

Marine microbes on the map: Defining spatial scales of functional microbial biogeography in the ocean

by

Cora Hörstmann

a Thesis submitted in partial fulfillment
of the requirements for the degree of

**Doctor of Philosophy
in Marine Microbiology**

Approved Dissertation Committee

Prof. Dr. Matthias Ullrich,
Jacobs University Bremen, Germany

Prof. Dr. Frank Oliver Glöckner,
Jacobs University Bremen, Germany;
AWI Bremerhaven, Germany

Prof. Dr. Anya M. Waite,
Ocean Frontier Institute,
Dalhousie University Halifax, Canada

Dr. Eric J. Raes,
Minderoo Foundation, Australia

Date of Defense: 23. Feb. 2022

Department of Life Sciences and Chemistry

*» It is interesting to contemplate an entangled bank,
clothed with many plants of many kinds, with birds
singing on the bushes, with various insects flitting about,
and with worms crawling through the damp earth, and to
reflect that these elaborately constructed forms, so
different from each other, and dependent on each other in
so complex a manner, have all been produced by laws
acting around us. «*

— Darwin

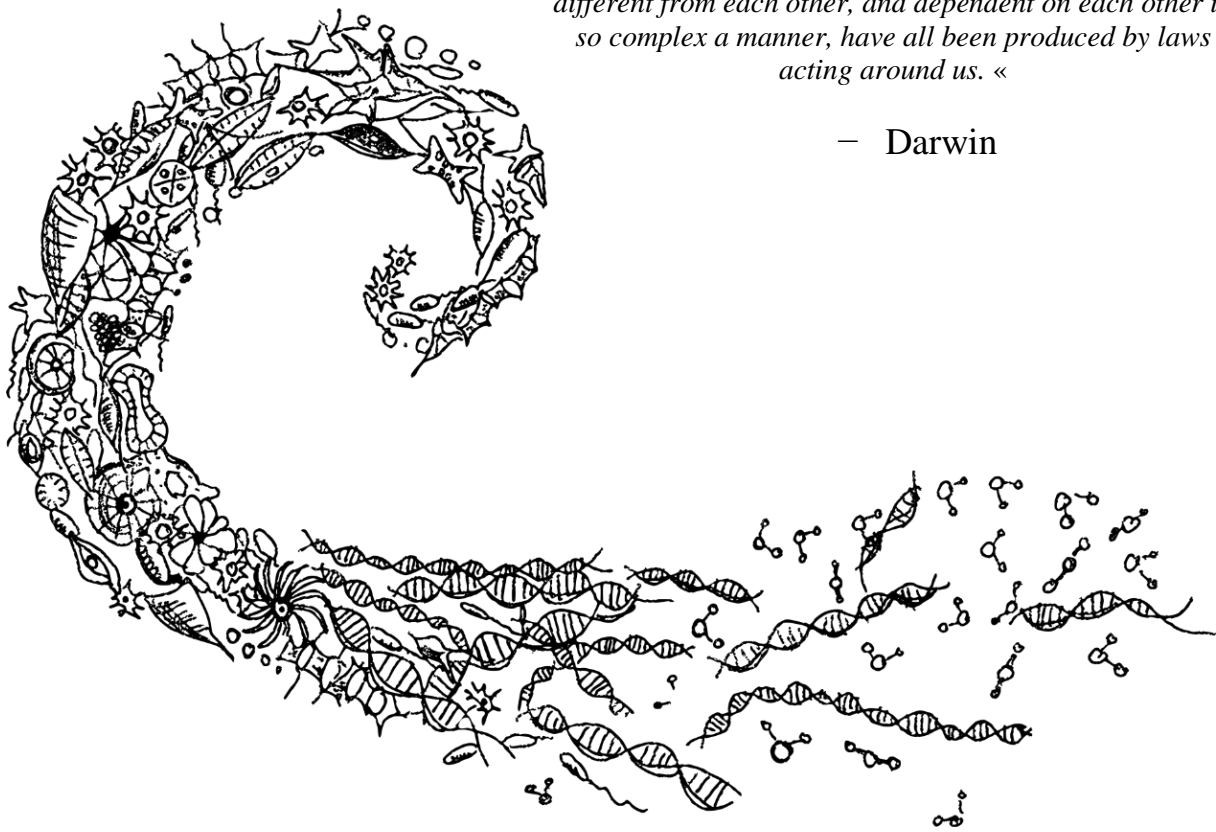


Table of Contents

I Summary	III
II Thesis Outline	V
III Glossary	IX
1. Background	1
1.1 Microbial diversity	2
1.2 Functional Microbial Biodiversity and its role in global carbon and nitrogen cycling	6
1.3 Microbial beta diversity and biogeography	9
1.4 Horizontal scales of microbial biodiversity	12
1.4.1 Observed bio-physical horizontal scales in the surface ocean	12
1.4.2 Methodological challenges that limit a global scaling of microbial observations	13
2. Objectives and hypotheses	15
3. Chapters	21
CHAPTER 1: Hydrographic fronts shape productivity, nitrogen fixation, and microbial community composition in the southern Indian Ocean and the Southern Ocean	23
Abstract	23
Chapter 1 1 Introduction	24
Chapter 1 2 Materials and methods	26
Chapter 1 3 Results	32
Chapter 1 4 Discussion	44
Chapter 1 5 Conclusion and outlook	49
CHAPTER 2: Microbial diversity through an oceanographic lens: refining the concept of ocean provinces through trophic-level analysis and productivity-specific length scales	51
Abstract	51
Chapter 2 1 Introduction	52
Chapter 2 2 Results	54
Chapter 2 3 Discussion	64
Chapter 2 4 Conclusion	70
Chapter 2 5 Materials and Methods	70

CHAPTER 3: Marine-terminating glaciers structure Arctic and sub-Arctic picoplankton diversity	79
Abstract	79
Chapter 3 1 Introduction.....	80
Chapter 3 2 Materials and Methods.....	82
Chapter 3 3 Results	87
Chapter 3 4 Discussion	96
Chapter 3 5 Conclusion	101
CHAPTER 4: Perspectives on Documenting Methods to Create Ocean Best Practices	103
Abstract	103
Chapter 4 1 Introduction.....	104
Chapter 4 2 What does it mean for a method to become a best practice (and why should we care)?	105
Chapter 4 3 How can we support the ocean community in creating best practices?	106
Chapter 4 4 Why are ethics and inclusiveness central to the documentation of methods and best practices?.....	108
Chapter 4 5 Concluding remarks and the path ahead	110
4. Synthesis.....	113
4.1 Alpha diversity, temperature and Primary Productivity	114
4.2 Mapping Microbes.....	117
4.3 Different levels of spatial scales: from a microbial community's perspective to regional and global scales.....	122
4.4 Future directions	127
5. Conclusions	131
Acknowledgement.....	133
6. References	135
Supplementary 1 Supplementary to Chapter 1	163
Supplementary 2 Supplementary to Chapter 2	175
Supplementary 3 Supplementary to Chapter 3	191
Supplementary 4.....	205
Statutory Declaration of Authorship.....	207

I Summary

Marine microorganisms have high functional and phylogenetic diversity and, through their high physiological activity, rapid turnover rates (hours to days). Contributing a high proportion (~70%) of total marine biomass, they sustain major elemental cycles, including those of carbon and nitrogen, impacting all (ocean) life on earth. The occurrence of different taxonomic units at one site (alpha diversity), and the differential distribution of microorganisms between sites (beta diversity) – and consequently the spatial distribution of their metabolic activity – is linked to complex ocean dynamics, including the continuous transport of microbial communities with ocean currents. It has been shown that the spatial distribution of microbial communities maps roughly to ocean provinces that are defined based on basic physical variables (temperature, salinity) and chlorophyll *a* concentrations. However, a more refined picture of microbial (alpha and beta) diversity patterns through an oceanographic lens is lacking. For example, it is unclear how physiological activity (e.g., primary productivity and N₂ fixation) covaries with phylogenetic beta diversity. Such information could provide insights into ecosystem function and thus add an important functional dimension to spatial scales of microbial biogeography.

In this thesis, I explore the relationship between microbial biodiversity (16S and 18S ribosomal RNA gene sequencing), and physiological rates of primary productivity, N₂ fixation in the surface ocean (0–40 m) in the Atlantic and the Indian Ocean, as well as coastal observations of Arctic and sub-Arctic fjords. I also include a consideration of chemical (dissolved inorganic nutrients, particulate organic matter) and physical environmental variables. I demonstrate how functional activity can be decoupled from phylogenetic diversity and observe environmental filtering of trophic functional groups in regions of high- and low chlorophyll *a*. I show that beta diversity patterns reflect previous

delineations of ocean provinces and can in turn also be used to refine their boundaries. Furthermore, the concept and quantification of a productivity-specific length scale can help to identify sample patchiness and scale sample diversity in relation to small spatial scales within the mosaic of marine microbial diversity structure. I show how microbial communities in fjords disperse regionally, and form both regional and within-fjord signals with different co-occurrence patterns between fjords with and without marine-terminating glaciers. My analyses highlight that mapping of microbial diversity patterns is scale-dependent, providing new insights into spatial scales and patchiness of microbial functional biogeography. The scales identified in this thesis can be used to improve sampling design and analyses across temporal and spatial scales, ultimately advancing ecosystem monitoring. Designing effective ecosystem monitoring is fundamentally an ecological challenge of identifying relevant sampling scales. It is, however, also a methodological challenge, of data inter-compatibility through consistent method documentation that can also be particularly relevant to interdisciplinary research such as the biophysical-coupling in this thesis. To contribute to closing this methodological gap, I provide perspectives on best practices in method documentation that aim to make scientific analyses more confidently intercomparable.

