiona A McLaughlin, Connie Lovejoy, and Eddy C Carmack (2009). "Smallest algae thrive as the Arctic Ocean freshens". In: *Science* 326.5952, pp. 539–539 (cit. on p. 18).

Lightbody, Gaye, Valeriia Haberland, Fiona Browne, Laura Taggart, Huiru Zheng, Eileen Parkes, and Jaine K Blayney (2019). "Review of applications of high-throughput sequencing in personalized medicine: barriers and facilitators of future progress in research and clinical application". In: *Briefings in bioinformatics* 20.5, pp. 1795–1811 (cit. on p. 20).

Lindsey, Rebecca (2009). "Climate and earth's energy budget". In: *NASA Earth Observatory* 680 (cit. on p. 3).

Lomartire, Silvia, João C Marques, and Ana MM Gonçalves (2021). "The key role of zooplankton in ecosystem services: A perspective of interaction between zooplankton and fish recruitment". In: *Ecological Indicators* 129, p. 107867 (cit. on p. 13).

Long, Zhenxia, Will Perrie, Minghong Zhang, and Yazhou Liu (2024). "Responses of Atlantic water inflow through Fram Strait to Arctic storms". In: *Geophysical Research Letters* 51.6, e2023GL107777 (cit. on p. 4).

Lotka, Alfred J (1920). "Analytical note on certain rhythmic relations in organic systems". In: *Proceedings of the National Academy of Sciences* 6.7, pp. 410–415 (cit. on p. 142).

Lovejoy, Connie, Warwick F Vincent, Sylvia Bonilla, Suzanne Roy, Marie-Josée Martineau, Ramon Terrado, Marianne Potvin, Ramon Massana, and Carlos Pedrós-Alió (2007). "DISTRIBUTION, PHYLOGENY, AND GROWTH OF COLD-ADAPTED PICOPRASINOPHYTES IN ARCTIC SEAS 1". In: *Journal of Phycology* 43.1, pp. 78–89 (cit. on p. 11).

Luo, Jessica Y, Jean-Olivier Irisson, Benjamin Graham, Cedric Guigand, Amin Sarafraz, Christopher Mader, and Robert K Cowen (2018). "Automated plankton image analysis using convolutional neural networks". In: *Limnology and Oceanography: methods* 16.12, pp. 814–827 (cit. on p. 24).

Macdonald, Robie W, T Harner, and J Fyfe (2005). "Recent climate change in the Arctic and its impact on contaminant pathways and interpretation of temporal trend data". In: *Science of the total environment* 342.1-3, pp. 5–86 (cit. on p. 21).

MacKay, David JC (2003). *Information theory, inference and learning algorithms*. Cambridge university press (cit. on p. 142).

Mahlstein, Irina and Reto Knutti (2012). "September Arctic sea ice predicted to disappear near 2 C global warming above present". In: *Journal of Geophysical Research: Atmospheres* 117.D6 (cit. on p. 15).

Malcai, Ofer, Ofer Biham, Peter Richmond, and Sorin Solomon (2002). "Theoretical analysis and simulations of the generalized Lotka-Volterra model". In: *Physical Review E* 66.3, p. 031102 (cit. on p. 142).

Manzoni, Claudia, Demis A Kia, Jana Vandrovcova, John Hardy, Nicholas W Wood, Patrick A Lewis, and Raffaele Ferrari (2018). "Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences". In: *Briefings in bioinformatics* 19.2, pp. 286–302 (cit. on p. 20).

Marsland III, Robert, Wenping Cui, Joshua Goldford, Alvaro Sanchez, Kirill Korolev, and Pankaj Mehta (2019). "Available energy fluxes drive a transition in the diversity, stability, and functional structure of microbial communities". In: *PLoS computational biology* 15.2, e1006793 (cit. on p. 143).

Masenya, Kedibone, Madira Coutlyne Manganyi, and Tshegofatso Bridget Dikobe (2024). "Exploring Cereal Metagenomics: Unravelling Microbial Communities for Improved Food Security". In: *Microorganisms* 12.3, p. 510 (cit. on p. 20).

Maslanik, JA, C Fowler, J Stroeve, S Drobot, J Zwally, D Yi, and W Emery (2007). "A younger, thinner Arctic ice cover: Increased potential for rapid, extensive sea-ice loss". In: *Geophysical Research Letters* 34.24 (cit. on p. 15).

Masson-Delmotte, V., M. Schulz, A. Abe-Ouchi, J. Beer, J. Ganopolski, J. F. González Rouco, E. Jansen, K. Lambeck, J. Luterbacher, T. Naish, T. Osborn, B. Otto-Bliesner, T. Quinn, R. Ramesh, M. Rojas, X. Shao, and A. Timmermann (2013). "Information from paleoclimate archives". In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Doschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley. Cambridge, UK: Cambridge University Press, pp. 383–464. DOI: 10.1017/CBO9781107415324.013 (cit. on p. 4).

Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (2021a). "Summary for Policymakers". In: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou. Figure SPM.8. Cambridge, UK and New York, NY, USA: Cambridge University Press, pp. 3–32. DOI: 10.1017/9781009157896.001 (cit. on pp. 17, 184).

Masson-Delmotte, Valérie, Panmao Zhai, Anna Pirani, Sarah L Connors, Clotilde Péan, Sophie Berger, Nada Caud, Y Chen, L Goldfarb, MI Gomis, et al. (2021b). "Climate change 2021: the physical science basis". In: *Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change* 2.1, p. 2391 (cit. on p. 3).

Matantseva, OV and SO Skarlato (2013). "Mixotrophy in microorganisms: ecological and cytophysiological aspects". In: *Journal of Evolutionary Biochemistry and Physiology* 49, pp. 377–388 (cit. on p. 11).

Mazur, James E (2006). "Mathematical models and the experimental analysis of behavior". In: *Journal of the experimental analysis of behavior* 85.2, pp. 275–291 (cit. on p. 23).

McLaughlin, Fiona A and Eddy C Carmack (2010). "Deepening of the nutricline and chlorophyll maximum in the Canada Basin interior, 2003–2009". In: *Geophysical research letters* 37.24 (cit. on p. 18).

McLaughlin, Fiona A, Eddy C Carmack, William J Williams, Sarah Zimmermann, Koji Shimada, and Motoyo Itoh (2009). "Joint effects of boundary currents and thermohaline intrusions on the warming of Atlantic water in the Canada Basin, 1993–2007". In: *Journal of Geophysical Research: Oceans* 114.C1 (cit. on p. 15).

Meier, Walter N, Greta K Hovelsrud, Bob EH Van Oort, Jeffrey R Key, Kit M Kovacs, Christine Michel, Christian Haas, Mats A Granskog, Sebastian Gerland, Donald K Perovich, et al. (2014). "Arctic sea ice in transformation: A review of recent observed changes and impacts on biology and human activity". In: *Reviews of Geophysics* 52.3, pp. 185–217 (cit. on p. 9).