In conclusion, my thesis provides insights into an oceanographic definition of pelagic microbial ecosystem boundaries, and the impact of (functional) microbial diversity on primary productivity. The presented approaches and compiled information will support mapping microbiomes to relevant oceanographic scales and has potential implications for new approaches in researching, observing, and monitoring of all trophic layers in marine ecosystems.

II Thesis Outline:

The presented thesis consists of an introductory background (**Section 1**), the study aims (**Section 2**), three research articles and one perspective article (**Chapter 1–3; Section 3**), a synthesis of the presented research based on the study aims, anticipated future directions (**Section 4**), and a conclusion (**Section 5**). Specifically, chapters of the presented thesis are subdivided in the following four manuscripts:

Chapter 1:

This chapter consists of a manuscript published in *Biogeosciences* under the CC BY-4.0 license, and parts of this study were already published as part of Cora Hörstmann's (CH's) master thesis. (Notably, chl *a* concentrations were corrected as part of this thesis, also highlighted in the corrigendum of the published manuscript). During her time as a PhD student, CH investigated the eukaryotic community composition of the presented transect, including amplicon 18S rRNA gene PCR, Illumina sequencing library preparation, sequencing, and statistical analysis. Further, CH performed a general dissimilarity model (GDM) analysis of both prokaryotic and eukaryotic amplicon sequences to investigate the relationship between primary productivity and di-nitrogen fixation measurements in relation to microbial beta diversity. In her work, CH was supervised by Eric J. Raes (experimental design, sample processing), Uwe John (amplicon sequencing and genomic analysis), Pier Luigi Buttigieg (statistical advice), and Anya M. Waite (experimental design, sample processing). All co-authors contributed to the data interpretation and write-up of the manuscript.

Published: **Hörstmann, C.**, Raes, E.J., Buttigieg, P.L., Lo Monaco, C., John, U., Waite, A.M (2021). Hydrographic fronts shape productivity, nitrogen fixation, and microbial

community composition in the South Indian Ocean and the Southern Ocean, *Biogeosciences*, <https://doi.org/10.5194/bg-2021-52>

Master thesis: **Hörstmann, C.**, (2018). How do pro- and eukaryotic microbial communities impact nitrogen and carbon processes in the South Indian Ocean and the French Southern and Antarctic Lands? Master thesis, Universität Bremen.

hdl:[10013/epic.43436e95-5418-410f-903c-e6a06b0bae19](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0011-9)

Chapter 2:

This chapter consists of a manuscript published in *Environmental Microbiology* under the CC BY-4.0 license (see below). CH led this study including sampling and experimental design, ship-board sampling (particulate organic matter, dissolved inorganic nutrients, flow cytometry, genomic DNA samples) and stable isotope (^{13}C) incubation experiments, sample processing (particulate organic matter, stable isotope pre-processing and shipping, flow cytometry, DNA extraction and amplicon library preparation, sequence analysis), statistical analysis and manuscript writing. All co-authors contributed to data analysis, data interpretation and writing the manuscript.

Published: **Hörstmann, C.**, Buttigieg, P.L., John, U., Raes, E.J., Wolf-Gladrow, D., Bracher, A., Waite, A.M (2021). Microbial diversity through an oceanographic lens: refining the concept of ocean provinces through trophic-level analysis and productivity-specific length scales, *Environmental Microbiology*, doi: [10.1111/1462-2920.15832](https://doi.org/10.1111/1462-2920.15832)

Chapter 3:

Cora Hörstmann co-designed the presented study. Preliminary analyses were made as part of Pauline Thomé's bachelor thesis. CH co-supervised Pauline Thomé in her lab work, statistical analysis, and data interpretation. CH performed all presented statistical analyses

and drafted the manuscript. All co-authors contributed to data interpretation and writing the manuscript. The manuscript is currently in review in *Science Advances*.

Bachelor thesis: Thomé, P. (2020) Spatial biodiversity patterns of bacterio- and picoplankton communities in Arctic fjords; Bachelor thesis, Georg-August Universität zu Göttingen. hdl:[10013/epic.9d50f18c-f636-4c5f-ab33-8846853c6064](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63644-p0011-9)

Manuscript: Hörstmann, C., Hattermann, T., Thomé, P., Buttigieg, P. L., Morel, I., Waite, A. M., John, U. Marine-terminating glaciers structure Arctic and sub-Arctic picoplankton diversity

Chapter 4:

Chapter 4 consists of a peer-reviewed perspective article that CH led and that is published in *Frontiers in Marine Science* under the CC BY-4.0 license. The perspectives provided, although addressed to a wider ocean observing audience, have significant implications for the scientific work presented in this thesis and are therefore considered as a reflective and outlook chapter of the research in this thesis. All co-authors co-developed the perspectives and contributed to the writing of the manuscript.

Published: Hörstmann, C., Buttigieg, P.L., Simpson, P., Pearlman, J, Waite, A.M (2021). Perspectives on Documenting Methods to Create Ocean Best Practices, *Frontiers in Marine Science*. 7:556234. doi: [10.3389/fmars.2020.556234](https://doi.org/10.3389/fmars.2020.556234)

III Glossary

Alpha diversity. Total of all taxonomic units, here amplicon sequence variants (ASVs) (i.e., sample richness) and their relative abundance (i.e., sample evenness) within one sample.

Beta diversity. The extent of change of all taxonomic units considering presence, absence and their relative abundance between two (or more) sites.

Environmental filtering. describes the presence of abiotic variables that do not allow the establishment or survival of particular species within a location (Kraft et al. 2015).

Functional traits. Traits that are connected to an organisms' fitness (growth, reproduction) within a particular environment (Litchman et al. 2015).

Heterotrophy. The dependence of organisms to take up and grow on organic carbon sources.

Niche space (Hutchinsonian niche). Here, as defined by Hutchinson (1957), refers to the ecological niche of a species as an n-dimensional hypervolume shaped by the environment under which conditions a species can exist indefinitely and was further extended by Maguire (1973) with species demographics. Pironon et al. (2018), investigated the importance of considering demographic distribution and the distribution of species in their niche space.

Microbial residence time. The time a microbe resides in a system (Mansfeldt et al. 2019) often used to understand dispersal and speciation of marine microorganisms.

Mixotrophy. Here, referring to the mixotrophic lifestyle of marine eukaryotes that include both the ability to fix inorganic carbon via photosynthesis and the ability to take up organic material via phagocytosis. Notably, the presence of proteorhodopsin is not considered in this thesis' analyses. Likewise, prokaryotes were generally not categorized according to their energy source.

Photoautotrophy. Organisms that use light energy to produce organic molecules from inorganic carbon sources. This includes various taxonomic groups of marine eukaryotes and cyanobacteria.

Primary productivity. The rate at which autotrophic organisms produce multi-carbon compounds for metabolic processes using energy from either chemical reaction (chemolithoautotrophs) or from sunlight (photoautotrophs). In this thesis, I am only referring to photoautotrophy measured by stable isotope carbon (^{13}C) fixation incubation experiments.

Functional redundancy/Partial functional redundancy. Organisms share the same set of functions (strict functional redundancy), or share some functions, but differ in other ecological requirements (Galand et al. 2018).

Omics. Here, referring to biomolecular methods including metabarcoding (amplicon sequencing), (meta)genomics, (meta)transcriptomics, proteomics, and metabolomics.

1. Background

Understanding the interplay of the mechanisms shaping microbial diversity in the ocean is key to the research of community ecology, biogeography, and also rapidly becoming urgent for observation and management design of marine ecosystems (Barnosky et al. 2012; Mittelbach and Schemske 2015). Thus, interdisciplinary, multi-dimensional approaches are needed to disentangle the environmental and biological factors orchestrating microbial diversity patterns in the ocean. Yet, microbial diversity, productivity, and oceanographic processes are less understood in the unified context of biophysical coupling, except in a few studies (e.g. the “dual-lens” approach of investigating a biological reaction rate in relation to oceanographic connectivity that can provide insights into the balance between hydrological and ecological processes; Oldham et al. 2013). Simultaneous measurements of biological and physical processes allow us to study microbial biodiversity changes at scales that are relevant for ecological and biogeochemical processes: from small horizontal changes within single biomass turnover (a microbe’s perspective) to large-scale productivity regimes that are usually remotely observed through chlorophyll *a* (chl *a*) satellite images. Properly measuring physical processes at these different scales, and the biogeochemical processes and mechanisms therein, were the motivation for this thesis. In the following background information, I will introduce the concept of microbial diversity and functional microbial diversity and set this in the context of microbial biogeography and the different mechanisms that play a role in creating ocean microbiomes.

1.1 Microbial diversity

Marine microorganisms, including both prokaryotic (bacteria and archaea) and unicellular eukaryotic (protists) organisms, make up the largest fraction (~70%) of biomass in the ocean (Bar-on et al. 2018). Microorganisms form the base of the marine food web and harbor a rich diversity which is predicted to range between 2.2–4.3 million prokaryotic operational taxonomic units (OTUs; 16S-V4 gene clusters at 97% similarity) worldwide (Louca et al. 2019) with another total estimate of 200 000–250 000 eukaryotic species (Debroas et al. 2017). Yet, we are only scratching the surface of identifying *who is there*, as the true number of microbial OTUs remains unknown (reflected in Dance 2020). Microorganisms typically have large population sizes and fast generation turnover between hours and days (Martiny et al. 2006) and thus both indicate and also shape biological change (Vincent 2010). The dynamic change of microbial taxa across temporal and spatial scales, as well as their high richness, requires specific methods and concepts to study their diversity and distribution.

High throughput sequencing technologies allow us to better understand the vast diversity and biological processes of ocean microorganisms (Vaulot et al. 2008; Caporaso et al. 2011). However, our understanding is limited by both sampling incompleteness and methodological biases (e.g. primer bias, preference of the polymerase for sequences with specific guanine-cytosine (GC) content; see van der Loos and Nijland (2020) for further details) that are required to disentangle the complex interplay of microbial organisms within one sample, region, and across entire ocean basins. Solutions to counter these biases and to make samples better comparable are methodological developments (e.g. mock communities; Yeh et al. 2018), but also post-processing, analytical methods that address the compositionality of the data such as alpha diversity measures (Chao et al. 2014).

The concept of alpha diversity is composed of different attributes of the observed community and can be context-specifically weighted and interpreted (Swingland 2001). There are several commonly used alpha diversity indices to estimate microbial diversity.

In my thesis, I adopt the measure of Hill numbers (Hill 1973). Hill numbers indicate the relative emphasis on rare or more abundant taxa, and are particularly suitable for the high diversity of microbial taxa and the associated sampling bias in its quantification (Kang et al. 2016). Hill numbers are parameterized by an order q and include diversity (D) Richness ($q = 0$), Shannon diversity ($q = 1$), and Simpson diversity ($q = 2$):

$$D_q = (\sum_{i=1}^S p_i^q)^{1/(1-q)} \quad (1)$$

Where S is the total number of observed taxa in a sample and p_i indicates the relative abundance of each taxonomic unit and q the individual Hill numbers. The individual Hill numbers and their application in microbial ecology are outlined in **box 1**.

Generally, these diversity measures help to understand the “ecological potential” to support species and gives us an idea about ecological niches. Changes in alpha diversity can act as an numerical indicator for ecosystem change (Crummett 2020) and allow us to identify microbial diversity patterns on the map that relate to physical oceanographic features ranging from small, internal hydrographic motions to large, latitudinal trends across ocean basins (Raes et al. 2018).

BOX 1 Alpha diversity measures (Hill numbers)

Richness: If $q = 0$, the abundance of individual taxa is not considered in equation (1) then

$$D_0 = (\sum_{i=1}^S p_i^0) \quad (2)$$

and thus simply returns the total number of taxa (S), i.e. the sample richness. Richness can be largely biased by sampling and sequencing depth, and is therefore relatively more prone to error than other diversity measures. Mostly, microbial samples are rarefied (randomly subsampling the sample and observation how often a new species is detected) to observe the possibility of capturing the true richness in a sample (i.e. rarefaction curve plateaus; see below).

Richness does not provide any information about relative abundances and thus lacks a crucial component of diversity from an ecological perspective (Pielou 1975; Chao et al. 2014). Therefore,

henceforth, I will consider richness independently and refer to one of the following indices as the diversity indices.

Shannon diversity: For $q = 1$, equation (1) is undefined, but if q tends to 1 it returns

$$D_1 = \lim_{q \rightarrow 1} D_1 = \exp\left(-\sum_{i=1}^S p_i \ln(p_i)\right) \quad (3)$$

(For mathematical proof see Hill 1973). Thus, individual taxa are weighted proportionally to their relative abundance (Chao et al. 2014). Shannon diversity is particularly useful to gain insights into the current state of the system, i.e. who is dominating and how balanced is the community (sample evenness) is without losing too much information about the sample richness. A comparative study of multiple diversity indices by Reese and Dunn (2018) revealed that the Shannon diversity is a robust metric and good for cross-comparing different studies as it is the most widely used metric in the field.

Simpson diversity: If $q = 2$ then consequently more weight is put on the taxa with a proportionally higher relative abundance and turns equation (1) into

$$D_2 = \frac{1}{\sum_{i=1}^S p_i^2} \quad (4)$$

The Simpson diversity is the most robust measure when true richness is unknown, as it largely neglects the rare biosphere. Simultaneously, it puts more weight on sample evenness and therefore provides insights into the sample complexity. For example, assuming two microbial communities with the same number of entities (S), Simpson diversity is relatively lower in a bloom situation where a single or few species dominate the community (Community B) and is relatively higher if multiple species are equally high abundant (Community A).

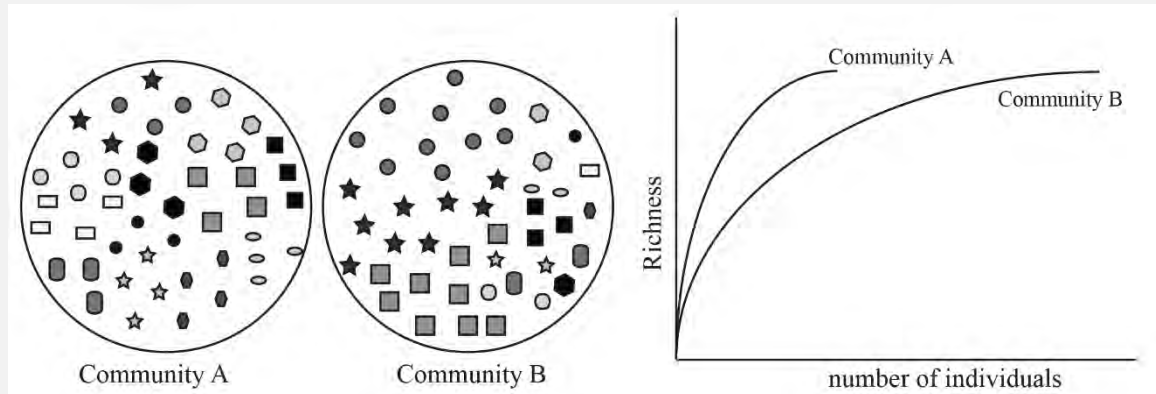


Figure B1.1 Two hypothetical communities A and B with the same sample richness (S) but different relative abundances in their communities: Community A is relatively even while Community B has few taxonomic units that dominate the community. The rarefaction curve of community A is much steeper as the probability to detect a new species with increasing sampling depth is much higher, while the chance to re-sample individuals which belong to the few high abundant taxonomic units is higher in Community B.

Two major environmental attributes have been repeatedly linked to variations in alpha diversity: (1) alpha diversity increases with increasing temperature due to favourable kinetics of biological processes (Rohde 1992), and (2) diversity increases with increasing primary productivity following the concept that “a larger pie can be divided into more pieces” (*sensu* Fuhrman et al. 2008), meaning that more available resources can support a more diverse community (e.g. Connell and Orias 1964). Notably, most studies focus on the concept that resources support microbial diversity, while a support of primary productivity through a more efficient resource use by a diverse microbial community can equally apply (Cardinale et al. 2009).

Ocean ecosystems are experiencing increasing anthropogenic pressures, including shifts in sea surface temperatures (IPCC 2021) that will also impact both microbial diversity and primary productivity. Defining and measuring the spatial scales of these ecosystem-level changes is key to a better understanding of the different mechanisms that drive changes in microbial diversity. Thus, sampling at high spatial and temporal resolution is needed to map the changes of microbial diversity. Furthermore, it is essential to understand which roles individual taxa and communities, and their co-occurrence and co-exclusion relationships, play within the ecosystem. In an analysis of microbial diversity in the context of a global climate prediction model, Tonelli et al. (2021) observed a decrease in microbial richness and diversity in the Southern Ocean with increasing sea surface temperature (between 2015 and 2100), a trend that has also been described for phytoplankton diversity in tropical regions (Thomas et al. 2012). Through the coupling of phytoplankton biodiversity and primary productivity may also change (Tilman et al. 1996), as well as changes in bacterial production rates (Sarmiento et al. 2010), but disentangling the drivers of these processes remains challenging. We are currently missing information to accurately map the consequences, how functional diversity and functional redundancy (organisms share the same set of functions or have overlaps in their functions) could ‘buffer’ the loss of species (Wolf et al. 2021). Thus, trait-based approaches are urgently needed such as trophic-disaggregated analysis of microbial diversity (reviewed in Seibold et al. 2018). Through such analyses, microbial diversity could be more directly related to uptake and

cycling of key elements such as carbon (C) and nitrogen (N) across the multiple levels of spatial and temporal scales (Worden et al. 2015; Hutchins and Capone 2022).