Meiners, Klaus, Robin Brinkmeyer, Mats A Granskog, and Antti Lindfors (2004). "Abundance, size distribution and bacterial colonization of exopolymer particles in Antarctic sea ice (Bellingshausen Sea)". In: *Aquatic Microbial Ecology* 35.3, pp. 283–296 (cit. on p. 9).

Melnikov, Igor A (1987). "Ecology of mass accumulations of colonial diatom algae under drifting Arctic ice". In: *Oceanology* 27, p. 233 (cit. on p. 8).

Meredith, Michael, Martin Sommerkorn, S Cassota, Chris Derksen, Alexey Ekaykin, Anne Hollowed, Gary Kofinas, Andrew Mackintosh, Jessica Melbourne-Thomas, MMC Muelbert, et al. (2019). "Polar regions". In: (cit. on pp. 3, 5, 140).

Metfies, Katja, Wilken-Jon von Appen, Estelle Kilias, Anja Nicolaus, and Eva-Maria Nöthig (2016). "Biogeography and photosynthetic biomass of arctic marine pico-eukaroytes during summer of the record sea ice minimum 2012". In: *PLoS One* 11.2, e0148512 (cit. on p. 11).

Miettinen, Arto, Nalan Koç, and Katrine Husum (2013). "Appearance of the Pacific diatom Neodenticula seminae in the northern Nordic Seas—an indication of changes in Arctic sea ice and ocean circulation". In: *Marine Micropaleontology* 99, pp. 2–7 (cit. on p. 18).

Milanese, Alessio, Daniel R Mende, Lucas Paoli, Guillem Salazar, Hans-Joachim Ruscheweyh, Miguelangel Cuenca, Pascal Hingamp, Renato Alves, Paul I Costea, Luis Pedro Coelho, et al. (2019). "Microbial abundance, activity and population genomic profiling with mOTUs2". In: *Nature communications* 10.1, p. 1014 (cit. on p. 20).

Mitra, Aditee, Kevin J Flynn, Urban Tillmann, John A Raven, David Caron, Diane K Stoecker, Fabrice Not, Per J Hansen, Gustaaf Hallegraeff, Robert Sanders, et al. (2016). "Defining planktonic protist functional groups on mechanisms for energy and nutrient acquisition: incorporation of diverse mixotrophic strategies". In: *Protist* 167.2, pp. 106–120 (cit. on p. 11).

Mock, Thomas, William Boulton, John-Paul Balmonte, Kevin Barry, Stefan Bertilsson, Jeff Bowman, Moritz Buck, Gunnar Bratbak, Emelia J. Chamberlain, Michael Cunliffe, Jessie Creamean, Oliver Ebenhöh, Sarah Lena Eggers, Allison A. Fong, Jessie Gardner, Rolf Gradinger, Mats A. Granskog, Charlotte Havermans, Thomas Hill, Clara J. M. Hoppe, Kerstin Korte, Aud Larsen, Oliver Müller, Anja Nicolaus, Ellen Oldenburg, Ovidiu Popa, Swantje Rogge, Hendrik Schäfer, Katyanne Shoemaker, Pauline Snoeijs-Leijonmalm, Anders Torstensson, Klaus Valentin, Anna Vader, Kerrie Barry, I.-M. A. Chen, Alicia Clum, Alex Copeland, Chris Daum, Emiley Eloe-Fadrosh, Brian Foster, Bryce Foster, Igor V. Grigoriev, Marcel Huntemann, Natalia Ivanova, Alan Kuo, Nikos C. Kyrpides, Supratim Mukherjee, Krishnaveni Palaniappan, T. B. K. Reddy, Asaf Salamov, Simon Roux, Neha Varghese, Tanja Woyke, Dongying Wu, Richard M. Leggett, Vincent Moulton, and Katja Metfies (Oct. 2022). "Multiomics in the central Arctic Ocean for benchmarking biodiversity change". In: *PLOS Biology* 20.10, pp. 1–6. DOI: 10.1371/journal.pbio.3001835. URL: https://doi.org/10.1371/journal.pbio.3001835 (cit. on pp. 2, 26, 141, 145, 183).

Mock, Thomas, Stuart J Daines, Richard Geider, Sinead Collins, Metodi Metodiev, Andrew J Millar, Vincent Moulton, and Timothy M Lenton (2016). "Bridging the gap between omics

and earth system science to better understand how environmental change impacts marine microbes". In: *Global Change Biology* 22.1, pp. 61–75 (cit. on p. 22).

Mock, Thomas and Rolf Gradinger (1999). "Determination of Arctic ice algal production with a new in situ incubation technique". In: *Marine Ecology Progress Series* 177, pp. 15–26 (cit. on p. 8).

Montgomery, Douglas C, Cheryl L Jennings, and Murat Kulahci (2015). *Introduction to time series analysis and forecasting*. John Wiley & Sons (cit. on p. 21).

Morey, Marcos, Ana Fernandez-Marmiesse, Daisy Castineiras, Jose M Fraga, Maria L Couce, and Jose A Cocho (2013). "A glimpse into past, present, and future DNA sequencing". In: *Molecular genetics and metabolism* 110.1-2, pp. 3–24 (cit. on p. 19).

Mudelsee, Manfred (2010). "Climate time series analysis". In: *Atmospheric and* 397 (cit. on p. 21).

Mueter, Franz J, Benjamin Planque, George L Hunt Jr, Irene D Alabia, Toru Hirawake, Lisa Eisner, Padmini Dalpadado, Melissa Chierici, Kenneth F Drinkwater, Naomi Harada, et al. (2021). "Possible future scenarios in the gateways to the Arctic for Subarctic and Arctic marine systems: II. Prey resources, food webs, fish, and fisheries". In: *ICES Journal of Marine Science* 78.9, pp. 3017–3045 (cit. on p. 24).

Mustaffa, Nur Ili Hamizah, Liisa Kallajoki, Johanna Biederbick, Franziska Isabell Binder, Alexandra Schlenker, and Maren Striebel (2020). "Coastal ocean darkening effects via terrigenous DOM addition on plankton: an indoor mesocosm experiment". In: *Frontiers in Marine Science* 7, p. 547829 (cit. on p. 143).

Negrete-Garcıéa, Gabriela, Jessica Y Luo, Colleen M Petrik, Manfredi Manizza, and Andrew D Barton (2024). "Changes in Arctic Ocean plankton community structure and trophic dynamics on seasonal to interannual timescales". In: *EGUsphere* 2024, pp. 1–32 (cit. on p. 18).

Nöthig, Eva-Maria, Astrid Bracher, Anja Engel, Katja Metfies, Barbara Niehoff, Ilka Peeken, Eduard Bauerfeind, Alexandra Cherkasheva, Steffi Gäbler-Schwarz, Kristin Hardge, et al. (2015). "Summertime plankton ecology in Fram Strait—a compilation of long-and short-term observations". In: *Polar Research* 34.1, p. 23349 (cit. on pp. 18, 21).

Notz, Dirk and Julienne Stroeve (2016). "Observed Arctic sea-ice loss directly follows anthropogenic CO2 emission". In: *Science* 354.6313, pp. 747–750 (cit. on p. 15).

Nürnberg, Dirk, Ingo Wollenburg, Dirk Dethleff, Hajo Eicken, Heidemarie Kassens, Tom Letzig, Erk Reimnitz, and Jörn Thiede (1994). "Sediments in Arctic sea ice: Implications for entrainment, transport and release". In: *Marine geology* 119.3-4, pp. 185–214 (cit. on p. 8).

O'Mahony, Niall, Sean Campbell, Anderson Carvalho, Suman Harapanahalli, Gustavo Velasco Hernandez, Lenka Krpalkova, Daniel Riordan, and Joseph Walsh (2020). "Deep learning

vs. traditional computer vision". In: *Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 1 1*. Springer, pp. 128–144 (cit. on p. 24).

Oldenburg, Ellen et al. (2023a). "Sea-ice melt acts as a barrier establishment of temperate Atlantic taxa in atlantified Arctic Ocean". In: *TBA* 42.42, p. 43 (cit. on pp. 31, 32, 185).

Oldenburg, Ellen, Raphael M Kronberg, Katja Metfies, Wilken-Jon von Appen, Matthias Wietz, Christina Bienhold, Ovidiu Popa, and Oliver Ebenhöh (2024a). "Beyond blooms: A novel time series analysis framework predicts seasonal keystone species and sheds light on Arctic pelagic ecosystem stability". In: *bioRxiv*, pp. 2024–03 (cit. on pp. 28, 91, 92, 145, 185).

Oldenburg, Ellen, Raphael M Kronberg, Barbara Niehoff, Oliver Ebenhöh, and Ovidiu Popa (2023b). "DeepLOKI-a deep learning based approach to identify zooplankton taxa on high-resolution images from the optical plankton recorder LOKI". In: *Frontiers in Marine Science* 10, p. 1280510 (cit. on pp. 24, 28, 117, 118, 145).

Oldenburg, Ellen, Ovidiu Popa, Matthias Wietz, Wilken-Jon von Appen, Sinhue Torres-Valdes, Christina Bienhold, Oliver Ebenhöh, and Katja Metfies (2024b). "Sea-ice melt determines seasonal phytoplankton dynamics and delimits the habitat of temperate Atlantic taxa as the Arctic Ocean atlantifies". In: *ISME communications* 4.1, ycae027 (cit. on pp. 28, 145).

Orcutt, Beth N, Jason B Sylvan, Nina J Knab, and Katrina J Edwards (2011). "Microbial ecology of the dark ocean above, at, and below the seafloor". In: *Microbiology and molecular biology reviews* 75.2, pp. 361–422 (cit. on p. 7).

Osadchiev, Alexander, Ekaterina Kuskova, and Vladimir Ivanov (2024). "The roles of river discharge and sea ice melting in formation of freshened surface layers in the Kara, Laptev, and East Siberian seas". In: *Frontiers in Marine Science* 11, p. 1348450 (cit. on p. 1).

Overland, James E and Muyin Wang (2013). "When will the summer Arctic be nearly sea ice free?" In: *Geophysical Research Letters* 40.10, pp. 2097–2101 (cit. on p. 15).

Overland, James E, Muyin Wang, John E Walsh, and Julienne C Stroeve (2014). "Future Arctic climate changes: Adaptation and mitigation time scales". In: *Earth's Future* 2.2, pp. 68–74 (cit. on p. 15).

Pabi, Sudeshna, Gert L van Dijken, and Kevin R Arrigo (2008). "Primary production in the Arctic Ocean, 1998–2006". In: *Journal of Geophysical Research: Oceans* 113.C8 (cit. on p. 18).

Pace, Norman R (1997). "A molecular view of microbial diversity and the biosphere". In: *Science* 276.5313, pp. 734–740 (cit. on p. 19).

– (2009). "Mapping the tree of life: progress and prospects". In: *Microbiology and molecular biology reviews* 73.4, pp. 565–576 (cit. on p. 19).

Pace, Norman R, Jan Sapp, and Nigel Goldenfeld (2012). "Phylogeny and beyond: Scientific, historical, and conceptual significance of the first tree of life". In: *Proceedings of the National Academy of Sciences* 109.4, pp. 1011–1018 (cit. on p. 19).

Pagano, Anthony M and Terrie M Williams (2021). "Physiological consequences of Arctic sea ice loss on large marine carnivores: unique responses by polar bears and narwhals". In: *Journal of Experimental Biology* 224.Suppl_1, jeb228049 (cit. on p. 5).

Parrott, Lael, Clément Chion, Rodolphe Gonzalès, and Guillaume Latombe (2012). "Agents, individuals, and networks: modeling methods to inform natural resource management in regional landscapes". In: *Ecology and Society* 17.3 (cit. on p. 23).

Passow, Uta (1991). "Species-specific sedimentation and sinking velocities of diatoms". In: *Marine biology* 108, pp. 449–455 (cit. on p. 10).

Paul, Carsten, Anna Reunamo, Elin Lindehoff, Johanna Bergkvist, Michaela A Mausz, Henrik Larsson, Hannes Richter, Sten-Åke Wängberg, Piia Leskinen, Ulf Båmstedt, et al. (2012). "Diatom derived polyunsaturated aldehydes do not structure the planktonic microbial community in a mesocosm study". In: *Marine drugs* 10.4, pp. 775–792 (cit. on p. 8).

Paulsen, Maria L, Hugo Dore, Laurence Garczarek, Lena Seuthe, Oliver Mueller, Ruth-Anne Sandaa, Gunnar Bratbak, and Aud Larsen (2016). "Synechococcus in the Atlantic gateway to the Arctic Ocean". In: *Frontiers in Marine Science* 3, p. 191 (cit. on p. 18).

Pawlowicz, Richard, JF Lynch, WB Owens, PF Worcester, WML Morawitz, and PJ Sutton (1995). "Thermal evolution of the Greenland Sea Gyre in 1988–1989". In: *Journal of Geophysical Research: Oceans* 100.C3, pp. 4727–4750 (cit. on p. 1).

Pearson, Gareth A, Asuncion Lago-Leston, Fernando Cánovas, Cymon J Cox, Frederic Verret, Sebastian Lasternas, Carlos M Duarte, Susana Agusti, and Ester A Serrão (2015). "Metatranscriptomes reveal functional variation in diatom communities from the Antarctic Peninsula". In: *The ISME journal* 9.10, pp. 2275–2289 (cit. on p. 141).

Perovich, Donald, Walt Meier, M Tschudi, S Hendricks, AA Petty, D Divine, S Farrell, S Gerland, C Haas, L Kaleschke, et al. (2020). "Arctic report card 2020: Sea ice". In: (cit. on pp. 3, 6).