1.2 Functional Microbial Biodiversity and its role in global carbon and nitrogen cycling

Microorganisms are the key drivers for the cycling of C and N (**Box 2**). Microorganisms are responsible for around ~50% of global C fixation via photosynthesis (Field et al. 1998; McClain 2009), and recycle around 50% of the C from primary productivity (Anderson and Ducklow 2001) within the microbial loop (reviewed in Azam and Malfatti 2007). Microbial-mediated nutrient recycling and biological N₂ fixation also support ecosystem productivity through the provision of essential nutrients, tightly coupling C and N cycling. However, flux measurements quantifying nutrient uptake, such as C and N cycling, across ocean ecosystems are rarely coupled to whole-community microbial biodiversity assessments, except in a few studies (e.g. Peter et al. 2011; Raes et al. 2018).

For example, recent metagenomic analyses revealed that *Trichodesmium spp.*, which was long considered the champion of marine N₂ fixation, can be non-diazotrophic under unfavorable, low-oxygen conditions (Delmont 2021). In turn, the discovery of heterotrophic N₂ fixation has significantly changed our understanding of the contribution and importance of marine N₂ fixation (e.g. Candidatus *Atelocyanobacterium Thalassa*; Thompson et al. 2012), highlighting the importance of whole-community analysis to identify activity (or community potential to become active) in researching global elemental cycles.

Of course, the taxonomic community composition does not provide a clear picture of ecosystem function without flux measurements either, because the metabolic pathways, such as C-fixation, are present in co-existing, taxonomically distinct organisms. This

functional redundancy can then – to a certain extent – buffer dynamic change of individual taxa in relation to overall changes of major elemental cycles such as C and N (Louca et al. 2016, 2018). In a meta-analysis of multiple studies, Graham et al. (2016) found that the prokaryotic community diversity (based on 16S rRNA gene analyses) was only a weak predictor for C and N cycling. These observations align with results by (Louca 2021) who found that most prokaryotic clades – and thus their metabolic function – are globally distributed and local differences are mainly a result of local environmental filtering.

However, if the selective forces of environmental filtering change, then (biologically) the set of beneficial traits that influence an organism's fitness in a specific environment also change (Violle et al. 2007). Thus, microbial communities performing environmental functions need to be considered in a context of this “partial redundancy”. Galand et al. (2018) showed that community shifts in diversity were associated with a change in functional traits across seasonal scales and biome-level spatial scales. Setting C and N fluxes in relation to *whole-community* biological traits will help to identify cascading effects (e.g. changes in trophic coupling in the community affect C-remineralization and C-export efficiency; Stukel et al. 2011; Le Moigne et al. 2015). Including microbial diversity and microbial trait analysis in researching C and N fluxes can reveal the mechanisms behind changes in these processes and help to better understand and predict ecosystem functions.

BOX 2: Microbial role in global C and N cycling

Carbon (C) and nitrogen (N) are the two of the most important elements sustaining life on earth (Fig. B1 a). The biggest reservoir for C (~80%) is in the ocean, in the form of dissolved inorganic carbon (DIC), and a small fraction of particulate organic carbon (POC) (Friedlingstein et al. 2019). C can be exported to the deep sea through the biological and physical C pump where it is buried on time scales of the order of 1000–10 000 years. However, most of the C and N is recycled in the upper ocean within the microbial loop and thus again made available for primary productivity (Fig. B2.1b). Biological N₂ fixation, terrestrial N input and upwelling of N can promote photosynthetic, autotrophic plankton growth and, by that, is tightly coupled to the microbial loop (Fig. B2.1b, c).

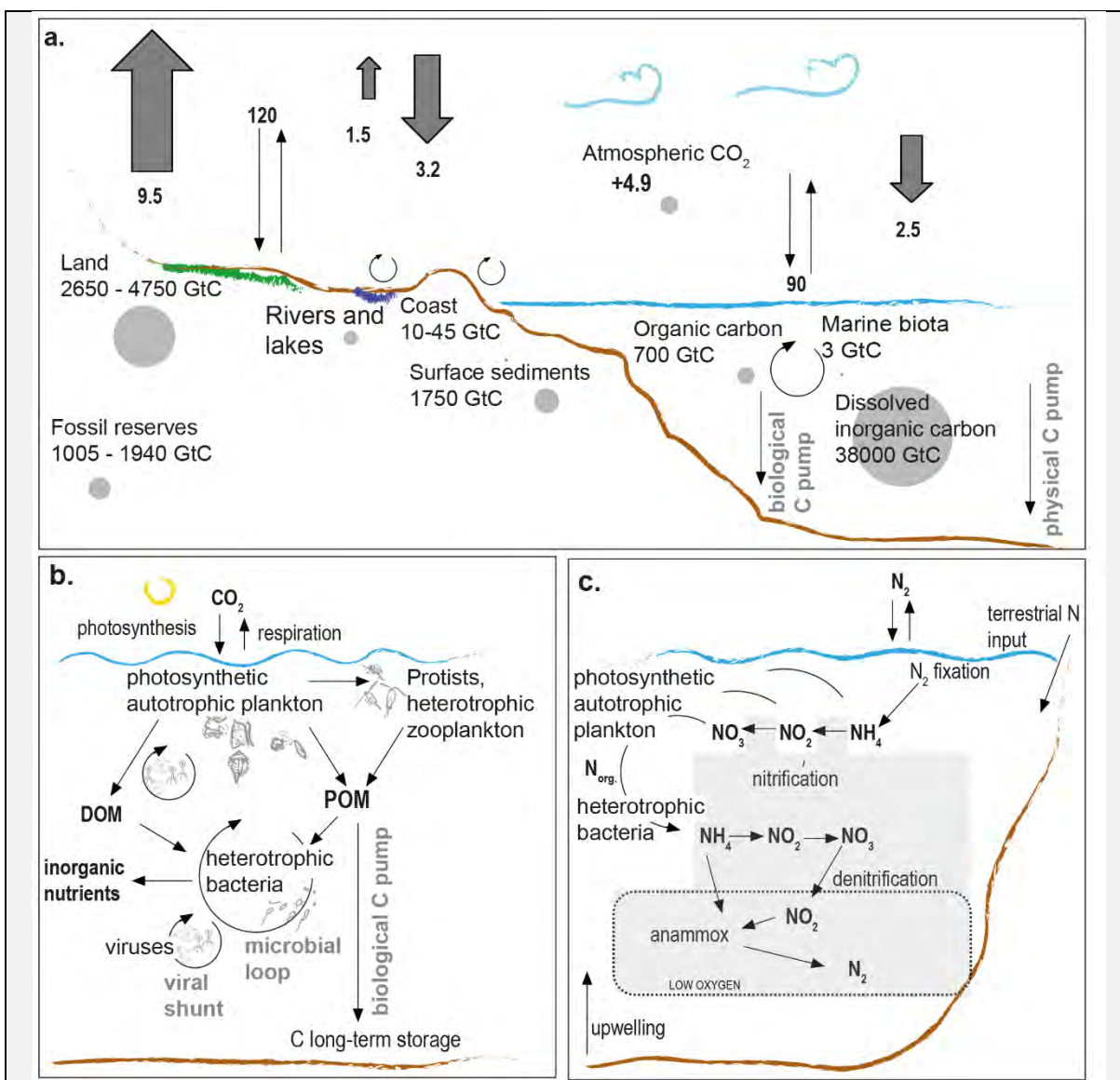


Figure B2.1 a. Global carbon cycle adapted and modified after (Friedlingstein et al. 2019). Scheme is simplified with a focus on oceanic C cycling. Land C budget includes permafrost, vegetation and soil. Fossil C budget includes coal, oil and gas reserves. Numbers indicate and small arrows indicate fluxes (in Gt C year⁻¹) and stocks (in Gt C) of carbon. Bold arrows indicate anthropogenic perturbation of C cycling through fossil fuels, land use change, land uptake and ocean uptake (uncertainties are not shown). Atmospheric increase is 4.9 Gt C year⁻¹ with a relatively small uncertainty of 0.02 Gt C year⁻¹ (not represented in the figure); **b.** Biological carbon pump and microbial loop adapted and modified after (Buchan et al. 2014). In brief, photosynthetic autotrophic plankton fixes inorganic C into organic C via photosynthesis. Phytoplankton are grazed by heterotrophic protists and zooplankton, are lysed by viruses, releasing dissolved organic matter (DOM), or die and form particulate organic matter (POM) that includes particulate organic C, N and phosphorus (P). DOM and POM are recycled by heterotrophic bacteria within the microbial loop and thus make available nutrients for phytoplankton growth. Viruses contribute and promote the release of POM and DOM through microbial cell lysis (viral shunt; not discussed in this thesis). A small fraction of the POM sinks to the seafloor in the form of aggregates (‘marine snow’) and can be stored on geological time scales (biological C pump); **c.** marine nitrogen (N) cycle. Light grey indicates N processes not discussed in this thesis.