Perovich, Donald, Walter Meier, Mark Tschudi, Sinead Farrell, Stefan Hendricks, Sebastian Gerland, et al. (2017). "Sea Ice". In: *Arctic Report Card: Update for 2017* (cit. on p. 15).

Perovich, Donald K and Christopher Polashenski (2012). "Albedo evolution of seasonal Arctic sea ice". In: *Geophysical Research Letters* 39.8 (cit. on p. 3).

Perrette, Mahé, A Yool, GD Quartly, and Ekaterina E Popova (2011). "Near-ubiquity of ice-edge blooms in the Arctic". In: *Biogeosciences* 8.2, pp. 515–524 (cit. on p. 5).

Petchey, Owen L, Andrew Gonzalez, and Howard B Wilson (1997). "Effects on population persistence: the interaction between environmental noise colour, intraspecific competition and space". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 264.1389, pp. 1841–1847 (cit. on p. 21).

Petrich, Chris and Hajo Eicken (2010). "Growth, structure and properties of sea ice". In: *Sea ice* 2, pp. 23–77 (cit. on pp. 7, 8).

Pfirman, SL, Roger Colony, Dirk Nürnberg, Hajo Eicken, and I Rigor (1997). "Reconstructing the origin and trajectory of drifting Arctic sea ice". In: *Journal of Geophysical Research: Oceans* 102.C6, pp. 12575–12586 (cit. on p. 8).

Pimm, Stuart L and Andrew Redfearn (1988). "The variability of population densities". In: *Nature* 334.6183, pp. 613–614 (cit. on p. 21).

Pinhassi, Jarone, Maria Montserrat Sala, Harry Havskum, Francesc Peters, Oscar Guadayol, Andrea Malits, and Celia Marrasé (2004). "Changes in bacterioplankton composition under different phytoplankton regimens". In: *Applied and Environmental Microbiology* 70.11, pp. 6753–6766 (cit. on p. 18).

planktonnet.awi.de (2024). *PlanktonNet: A Global Biodiversity Database for Plankton Images*. Accessed: 2024-01-01. URL: `https://planktonnet.awi.de/` (cit. on pp. 12, 183).

Polyak, Leonid, Richard B Alley, John T Andrews, Julie Brigham-Grette, Thomas M Cronin, Dennis A Darby, Arthur S Dyke, Joan J Fitzpatrick, Svend Funder, Marika Holland, et al. (2010). "History of sea ice in the Arctic". In: *Quaternary Science Reviews* 29.15-16, pp. 1757–1778 (cit. on pp. 7, 15).

Polyakov, Igor V, John E Walsh, and Ronald Kwok (2012). "Recent changes of Arctic multiyear sea ice coverage and the likely causes". In: *Bulletin of the American Meteorological Society* 93.2, pp. 145–151 (cit. on pp. 15, 18).

Pomeroy, Lawrence R and William J Wiebe (2001). "Temperature and substrates as interactive limiting factors for marine heterotrophic bacteria". In: *Aquatic Microbial Ecology* 23.2, pp. 187–204 (cit. on p. 8).

Popa, Ovidiu, Ellen Oldenburg, and Oliver Ebenhöh (2020a). "From sequence to information". In: *Philosophical Transactions of the Royal Society B* 375.1814, p. 20190448 (cit. on p. 21).

– (2020b). "From sequence to information". In: *Philosophical Transactions of the Royal Society B* 375.1814, p. 20190448 (cit. on p. 145).

Poretsky, Rachel, Luis M Rodriguez-R, Chengwei Luo, Despina Tsementzi, and Konstantinos T Konstantinidis (2014). "Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics". In: *PloS one* 9.4, e93827 (cit. on p. 20).

Posch, Thomas, Bettina Eugster, Francesco Pomati, Jakob Pernthaler, Gianna Pitsch, and Ester M Eckert (2015). "Network of interactions between ciliates and phytoplankton during spring". In: *Frontiers in Microbiology* 6, p. 156959 (cit. on p. 11).

Post, Eric, Uma S Bhatt, Cecilia M Bitz, Jedediah F Brodie, Tara L Fulton, Mark Hebblewhite, Jeffrey Kerby, Susan J Kutz, Ian Stirling, and Donald A Walker (2013). "Ecological consequences of sea-ice decline". In: *science* 341.6145, pp. 519–524 (cit. on p. 5).

Price, PB, RA Rohde, and RC Bay (2009). "Fluxes of microbes, organic aerosols, dust, sea-salt Na ions, non-sea-salt Ca ions, and methanesulfonate onto Greenland and Antarctic ice". In: *Biogeosciences* 6.3, pp. 479–486 (cit. on p. 8).

Priest, Taylor, Wilken-Jon von Appen, Ellen Oldenburg, Ovidiu Popa, Sinhué Torres-Valdés, Christina Bienhold, Katja Metfies, William Boulton, Thomas Mock, Bernhard M Fuchs, et al. (2023). "Atlantic water influx and sea-ice cover drive taxonomic and functional shifts in Arctic marine bacterial communities". In: *The ISME Journal* 17.10, pp. 1612–1625 (cit. on pp. 28, 145).

Priest, Taylor, Wilken-Jon von Appen, Ellen Oldenburg, Ovidiu Popa, Sinhué Torres-Valdés, Christina Bienhold, Katja Metfies, Bernhard M Fuchs, Rudolf Amann, Antje Boetius, et al. (2022). "Variations in Atlantic water influx and sea-ice cover drive taxonomic and functional shifts in Arctic marine bacterial communities". In: *bioRxiv*, pp. 2022–08 (cit. on pp. 49, 50).

Priest, Taylor, Ellen Oldenburg, Ovidiu Popa, Bledina Dede, Katja Metfies, Wilkon-Jon von Appen, Sinhué Torres-Valdés, Christina Bienhold, Bernhard M Fuchs, Rudolf Amann, et al. (2024). "Seasonal recurrence and modular assembly of an Arctic pelagic marine microbiome". In: *bioRxiv*, pp. 2024–05 (cit. on pp. 28, 67, 68, 145, 185).

Ramanan, Rishiram, Byung-Hyuk Kim, Dae-Hyun Cho, Hee-Mock Oh, and Hee-Sik Kim (2016). "Algae–bacteria interactions: evolution, ecology and emerging applications". In: *Biotechnology advances* 34.1, pp. 14–29 (cit. on p. 10).

Ramasamy, Kesava Priyan, Lovely Mahawar, Raju Rajasabapathy, Kottilil Rajeshwari, Cristina Miceli, and Sandra Pucciarelli (2023). "Comprehensive insights on environmental adaptation strategies in Antarctic bacteria and biotechnological applications of cold adapted molecules". In: *Frontiers in Microbiology* 14, p. 1197797 (cit. on p. 8).

Rantanen, Mika, Alexey Yu Karpechko, Antti Lipponen, Kalle Nordling, Otto Hyvärinen, Kimmo Ruosteenoja, Timo Vihma, and Ari Laaksonen (2022). "The Arctic has warmed nearly four times faster than the globe since 1979". In: *Communications earth & environment* 3.1, p. 168 (cit. on p. 15).

Rapp, Josephine Z (2018). "Diversity and function of microbial communities in the Arctic Ocean". PhD thesis. Universität Bremen (cit. on p. 18).

Reigstad, Marit and Paul Wassmann (2007). "Does Phaeocystis spp. contribute significantly to vertical export of organic carbon?" In: *Phaeocystis, major link in the biogeochemical cycling of climate-relevant elements*, pp. 217–234 (cit. on p. 11).

Reuter, Jason A, Damek V Spacek, and Michael P Snyder (2015). "High-throughput sequencing technologies". In: *Molecular cell* 58.4, pp. 586–597 (cit. on p. 19).

Rhein, M, SR Rintoul, S Aoki, E Campos, D Chambers, RA Feely, S Gulev, GC Johnson, SA Josey, A Kostianoy, et al. (2013). "Observations: ocean in climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change". In: *Fifth assessment report of the Intergovernmental Panel on Climate Change*, pp. 255–316 (cit. on p. 4).

Riedel, Andrea, Christine Michel, Michel Gosselin, and Bernard LeBlanc (2007). "Enrichment of nutrients, exopolymeric substances and microorganisms in newly formed sea ice on the Mackenzie shelf". In: *Marine Ecology Progress Series* 342, pp. 55–67 (cit. on p. 9).

Rothrock, DA, DB Percival, and M Wensnahan (2008). "The decline in arctic sea-ice thickness: Separating the spatial, annual, and interannual variability in a quarter century of submarine data". In: *Journal of Geophysical Research: Oceans* 113.C5 (cit. on p. 15).

Rotter, Ana, Ariola Bacu, Michèle Barbier, Francesco Bertoni, Atle M Bones, M Leonor Cancela, Jens Carlsson, Maria F Carvalho, Marta Cegłowska, Meltem Conk Dalay, et al. (2020). "A new network for the advancement of marine biotechnology in Europe and beyond". In: *Frontiers in marine science* 7, p. 278 (cit. on p. 20).

Roukaerts, Arnout, Florian Deman, Fanny Van der Linden, Gauthier Carnat, Arne Bratkic, Sebastien Moreau, Delphine Lannuzel, Frank Dehairs, Bruno Delille, Jean-Louis Tison, et al. (2021). "The biogeochemical role of a microbial biofilm in sea ice: Antarctic landfast sea ice as a case study". In: *Elem Sci Anth* 9.1, p. 00134 (cit. on p. 9).

Rudels, B (2012). "Arctic Ocean circulation and variability–advection and external forcing encounter constraints and local processes". In: *Ocean Science* 8.2, pp. 261–286 (cit. on p. 6).

Rudels, B, Ursula Schauer, G Björk, M Korhonen, S Pisarev, Benjamin Rabe, and Andreas Wisotzki (2013). "Observations of water masses and circulation with focus on the Eurasian Basin of the Arctic Ocean from the 1990s to the late 2000s". In: *Ocean Science* 9.1, pp. 147–169 (cit. on pp. 1, 7, 18).

Rudels, Bert and Hans J Friedrich (2000). "The transformations of Atlantic water in the Arctic Ocean and their significance for the freshwater budget". In: *The freshwater budget of the Arctic Ocean*. Springer, pp. 503–532 (cit. on p. 1).

Rudels, Bert, E Peter Jones, Ursula Schauer, and Patrick Eriksson (2004). "Atlantic sources of the Arctic Ocean surface and halocline waters". In: *Polar research* 23.2, pp. 181–208 (cit. on p. 1).

Salcher, Michaela M, Jakob Pernthaler, and Thomas Posch (2011). "Seasonal bloom dynamics and ecophysiology of the freshwater sister clade of SAR11 bacteria 'that rule the waves'(LD12)". In: *The ISME journal* 5.8, pp. 1242–1252 (cit. on p. 6).

Schauer, Kristine, Dmitry A Rodionov, and Hilde de Reuse (2008). "New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'?" In: *Trends in biochemical sciences* 33.7, pp. 330–338 (cit. on p. 4).

Schauer, Ursula, Eberhard Fahrbach, Svein Osterhus, and Gerd Rohardt (2004). "Arctic warming through the Fram Strait: Oceanic heat transport from 3 years of measurements". In: *Journal of Geophysical Research: Oceans* 109.C6 (cit. on p. 4).

Schloss, Patrick D, Dirk Gevers, and Sarah L Westcott (2011). "Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies". In: *PloS one* 6.12, e27310 (cit. on p. 20).

Schweiger, Axel J, Kevin R Wood, and Jinlun Zhang (2019). "Arctic sea ice volume variability over 1901–2010: A model-based reconstruction". In: *Journal of Climate* 32.15, pp. 4731–4752 (cit. on p. 15).

Serreze, Mark C and Roger G Barry (2011). "Processes and impacts of Arctic amplification: A research synthesis". In: *Global and planetary change* 77.1-2, pp. 85–96 (cit. on pp. 3, 6).

Serreze, Mark C, Marika M Holland, and Julienne Stroeve (2007). "Perspectives on the Arctic's shrinking sea-ice cover". In: *science* 315.5818, pp. 1533–1536 (cit. on p. 15).

Sigman, Daniel M and Mathis P Hain (2012). "The biological productivity of the ocean". In: *Nature Education Knowledge* 3.10, p. 21 (cit. on p. 6).

Slagstad, Dag, Paul FJ Wassmann, and Ingrid Ellingsen (2015). "Physical constrains and productivity in the future Arctic Ocean". In: *Frontiers in Marine Science* 2, p. 85 (cit. on pp. 1, 18).

Smedsrud, Lars H, Mari H Halvorsen, Julienne C Stroeve, Rong Zhang, and Kjell Kloster (2017). "Fram Strait sea ice export variability and September Arctic sea ice extent over the last 80 years". In: *The Cryosphere* 11.1, pp. 65–79 (cit. on p. 15).

Smedsrud, Lars H, Morven Muilwijk, Ailin Brakstad, Erica Madonna, Siv K Lauvset, Clemens Spensberger, Andreas Born, Tor Eldevik, Helge Drange, Emil Jeansson, et al. (2022). "Nordic Seas heat loss, Atlantic inflow, and Arctic sea ice cover over the last century". In: *Reviews of Geophysics* 60.1, e2020RG000725 (cit. on p. 4).

Smedsrud, Lars Henrik, Anders Sirevaag, Kjell Kloster, Asgeir Sorteberg, and Stein Sandven (2011). "Recent wind driven high sea ice area export in the Fram Strait contributes to Arctic sea ice decline". In: *The Cryosphere* 5.4, pp. 821–829 (cit. on p. 5).