1.3 Microbial beta diversity and biogeography

Surface ocean currents, fronts and eddies shape the mosaic of niches and ecosystem structures underlying phylogenetic and functional microbial biodiversity in the ocean. In principle, a given microbial community can move globally, passively dispersed through ocean currents, but local environmental factors constrain growth of different microbial assemblages (*'Everything is everywhere but the environment selects'*; Baas Becking 1934). However, fast generation times and large population sizes (Elena and Lenski 2003) including historical attributes of communities, also contribute to observed diversity differences across spatial (and temporal) scales (Martiny et al. 2006; Hellweger et al. 2014).

Both biological processes and oceanographic features act to modify microbial beta diversity patterns between two or more sites or across temporal scales. These processes can be collectively understood as combinations of internal process including ecological determinism and stochasticity, and external processes orchestrating the resulting community assembly, and their collection of biological traits within an environment (e.g. Vellend et al. 2014; Fitzpatrick 2020). Deterministic processes include intrinsic processes and traits within the microbial community in response to niche properties such as nutrient acquisition or growth rates (here, referring to Hutchinsonian niche; (Hutchinson 1957), and microbial interactions with each other via grazing, parasitism, symbiosis and competition (Vellend and Agrawal 2010). Stochastic ecological processes include random colonization, and extinction and can be largely driven by external processes acting on the community such as horizontal dispersal (dispersal-limitation theory; MacArthur and Wilson 1967) and vertical mixing within the mixed layer (Cheng et al. 2020). Depending on environmental conditions, different combinations of deterministic, stochastic and external processes control microbial beta diversity (Vilmi et al. 2021).

Microbial biogeography offers a framework to reflect on the distribution of microorganisms in space and over geological time scales (Martiny et al. 2006). Marine

microbes disperse through advective transport (Villarino et al. 2018) following a ‘microbial conveyor belt’ over 1000 years that allows them to diversify under different local environmental conditions across different space and timescales (Mestre and Höfer 2021). Locally, they adapt to local environmental conditions such as temperature gradients that result in ‘breakpoints’ of taxonomic beta diversity patterns between polar and temperate algal microbiomes (Martin et al. 2021). The interplay between physical, chemical, and biological processes ultimately results in the formation of ocean provinces on scales from small coastal provinces such as the Canary coastal province to large ocean subtropical gyres (Longhurst 2007 ; **Box 3**, Fig. B2.1).

Longhurst provinces have been shown to overlap with the beta diversity patterns of microorganisms (e.g. Milici et al. 2016; Raes et al. 2018). However, microbes often have the potential to respond rapidly to environmental changes, resulting in small-scale dynamics and patchiness within ecosystems (Zwirgmaier et al. 2008; Zinger et al. 2011; Gibbons et al. 2013; Goodwin et al. 2017). The spatial scale of these ecosystem patches such as ocean filaments (Fadeev et al. 2021), fronts (Mousing et al. 2016), and eddies (Bolaños et al. 2020) are rarely observed across spatial scales of microbial community turnover time, and are only in a few studies directly linked to the intrinsic biological processes.

BOX 3: physico-biological scales and categorization in the ocean

To delineate the ocean’s seascape into large ecosystems, Longhurst (2007) defined ocean provinces based on physical (temperature, salinity) and biological (chl *a*) parameters (Fig. B3.1). The boundaries of these provinces are rationalized by the physical circulation of the ocean, and not the distribution pattern of individual species. Longhurst provinces are a widely used categorization of the ocean for biogeochemical models (Vichi et al. 2011) and phytoplankton bio-optical observations (Taylor et al. 2011).

Naturally, these boundaries are not static and need to be refined to resolve temporal variations of physical oceanographic features. Due to temporal and spatial constraints, Longhurst provinces do not resolve all dynamics that shape plankton distributions (Fig. B3.2). However, emerging interdisciplinary approaches and application of different sensors and observation methods allow more refined understanding of ecosystem dynamics across spatial and temporal scales (reviewed in Dickey 2003).

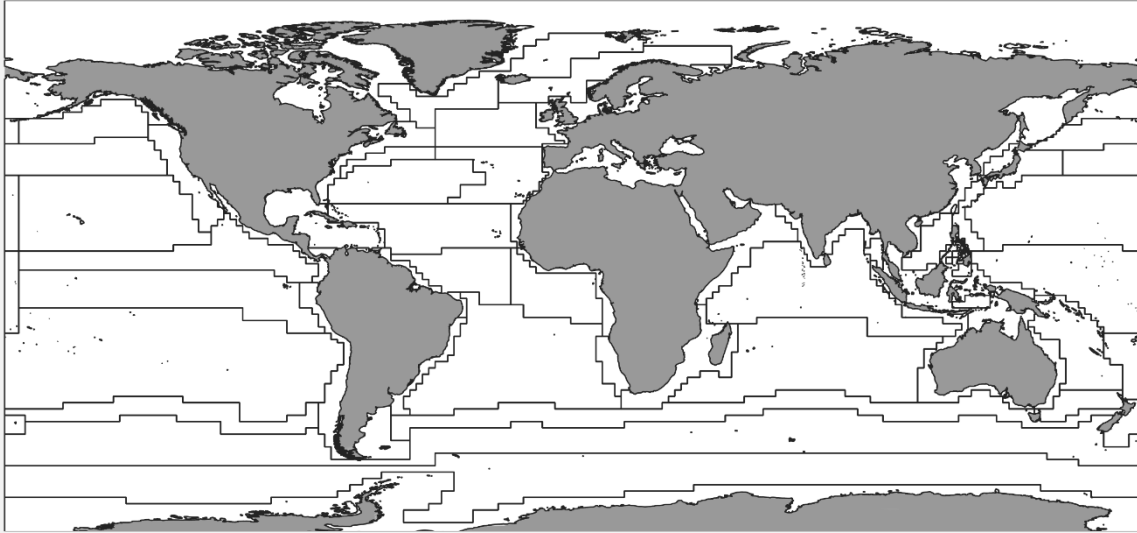
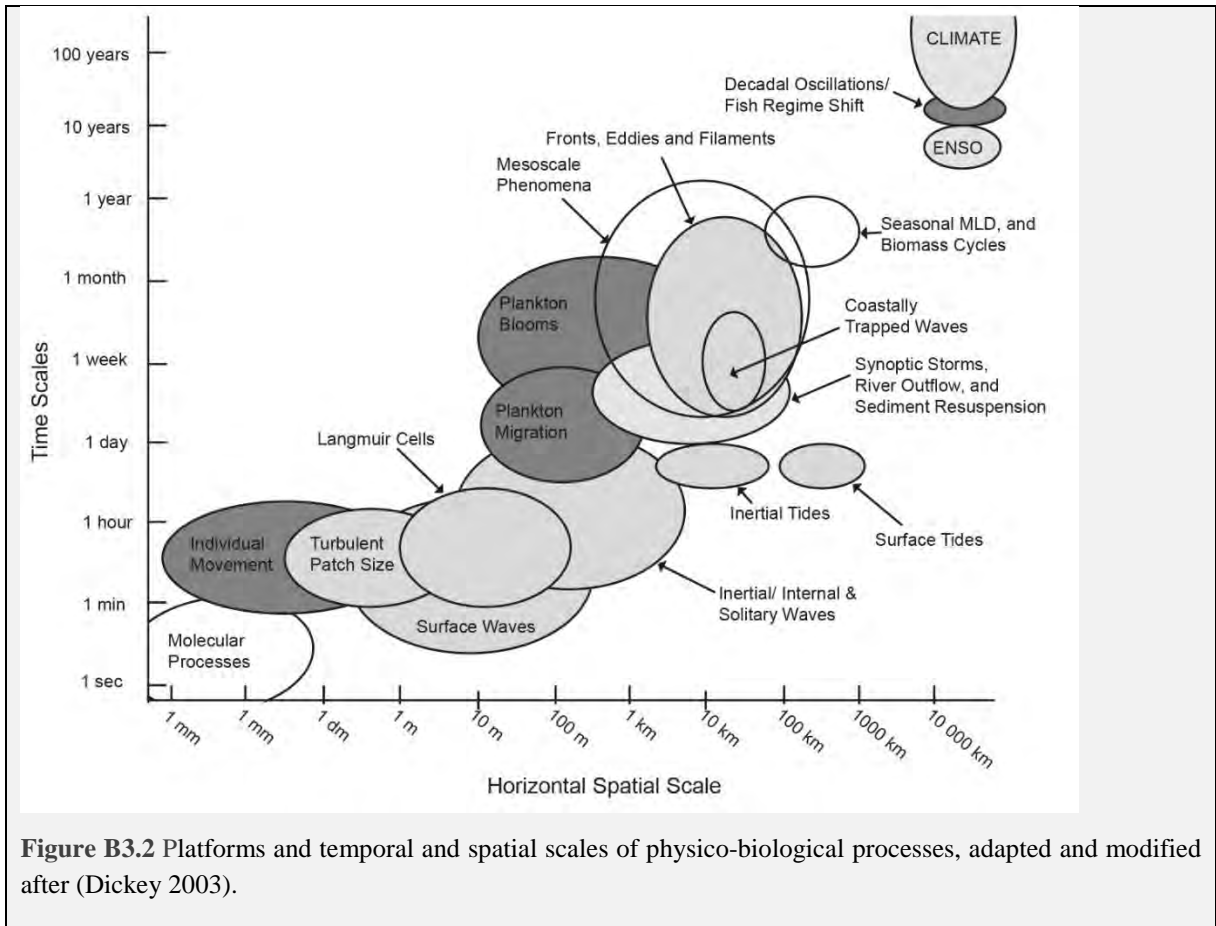


Figure B3.1 Longhurst provinces of the ocean (adapted from Flanders Marine Institute (2009). Longhurst Provinces version 4. Available online including extended version of individual names at <https://www.marineregions.org/>).