Snape, Thomas J and Piers M Forster (2014). "Decline of Arctic sea ice: Evaluation and weighting of CMIP5 projections". In: *Journal of Geophysical Research: Atmospheres* 119.2, pp. 546–554 (cit. on p. 15).

Soltwedel, Thomas, Eduard Bauerfeind, Melanie Bergmann, Astrid Bracher, Nataliya Budaeva, Kathrin Busch, Alexandra Cherkasheva, Kirsten Fahl, Katarzyna Grzelak, Christiane Hasemann, et al. (2016). "Natural variability or anthropogenically-induced variation? Insights from 15 years of multidisciplinary observations at the arctic marine LTER site HAUSGARTEN". In: *Ecological Indicators* 65, pp. 89–102 (cit. on pp. 1, 19).

Soltwedel, Thomas, Eduard Bauerfeind, Melanie Bergmann, Nataliya Budaeva, Eveline Hoste, Nina Jaeckisch, Karen von Juterzenka, Jens Matthießen, VADIM Mokievsky, E-M Nöthig, et al. (2005). "HAUSGARTEN: multidisciplinary investigations at a deep-sea, long-term observatory in the Arctic Ocean". In: *Oceanography* 3 (cit. on p. 1).

Soltwedel, Thomas, Ursula Schauer, Olaf Boebel, Eva-Maria Nöthig, Astrid Bracher, Katja Metfies, Ingo Schewe, Antje Boetius, and Michael Klages (2013). "FRAM-FRontiers in Arctic marine Monitoring Visions for permanent observations in a gateway to the Arctic Ocean". In: *2013 MTS/IEEE OCEANS-Bergen*. IEEE, pp. 1–7 (cit. on p. 2).

Spielhagen, Robert F, Kirstin Werner, Steffen Aagaard Sørensen, Katarzyna Zamelczyk, Evguenia Kandiano, Gereon Budeus, Katrine Husum, Thomas M Marchitto, and Morten Hald (2011). "Enhanced modern heat transfer to the Arctic by warm Atlantic water". In: *Science* 331.6016, pp. 450–453 (cit. on p. 18).

Spisla, Carsten (2021). "Marine zooplankton community responses to anthropogenic influences". PhD thesis. Christian-Albrechts-Universität zu Kiel (cit. on p. 13).

Spreen, Gunnar, Laura de Steur, Dmitry Divine, Sebastian Gerland, Edmond Hansen, and Ron Kwok (2020). "Arctic sea ice volume export through Fram Strait from 1992 to 2014". In: *Journal of Geophysical Research: Oceans* 125.6, e2019JC016039 (cit. on p. 15).

Staley, James T and Allan Konopka (1985). "Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats". In: *Annual review of microbiology* 39.1, pp. 321–346 (cit. on p. 19).

Steele, John H (1985). "A comparison of terrestrial and marine ecological systems". In: *Nature* 313.6001, pp. 355–358 (cit. on p. 21).

Steele, Michael and Suzanne Dickinson (2016). "The phenology of A rctic O cean surface warming". In: *Journal of Geophysical Research: Oceans* 121.9, pp. 6847–6861 (cit. on p. 15).

Steiner, Nadja S, Jeff Bowman, Karley Campbell, Melissa Chierici, Eeva Eronen-Rasimus, Marianne Falardeau, Hauke Flores, Agneta Fransson, Helena Herr, Stephen J Insley, et al. (2021). "Climate change impacts on sea-ice ecosystems and associated ecosystem services". In: *Elem Sci Anth* 9.1, p. 00007 (cit. on p. 8).

Stock, James H and Mark W Watson (2020). *Introduction to econometrics*. Pearson (cit. on p. 21).

Stockdale, Jennifer E, Jenny C Dunn, Simon J Goodman, Antony J Morris, Danae K Sheehan, Philip V Grice, and Keith C Hamer (2015). "The protozoan parasite Trichomonas gallinae causes adult and nestling mortality in a declining population of European Turtle Doves, Streptopelia turtur". In: *Parasitology* 142.3, pp. 490–498 (cit. on p. 13).

Stoecker, Diane K, Daniel E Gustafson, Megan MD Black, and Christine T Baier (1998). "Population dynamics of microalgae in the upper land-fast sea ice at a snow-free location". In: *Journal of Phycology* 34.1, pp. 60–69 (cit. on p. 8).

Stroeve, Julienne, Marika M Holland, Walt Meier, Ted Scambos, and Mark Serreze (2007). "Arctic sea ice decline: Faster than forecast". In: *Geophysical research letters* 34.9 (cit. on pp. 3, 15).

Stroeve, Julienne and Dirk Notz (2018). "Changing state of Arctic sea ice across all seasons". In: *Environmental Research Letters* 13.10, p. 103001 (cit. on p. 3).

Stroeve, Julienne, Mark Serreze, Sheldon Drobot, Shari Gearheard, Marika Holland, James Maslanik, Walt Meier, and Ted Scambos (2008). "Arctic sea ice extent plummets in 2007". In: *Eos, Transactions American Geophysical Union* 89.2, pp. 13–14 (cit. on p. 5).

Stroeve, Julienne C, Vladimir Kattsov, Andrew Barrett, Mark Serreze, Tatiana Pavlova, Marika Holland, and Walter N Meier (2012a). "Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations". In: *Geophysical Research Letters* 39.16 (cit. on p. 15).

Stroeve, Julienne C, Thorsten Markus, Linette Boisvert, J Miller, and A Barrett (2014). "Changes in Arctic melt season and implications for sea ice loss". In: *Geophysical Research Letters* 41.4, pp. 1216–1225 (cit. on pp. 15, 18).

Stroeve, Julienne C, Mark C Serreze, Marika M Holland, Jennifer E Kay, James Malanik, and Andrew P Barrett (2012b). "The Arctic's rapidly shrinking sea ice cover: a research synthesis". In: *Climatic change* 110, pp. 1005–1027 (cit. on pp. 5, 15).

Sugihara, George, Robert May, Hao Ye, Chih-hao Hsieh, Ethan Deyle, Michael Fogarty, and Stephan Munch (2012). "Detecting causality in complex ecosystems". In: *science* 338.6106, pp. 496–500 (cit. on p. 22).

Suzuki, Kenta, Shinji Nakaoka, Shinji Fukuda, and Hiroshi Masuya (2021). "Energy landscape analysis elucidates the multistability of ecological communities across environmental gradients". In: *Ecological Monographs* 91.3, e01469 (cit. on p. 23).

Swoboda, Steffen, Thomas Krumpen, Eva-Maria Nöthig, Katja Metfies, Simon Ramondenc, Jutta Wollenburg, Kirsten Fahl, Ilka Peeken, and Morten Iversen (2024). "Release of ballast material during sea-ice melt enhances carbon export in the Arctic Ocean". In: *PNAS Nexus*, pgae081 (cit. on p. 8).

Tableau (2024). "Time Series Analysis: Definition, Types & Techniques". In: 05.05.2024. URL: `https://www.tableau.com/learn/articles/time-series-analysis` (cit. on p. 21).

Talley, LD, GL Pickard, WJ Emery, JH Swift, and LD Talley (2011). "Arctic Ocean and Nordic seas". In: *Descriptive physical oceanography*, pp. 401–436 (cit. on p. 6).

Tamelander, Tobias, Marit Reigstad, Haakon Hop, and Tatjana Ratkova (2009). "Ice algal assemblages and vertical export of organic matter from sea ice in the Barents Sea and Nansen Basin (Arctic Ocean)". In: *Polar biology* 32, pp. 1261–1273 (cit. on p. 8).

Taylor, Patrick C, Ming Cai, Aixue Hu, Jerry Meehl, Warren Washington, and Guang J Zhang (2013). "A decomposition of feedback contributions to polar warming amplification". In: *Journal of Climate* 26.18, pp. 7023–7043 (cit. on p. 15).

Teeling, Hanno, Bernhard M Fuchs, Dörte Becher, Christine Klockow, Antje Gardebrecht, Christin M Bennke, Mariette Kassabgy, Sixing Huang, Alexander J Mann, Jost Waldmann, et al. (2012). "Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom". In: *Science* 336.6081, pp. 608–611 (cit. on p. 7).

Teeling, Hanno, Bernhard M Fuchs, Christin M Bennke, Karen Krüger, Meghan Chafee, Lennart Kappelmann, Greta Reintjes, Jost Waldmann, Christian Quast, Frank Oliver Glöckner, et al. (2016). "Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms". In: *elife* 5, e11888 (cit. on p. 7).

Tesi, Tommaso, Francesco Muschitiello, Gesine Mollenhauer, Stefano Miserocchi, Leonardo Langone, Chiara Ceccarelli, Giuliana Panieri, Jacopo Chiggiato, Alessio Nogarotto, Jens Hefter, et al. (2021). "Rapid Atlantification along the Fram Strait at the beginning of the 20th century". In: *Science advances* 7.48, eabj2946 (cit. on p. 4).

Thomas, David N, Gerhard Kattner, Ralph Engbrodt, Virginia Giannelli, Hilary Kennedy, Christian Haas, and Gerhard S Dieckmann (2001). "Dissolved organic matter in Antarctic sea ice". In: *Annals of Glaciology* 33, pp. 297–303 (cit. on p. 8).

Thomas, DN and GS Dieckmann (2002). "Antarctic sea ice–a habitat for extremophiles". In: *Science* 295.5555, pp. 641–644 (cit. on p. 7).

Thornton, Daniel CO (2014). "Dissolved organic matter (DOM) release by phytoplankton in the contemporary and future ocean". In: *European Journal of Phycology* 49.1, pp. 20–46 (cit. on p. 8).

Trego, Anna, Ciara Keating, Corine Nzeteu, Alison Graham, Vincent O'Flaherty, and Umer Zeeshan Ijaz (2022). "Beyond basic diversity estimates—analytical tools for mechanistic interpretations of amplicon sequencing data". In: *Microorganisms* 10.10, p. 1961 (cit. on p. 20).

Tremblay, J-É, S Bélanger, DG Barber, M Asplin, J Martin, G Darnis, L Fortier, Y Gratton, H Link, P Archambault, et al. (2011). "Climate forcing multiplies biological productivity in the coastal Arctic Ocean". In: *Geophysical Research Letters* 38.18 (cit. on p. 18).

Tremblay, Jean-Éric and Jonathan Gagnon (2009). "The effects of irradiance and nutrient supply on the productivity of Arctic waters: a perspective on climate change". In: *Influence of climate change on the changing arctic and sub-arctic conditions*. Springer, pp. 73–93 (cit. on p. 18).

Tremblay, Julien, Kanwar Singh, Edward S Kirton, Shaomei He, Tanja Woyke, and Susannah G Tringe (2015). "Primer and platform effects on 16S rRNA tag sequencing". In: *Frontiers in microbiology* 6, p. 150794 (cit. on p. 6).

Tschudi, Mark A, Julienne C Stroeve, and J Scott Stewart (2016). "Relating the age of Arctic sea ice to its thickness, as measured during NASA's ICESat and IceBridge campaigns". In: *Remote Sensing* 8.6, p. 457 (cit. on p. 15).

Tucker III, Walter B, John W Weatherly, Duane T Eppler, L Dennis Farmer, and Diane L Bentley (2001). "Evidence for rapid thinning of sea ice in the western Arctic Ocean at the end of the 1980s". In: *Geophysical Research Letters* 28.14, pp. 2851–2854 (cit. on p. 15).

Verdugo, Pedro, Alice L Alldredge, Farooq Azam, David L Kirchman, Uta Passow, and Peter H Santschi (2004). "The oceanic gel phase: a bridge in the DOM–POM continuum". In: *Marine chemistry* 92.1-4, pp. 67–85 (cit. on p. 8).

Verity, Peter G (1985). "Grazing, respiration, excretion, and growth rates of tintinnids 1". In: *Limnology and Oceanography* 30.6, pp. 1268–1282 (cit. on p. 11).

Vincent, Warwick F (2010). "Microbial ecosystem responses to rapid climate change in the Arctic". In: *The ISME journal* 4.9, pp. 1087–1090 (cit. on p. 18).

Volterra, Vito (1927). *Variazioni e fluttuazioni del numero d'individui in specie animali conviventi*. Vol. 2. Societá anonima tipografica" Leonardo da Vinci" (cit. on p. 142).

Von Quillfeldt, CH (2000). "Common diatom species in Arctic spring blooms: their distribution and abundance". In: (cit. on p. 10).

Walczowski, Waldemar, Agnieszka Beszczynska-Möller, Piotr Wieczorek, Malgorzata Merchel, and Agata Grynczel (2017). "Oceanographic observations in the Nordic Sea and Fram Strait in 2016 under the IO PAN long-term monitoring program AREX". In: *Oceanologia* 59.2, pp. 187–194 (cit. on p. 18).

Ward, R Matthew, Robert Schmieder, Gareth Highnam, and David Mittelman (2013). "Big data challenges and opportunities in high-throughput sequencing". In: *Systems Biomedicine* 1.1, pp. 29–34 (cit. on p. 20).

Wassmann, Paul (2011). *Arctic marine ecosystems in an era of rapid climate change* (cit. on pp. 3, 6, 18).

Wassmann, Paul, Carlos M Duarte, Susana Agusti, and Mikael K Sejr (2011). "Footprints of climate change in the Arctic marine ecosystem". In: *Global change biology* 17.2, pp. 1235–1249 (cit. on p. 6).

Weeks, Wilford F and Stephen F Ackley (1986). *The growth, structure, and properties of sea ice*. Springer (cit. on p. 7).