1.4 Horizontal scales of microbial biodiversity

1.4.1 Observed bio-physical horizontal scales in the surface ocean

The scales at which microbial diversity patterns vary also depend on the fluid connectivity of ocean systems (see Carr et al. 2003). Numerous beta diversity patterns have been described, such as latitudinal gradients (e.g. Hillebrand 2004; Fuhrman et al. 2008), distance-decay (Milici et al. 2016), species-abundance distributions (Matthews and Whittaker 2014; Locey and Lennon 2016), and temporal and spatial patterns (Preston 1960; Horner-Devine et al. 2004; Zhou et al. 2008). Therefore, to understand whether and how

microbial community composition is linked to a fundamental ecosystem function such as primary productivity requires systematic examinations across different environmental gradients, ocean regions and temporal scales. Dual-lense approaches (e.g. Oldham et al. 2013) that consider measurements of microbial communities along with chemical and physical contextual data, can help to shed light on the small and large scale pattern in this mosaic of the different temporal and spatial scales of biodiversity signals in the ocean.

On a small scale, a local community can experience a habitat change depending on both their internal biomass turnover and on the local/ regional hydrodynamics as well as the types of existing environmental gradients across which communities are transported (Jönsson and Watson 2016). Crump et al. (2004) measured the growth of microbial communities in the context of changing physical residence times along a salinity gradient in an estuary. They found seasonal changes mixing in communities (where residence time > growth rate), and the development of local communities at intermediate salinity (where residence time < growth rate). Similarly, a meta-analysis of 885 metagenome-assembled genomes revealed residence time as the 4th most important predictor of microbial beta diversity patterns in the global ocean (Faure et al. 2021). In the open ocean, understanding the timescales of advection in relation to those of intrinsic biological processes such as microbial growth and biomass turnover can help to assess sample patchiness, informing sampling scales and ultimately help us to better monitor marine ecosystem dynamics.

1.4.2 Methodological challenges that limit a global scaling of microbial observations

Monitoring ocean ecosystems effectively requires development of appropriate sampling scales for individual research campaigns and integrated analysis of multiple datasets. For example, the comparison of multiple datasets (Global Ocean Ship-based Hydrographic Investigations Program (BIO-GO-SHIP); <https://biogoship.org/>; TARA Oceans;

<https://oceans.taraexpeditions.org/>; GEOTRACES; <https://www.geotraces.org/>) that aligned in their methodological sampling design, and thus allowed new scientific insights into global patterns of ocean nutrient limitation (Ustick et al. 2021). Unfortunately, many ocean datasets have localized methodological adaptations and refinements, which are often not broadly accessible and/or well documented, which is the reason for the current reproducibility crisis (Lowndes et al. 2017).

Both the generated data and the underlying methodological complexity require structured, and consistent ways to enable global cross-comparisons of ocean observations. For ocean data FAIR (findable, accessible, interoperable, reproducible) data principles supported a culture of more open knowledge sharing in ocean observation (Tanhua et al. 2019). Specifically, in microbial observations this includes standardisation of essential metadata (e.g., Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS); Yilmaz et al. 2011), as well as reference material (e.g. mock communities, Parada et al. 2016). While the FAIR principles refer to data, it is also important to consider similar principles in method documentation.

Consensus on the optimal technical approach exists for a few large-scale campaigns such as the GO-SHIP report from the SCOR Working Group 154 (Boss et al. 2020) and networks such as the Genomic network for monitoring earth's ecosystems (Davies and Field 2012; Davies et al. 2014). However, a large number of localized methodologies (reviewed in Buttigieg et al. 2018) and organisational silos (reviewed in Révelard et al. 2021) persist; a situation that remains challenging in terms of developing openly accessible and reproducible methods that are benchmarked and inter-comparable with each other. Thus, strategies and guidelines are needed to make methods more complete and reproducible. A more open culture of method sharing and co-development could promote more inclusive science (Woodall et al. 2021) and the development of and consensus on globally interoperable, context-specific best practices, which will in turn enable us to study ocean systems at larger scales more confidently.

2. Objectives and hypotheses

In this thesis, I study microbial alpha and beta diversity patterns associated with changes in microbial mediated processes (specifically, primary productivity and N_2 fixation, as two important processes of marine C- and N-cycling, see **Box 2** for further details) in their hydrographic context. I resolve functional beta diversity through trophically disaggregated analyses assessed against biogeochemical gradients. I focus on surface oceanographic features such as fronts and currents and the advection of microbial communities therein. Specifically, I investigate the following five main objectives (**O**) of marine microbial ecology in the Atlantic and Indian Ocean as well as of Arctic and sub-Arctic Fjords (Fig. 2.1).

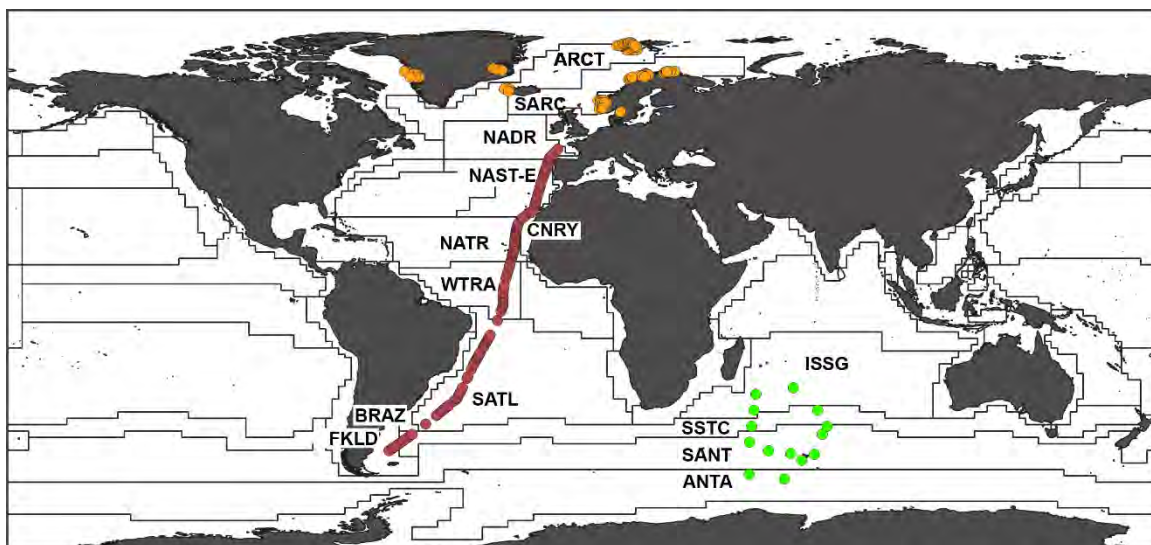


Figure 2.1. Datasets included in this thesis. Sampling sites analysed in Chapter 1 from of the MD206 expedition around Kerguelen island are indicated in orange; Sampling sites analysed in Chapter 2 from of the PS113 expedition across the Atlantic Ocean are indicated in red; Sampling sites analysed in Chapter 3 from of HE492, HE431, HE533, MSM21 and MSM56 expeditions to sub-Arctic and Arctic fjords are indicated in yellow. Province label are added for provinces studied

in this thesis: ANTA – Antarctic, ARCT – Atlantic Arctic, BRAZ – Brazilian Current coast, CNRY – Canary Current coast, FKLD – Falkland Current coast, ISSG – Indian South Subtropical gyre, NADR – North Atlantic drift, NAST-E – North Atlantic Subtropical gyre – East, NATR – North Atlantic Tropical gyre, SANT – Sub Antarctic water ring, SATL – South Atlantic Subtropical gyre, SARC – Atlantic Sub-Arctic, SSTC – South subtropical convergence, WTRA – Western Tropical Atlantic.

O1: Investigating temperature and primary productivity as two key structuring variables of microbial alpha diversity patterns in pelagic marine ecosystems.

Two general theories exist of microbial alpha diversity patterns: *Microbial diversity increases with increasing temperature across a latitudinal gradient (H1.1)* (Rohde 1992), and *higher microbial diversity is supported by increasing primary productivity (PP) until intermediate rates of PP (H1.2)* (Vallina et al. 2014) *within ocean provinces* (Raes et al. 2018). In **Chapter 1** and **Chapter 2**, I explored whether these previously described relationships of microbial diversity and their activity hold in the south Indian Ocean, and across a latitudinal transect in the Atlantic Ocean. Specifically, **Chapter 1** includes an observational study simultaneously assessing prokaryotic and eukaryotic diversity using 16S and 18S rRNA gene amplicon sequencing techniques combined with in situ rate measurements of PP and N₂ fixation, both microbially-mediated processes. In **Chapter 2**, I grouped the microbial 16S and 18S rRNA sequences into trophic functional groups based on literature research, a perspective that disentangles the direct and indirect relationships of microorganisms with PP and thus provides better insights into the food web and C-cycling.