Wefing, Anne-Marie, Núria Casacuberta, Marcus Christl, Nicolas Gruber, and John N Smith (2021). "Circulation timescales of Atlantic Water in the Arctic Ocean determined from anthropogenic radionuclides". In: *Ocean Science* 17.1, pp. 111–129 (cit. on p. 1).

Wegner, Carolyn, Kerstin Wittbrodt, JA Hölemann, MA Janout, Thomas Krumpen, V Selyuzhenok, A Novikhin, Ye Polyakova, Irina Krykova, Heidemarie Kassens, et al. (2017). "Sediment entrainment into sea ice and transport in the Transpolar Drift: A case study from the Laptev Sea in winter 2011/2012". In: *Continental Shelf Research* 141, pp. 1–10 (cit. on p. 8).

Wells, Mark L, Vera L Trainer, Theodore J Smayda, Bengt SO Karlson, Charles G Trick, Raphael M Kudela, Akira Ishikawa, Stewart Bernard, Angela Wulff, Donald M Anderson, et al. (2015). "Harmful algal blooms and climate change: Learning from the past and present to forecast the future". In: *Harmful algae* 49, pp. 68–93 (cit. on p. 7).

Wilson, Bryan, Oliver Müller, Eva-Lena Nordmann, Lena Seuthe, Gunnar Bratbak, and Lise Øvreås (2017). "Changes in marine prokaryote composition with season and depth over an Arctic polar year". In: *Frontiers in Marine Science* 4, p. 95 (cit. on pp. 6, 7).

Winter, Amos and William G Siesser (2006). *Coccolithophores*. Cambridge University Press (cit. on p. 11).

Winton, Michael (2011). "Do climate models underestimate the sensitivity of Northern Hemisphere sea ice cover?" In: *Journal of Climate* 24.15, pp. 3924–3934 (cit. on p. 15).

Woese, Carl R and George E Fox (1977). "Phylogenetic structure of the prokaryotic domain: the primary kingdoms". In: *Proceedings of the National Academy of Sciences* 74.11, pp. 5088–5090 (cit. on p. 19).

Woese, Carl R, Otto Kandler, and Mark L Wheelis (1990). "Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya." In: *Proceedings of the National Academy of Sciences* 87.12, pp. 4576–4579 (cit. on p. 19).

Woodgate, Rebecca A, Thomas J Weingartner, and Ron Lindsay (2012). "Observed increases in Bering Strait oceanic fluxes from the Pacific to the Arctic from 2001 to 2011 and their impacts on the Arctic Ocean water column". In: *Geophysical Research Letters* 39.24 (cit. on p. 1).

Worden, Alexandra Z, Jae-Hyeok Lee, Thomas Mock, Pierre Rouzé, Melinda P Simmons, Andrea L Aerts, Andrew E Allen, Marie L Cuvelier, Evelyne Derelle, Meredith V Everett, et al. (2009). "Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes Micromonas". In: *Science* 324.5924, pp. 268–272 (cit. on p. 11).

Wylie, Kristine M, Rebecca M Truty, Thomas J Sharpton, Kathie A Mihindukulasuriya, Yanjiao Zhou, Hongyu Gao, Erica Sodergren, George M Weinstock, and Katherine S Pollard (2012). "Novel bacterial taxa in the human microbiome". In: *PloS one* 7.6, e35294 (cit. on p. 20).

Yadav, Juhi, Avinash Kumar, and Rahul Mohan (2020). "Dramatic decline of Arctic sea ice linked to global warming". In: *Natural Hazards* 103, pp. 2617–2621 (cit. on p. 5).

Yamanouchi, Takashi and Kumiko Takata (2020). "Rapid change of the Arctic climate system and its global influences-Overview of GRENE Arctic climate change research project (2011–2016)". In: *Polar science* 25, p. 100548 (cit. on p. 9).

Yamuza-Magdaleno, Alba, Rocıéo Jiménez-Ramos, Isabel Casal-Porras, Fernando G Brun, and Luis G Egea (2024). "Long-term sediment organic carbon remineralization in different seagrass and macroalgae habitats: implication for blue carbon storage". In: *Frontiers in Marine Science* 11, p. 1370768 (cit. on p. 13).

Yang, Meng, Yubao Qiu, Lin Huang, Maoce Cheng, Jianguo Chen, Bin Cheng, and Zhengxin Jiang (2023). "Changes in sea surface temperature and sea ice concentration in the Arctic Ocean over the past two decades". In: *Remote Sensing* 15.4, p. 1095 (cit. on p. 5).

Ye, Yufang, Mohammed Shokr, Georg Heygster, and Gunnar Spreen (2016). "Improving multiyear sea ice concentration estimates with sea ice drift". In: *Remote Sensing* 8.5, p. 397 (cit. on p. 5).

Yung, Charmaine CM, Elvira Rey Redondo, Frederic Sanchez, Sheree Yau, and Gwenael Piganeau (2022). "Diversity and evolution of Mamiellophyceae: early-diverging phytoplanktonic green algae containing many cosmopolitan species". In: *Journal of Marine Science and Engineering* 10.2, p. 240 (cit. on p. 11).

Zege, E, A Malinka, I Katsev, A Prikhach, Georg Heygster, L Istomina, Gerit Birnbaum, and Pascal Schwarz (2015). "Algorithm to retrieve the melt pond fraction and the spectral albedo of Arctic summer ice from satellite optical data". In: *Remote Sensing of Environment* 163, pp. 153–164 (cit. on p. 15).

Zhang, Haohao, Xuezhi Bai, and Kaiwen Wang (2023). "Response of the Arctic sea ice–ocean system to meltwater perturbations based on a one-dimensional model study". In: *Ocean Science* 19.6, pp. 1649–1668 (cit. on p. 3).

Zhang, Zhihua, Donald Huisingh, and Malin Song (2019). "Exploitation of trans-Arctic maritime transportation". In: *Journal of Cleaner Production* 212, pp. 960–973 (cit. on p. 18).

Zhao, Jinping, Wenli Zhong, Yina Diao, and Yong Cao (2019). "The rapidly changing Arctic and its impact on global climate". In: *Journal of Ocean University of China* 18, pp. 537–541 (cit. on p. 5).

Zhou, Botao, Ziyi Song, Zhicong Yin, Xinping Xu, Bo Sun, Pangchi Hsu, and Haishan Chen (2024). "Recent autumn sea ice loss in the eastern Arctic enhanced by summer Asian-Pacific Oscillation". In: *Nature Communications* 15.1, pp. 1–10 (cit. on p. 5).

Zhou, Jin, Mindy L Richlen, Taylor R Sehein, David M Kulis, Donald M Anderson, and Zhonghua Cai (2018). "Microbial community structure and associations during a marine dinoflagellate bloom". In: *Frontiers in microbiology* 9, p. 1201 (cit. on p. 7).

List of Figures

Declaration of Authorship

laut §5 der Promotionsordnung vom 15.06.2018

Ich versichere an Eides statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Ort, Datum

Ellen Oldenburg