O2: Mapping microbes: How can we better grasp the structure and function of microbial beta diversity in the surface ocean?

Longhurst provinces can be used to delineate pelagic microbiomes (Faure et al. 2021) *with clear province boundaries* (Raes et al. 2018) (**H2.1**). However, less understood is whether

these beta diversity patterns also reflect microbial activity, as (partial) functional redundancy has been previously observed in marine ecosystems (Louca et al. 2016). Therefore, I investigated the following objective: **(O2.1)** *How can we link microbial phylogeny to their activity within the framework of ocean provinces?* In order to link phylogeny to activity, I classified 16S and 18S rRNA gene sequences in trophic functional groups based on literature research, and measured primary productivity and N₂ fixation across the Indian Ocean (**Chapter 1**) and Atlantic Ocean (**Chapter 2**). In **Chapter 1**, I focused on how the beta diversity pattern, primary productivity and N₂ fixation changes across a boundary in the South Indian Ocean to better understand the role of boundary currents for the structure and function of microbiomes.

More practically, as microbes have fast turnover rates (hours to days; Martiny et al. 2006) they may also be good indicators to reflect change. This resulted in the following objective **(O2.2)** *Can we use microbial beta diversity patterns to refine the boundaries of these existing delineations?* In **Chapter 2**, I refined ocean provinces based on microbial beta diversity profiles. This approach can help to better identify appropriate boundaries for microbial ecosystems and thus inform about scales of microbial diversity change and to identify sampling schemes and potential sample patchiness. I used a combination of satellite observations, in situ measurements of temperature, salinity, dissolved inorganic nutrients (N, P, Si), particulate organic matter (C, N) and chlorophyll *a* concentrations for a more holistic approach of the different environmental variables shaping microbial diversity.

O3: Assessing the effect of regional advection by surface currents on microbial biodiversity patterns

Advection is a key determinant for microbial beta diversity patterns in the ocean (Jönsson and Watson 2016) **(H3.1)**. Yet, the spatial scales of microbial biodiversity patterns are inadequately quantified in relation to temporal dimensions that are relevant to biogeochemical cycles and microbial biomass turnover rates. In **Chapter 2**, we calculated

productivity-specific length scales for each station across a latitudinal transect in the Atlantic Ocean. The productivity-specific length scale reflects the potential length scale of horizontal displacement of a microbial community over the period of biomass turnover (i.e., the inverse of specific primary productivity). This provides insights into forces driving diversity patterns within and across ocean provinces. Additionally, in **Chapter 3**, we estimated microbial transport – and thus the oceanographic connectivity between sampling sites of multiple time points. Specifically, we estimated the spatial scales for dispersal over a month, three month, six month, a year, and 5 year periods, using a hydrodynamic model to map large ensembles of Lagrangian drifter trajectories. This model analysis helped to resolve *oceanographic connectivity* of microbial communities between Arctic and sub-Arctic fjords.

O4: Resolving the complexity of regional variabilities: spatial scales of microbial beta diversity patterns in coastal fjords

In comparison to open ocean systems, the heterogeneity and variability of physical dynamics in coastal systems are not well resolved in the more coarsely defined Longhurst provinces (e.g. highlighted in Hardman-Mountford et al. 2008). Specifically, the Arctic region is marked by fragmented coastlines with fjords ranging deep into the terrestrial systems and productivity sustained by the input from marine-terminating glaciers. Shaped through these strong land-sea exchange processes, I hypothesized that: **(H4.1)** *Microbial beta diversity is strongly shaped by the presence of marine-terminating glaciers*. Previous work identified comparable links in fjords of the Southern Hemisphere (Maturana-Martínez et al. 2021). In **Chapter 3**, I present an analysis of multiple Arctic and sub-Arctic fjords located in the Atlantic sector of the Arctic. The presence of marine-terminating glaciers in some of these fjords, and the absence of such glaciers in others, allows us to better understand (co-)occurrences of microbial taxa in these high-latitude environments.

O5: Enabling global scaling of ocean observations through consistent and reproducible method documentation

The cross-comparability and meta-analyses of microbial datasets (such as the ones used in this thesis) strongly depend on how and which information we can access. Of course, this is not only an issue in microbial observation but spans across all ocean observing disciplines. Understanding of the underlying methodological choices and experimental designs allows better inter-comparability of ocean data. However, many method descriptions are incomplete, inaccessible or even lost; sharing its part to the existing reproducibility crisis of ocean data (Lowndes et al. 2017). In **Chapter 4**, we provide a perspective on how to document ocean methods more transparently and reproducibly. Through this, we pave the way for enhanced communication, collaboration and accessibility of methods to support the ocean community in developing, testing, and adopting scientific methods in more diverse contexts. This also supports a more effective use of available technologies and archiving systems such as the Ocean Best Practices System (OBPS), ultimately allowing the scaling of scientific observations more globally.

3. Chapters

In this thesis, I demonstrate how functional activity can be decoupled from phylogenetic diversity through e.g. the environmental filtering of specific trophic functional groups in regions of high- and low chlorophyll *a*, respectively. Further, large variations in temporal and spatial scales impact the establishment of local communities and subsequently define ocean provinces. I show that beta diversity patterns reflect these ocean provinces and can in turn be used to refine ecosystem boundaries. In that context, the presented calculations of sample-wise productivity-specific length scales can help to identify sample patchiness and scale sample diversity in relation to different spatial levels of marine ecosystem structure. I apply these concepts through an integrative analysis of multiple Arctic and sub-Arctic fjords including multiple expeditions and sampling years and show how microbial communities disperse and form regional and within-fjord signals with different co-occurrence patterns between fjords with and without marine-terminating glaciers. Lastly, I provide perspectives on how consistent method documentation can make individual observations better reproducible, ultimately helping to improve inter-comparability of multiple observations and scaling of science more globally.

CHAPTER 1: Hydrographic fronts shape productivity, nitrogen fixation, and microbial community composition in the southern Indian Ocean and the Southern Ocean

Authors: Cora Hörstmann, Eric J Raes, Pier Luigi Buttigieg, Claire Lo Monaco, Uwe John, Anya M Waite

Abstract

Biogeochemical cycling of carbon (C) and nitrogen (N) in the ocean depends on both the composition and activity of underlying biological communities and on abiotic factors. The Southern Ocean is encircled by a series of strong currents and fronts, providing a barrier to microbial dispersion into adjacent oligotrophic gyres. Our study region straddles the boundary between the nutrient-rich Southern Ocean and the adjacent oligotrophic gyre of the southern Indian Ocean, providing an ideal region to study changes in microbial productivity. Here, we measured the impact of C and N uptake on microbial community diversity, contextualized by hydrographic factors and local physico-chemical conditions across the Southern Ocean and southern Indian Ocean. We observed that contrasting physico-chemical characteristics led to unique microbial diversity patterns, with significant correlations between microbial alpha diversity and primary productivity (PP). However, we detected no link between specific PP (PP normalized by chlorophyll *a* concentration) and microbial alpha and beta diversity. Prokaryotic alpha and beta diversity were correlated with biological N₂ fixation, which is itself a prokaryotic process, and we detected measurable N₂ fixation to 60°S. While regional water masses have distinct microbial genetic fingerprints in both the eukaryotic and prokaryotic fractions, PP and N₂ fixation vary more gradually and regionally. This suggests that microbial phylogenetic diversity is more strongly

bounded by physical oceanographic features, while microbial activity responds more to chemical factors. We conclude that concomitant assessments of microbial diversity and activity are central to understanding the dynamics and complex responses of microorganisms to a changing ocean environment.

Chapter 1 / 1 Introduction

The Southern Ocean (SO), in particular its sub-Antarctic zone, is a major sink for atmospheric CO₂ (Constable et al. 2014). The SO is separated from the Indian South Subtropical Gyre (ISSG) by the South Subtropical Convergence province (SSTC), comprising the Subtropical Front (STF) and the Subantarctic Front (SAF). The SSTC is a zone of deep mixing and thus elevated nutrient concentrations (Longhurst 2007). Further, the SSTC has been shown to act as a transition zone both numerically and taxonomically for dominant populations of marine bacterioplankton (Baltar et al. 2016).

In this dynamic context, a key driver of microbial productivity is nutrient availability, especially through tightly coupled carbon (C) and nitrogen (N) cycles. The constant availability of nutrients through vertical mixing in frontal zones, such as the STF, enhances primary productivity (Le Fèvre 1987) and chlorophyll *a* (chl *a*) concentrations (Belkin and O'Reilly 2009). Primary productivity (PP) and specific primary productivity (P^B, meaning primary productivity per unit chl *a*) are reflected in the relative abundance of different phytoplankton size classes whose productivity values are, in turn, stimulated by nutrient injections via shallowing of mixed layer depth (MLD) at the SO fronts (Strass et al. 2002); decreasing the possibility of N limitation. However, N limitation can also biologically be alleviated through N₂ fixation mediated by diazotrophs, significantly contributing to the N pool in oligotrophic regions (Tang et al. 2019). In high-latitude regions, biological N₂ fixation could potentially have a large impact on productivity (Sipler et al. 2017). However, large disagreements exist between models of high latitude N₂ fixation and its coupling to microbial diversity due to sparse sampling in these regions (Tang et al. 2019).

Due to the dynamics of the region, conflicting observations, and climate-driven changes, resolving the coupling of microbial productivity and diversity is particularly important across the strong environmental gradients crossing the ISSG, through the SSTC into the SO. Indeed, climate variability has been shown to impact ocean productivity and thus influences the provision of resources to sustain ocean life (Behrenfeld et al. 2006). To date, observations of climate-change-related effects in this region of the SO have been synthesized only based on long-term nutrient concentration and physical (temperature and salinity) changes (Lo Monaco et al. 2010); however, these typically lack a microbial dimension. Microbial composition, activity, and C export may all be impacted by climate-driven changes in ocean dynamics (Evans et al. 2011) such as MLD shallowing, eddy formation, and poleward shifts of ocean fronts (Chapman et al. 2020). For a more holistic ecosystem-based understanding of this region, concomitant assessments of (1) steady-state biogeochemical processes through rate measurements of key elements (such as C and N) and (2) the microbial diversity that underpins it are essential enhancements to such long-term investigations.

Here, we measure the impact of C and N uptake on microbial community diversity, alongside the effects of hydrography (e.g., dispersal limitation) and local physico-chemical conditions across the Southern Ocean and southern Indian Ocean. We focused our investigation on surface communities, aiming to resolve horizontal surface variation. We used our observation to assess whether the following relationships – previously observed in related systems – hold in our study region:

1. Microbial diversity increases with increasing primary productivity (PP). Previous work has claimed that more resources support higher species richness until intermediate rates of PP (Fig. C1.1; Vallina et al. 2014) within ocean provinces (Raes et al. 2018).
2. Frontal systems are discrete ecological transition zones between regions that provide perspectives on the findings of Baltar et al., (2016; see above). These systems often separate water masses with distinct trophic structures (e.g., Albuquerque et al., 2021).
3. Microbial alpha and beta diversity are impacted by N₂ fixation, which is itself correlated with the presence of other available sources of N and/or temperature; this is to provide more evidence on the role of N₂ fixation to the N budget in high latitudes (see e.g., Sipler et al., 2017; Shiozaki et al., 2018).

To our knowledge, there are no concomitant evaluations of how surface gradients, microbial activity, and community composition relate to one another in this region. Here, we provide perspectives on these key relationships across the Indian South Subtropical Gyre (ISSG), the Subtropical Front (STF), and Subantarctic Front (SAF), and the SO comprising the Polar Front (PF) and Antarctic Zone (AZ).

Chapter 1 | 2 Materials and methods

Chapter 1 | 2.1 Study region, background data, and sample collection

Our study region ranged from Réunion in the Indian South Subtropical Gyre (ISSG) to south of the Kerguelen Islands in the Southern Ocean (56.5°S, 63.0°E; Fig. C1.1a) as part of a larger repeated “OISO” sampling program – (Océan Indien Service d’Observations; Metzl and Lo Monaco, 1998; <https://doi.org/10.17600/17009700>). Samples were collected as part of the VT153/OISO27 (MD206) cruise aboard the R/V Marion Dufresne from 6 January to 7 February 2017.

Physical and biogeochemical data, as well as metadata, were collected from a rosette equipped with Niskin bottles and a conductivity, temperature, depth sensor (CTD) (Sea-Bird SBE32) equipped with a SBE43 O₂ sensor and a Chelsea Aqua tracker fluorometer. OISO long-term data, starting in 1998, were used as a backdrop to our data collected in 2018 and allowed us to monitor changes in physical and chemical oceanographic properties over time (Supplementary 1 File A).

Chapter 1 | 2.2 Province delineation after Longhurst

We identified three main clusters (i.e., ocean provinces) and five subclusters (i.e., water masses) on a temperature–salinity plot (Fig. C1.1b). As an overview, we used CTD depth profiles to validate the vertical extent of water masses in our samples (Fig. C1.1c, d) and checked the

horizontal extent of the identified clusters using remote sensing data of sea surface temperature (Supplementary 1 Fig. S1). Additionally, we checked the horizontal boundaries of these clusters for matches in strong chl *a* concentration gradients as an approximate for biological component of ocean provinces, following the concept of Longhurst (2007). Satellite data were acquired from MODIS (<https://neo.sci.gsfc.nasa.gov/>, last access: 16 June 2021), with images processed by NASA Earth Observations (NEO) in collaboration with Gene Feldman and Norman Kuring, i.e., NASA OceanColor Group (Supplementary 1 Fig. S2). We calculated the geodesic distance between sites from latitude and longitude coordinates using the geodist package in R (v0.0.4; Padgham et al., 2020).

Chapter 1 | 2.3 Nutrient analysis

Dissolved inorganic nutrient concentrations, including phosphate (PO_4), silicate (Si), mono-nitrogen oxides (NO_x), nitrite (NO_2), and ammonium (NH_4), were assayed on a QuAatro39 continuous segmented flow analyzer (Seal Analytical) following widely used colorimetric methods (Armstrong 1951; Murphy and Riley 1962; Wood et al. 1967) with adaptations to particular needs for Seal Analytical QuAatro autoanalyzer. NH_4 was measured using the fluorometric method of K  rouel and Aminot (1997). Detection limits of these methods were $0.1 \mu\text{mol L}^{-1}$ for PO_4 , $0.3 \mu\text{mol L}^{-1}$ for Si, $0.03 \mu\text{mol L}^{-1}$ for NO_x , and $0.05 \mu\text{mol L}^{-1}$ for NH_4 .

Chapter 1 | 2.4 Dissolved inorganic nitrogen and carbon assimilation

At each CTD station, water samples to measure primary productivity (PP) and N_2 fixation were taken from the underway flow-through system (intake at 7 m). As the ship was moving during sampling, the distance between samples of the same station can range up to 15 km. Incubations were performed in acid-washed polycarbonate bottles on deck at ambient light conditions. All polycarbonate incubation bottles were rinsed prior to sampling with 10% HCl (3x), deionized H_2O (3x), and sampling water (2x). In order to obtain the natural abundance of particulate nitrogen (PN)

and particulate organic carbon (POC), which we used as a t_0 value to calculate the assimilation rates, 4 L of water was filtered onto a 25mm pre-combusted GF/F filter for each station.

N_2 fixation experiments were carried out in triplicate for each station. We used the combination of the bubble approach (Montoya et al. 1996) and the dissolution method (Mohr et al. 2010) proposed by Klawonn et al., (2015). The 4.5 L bottles were filled up headspace free. All incubations were initialized by adding a $^{15}N_2$ gas bubble with a volume of 10 mL. We used $^{15}N_2$ -labeled gas provided by Cambridge Isotope Laboratories (Tewksbury, MA). Bottles were gently rocked for 15 min. Finally, the remaining bubble was removed to avoid further equilibration between gas and the aqueous phase. After 24 h, a water subsample was transferred to a 12mL exetainer® and preserved with 100 μ L $HgCl_2$ solution for later determination of exact ^{15}N – ^{15}N concentration in solution. Natural $^{15}N_2$ was determined using membrane inlet mass spectrometry (MIMS; GAM200, IPI) for each station with an average enrichment of 3.8 ± 0.007 at% $^{15}N_2$ (mean SD; $n = 104$). Primary productivity was measured by adding $Na^{13}CO_3$ at a final ^{13}C concentration of 200 μ mol L^{-1} . Incubation bottles were incubated on board at ambient sea surface temperature (SST; water intake at 7 m) using a continuous-flow-through system. Temperature of both incubation bins was continuously measured. After 24 h, the C and N_2 fixation experiments were terminated by collecting the suspended particles from each bottle by gentle vacuum filtration through a 25 mm pre-combusted GF/F filter (< 10 kPa). Filters were snap-frozen in liquid nitrogen and stored at -80 °C while at sea. Filters with enriched (T24) and unenriched (T0) samples were acidified and dried overnight at 60 °C. Analysis of ^{15}N and ^{13}C incorporated was carried out by the isotopic laboratory at the University of California, Davis, California campus, using an Elementar Vario EL Cube or MICRO cube elemental analyzer (Elementar Analysensysteme GmbH, Hanau, Germany).

Carbon assimilation rates were calculated according to Knap et al., (1996), excluding the ^{14}C – ^{12}C conversion factor, and N_2 fixation was calculated according to Montoya et al., (1996). The minimum quantifiable rate was calculated according to Gradoville et al., (2017).

Chapter 1 | 2.5 Pigment analysis

For pigment analyses, 4 L of seawater was filtered (<10 kPa) on a 47mm Whatman GF/F filter and stored at -80°C until further analysis. High-performance liquid chromatography (HPLC) was carried out as described in Kiliyas et al., (2013) with the following modifications: 150 μL of the internal standard canthaxanthin was included to each sample. Samples were dissolved in 4 mL acetone and disrupted with glass beads using a Precellys 24 tissue homogenizer (Bertin Technologies, France) at 7200 rpm for 20 s. Detection of the sample at 440 nm absorbance was performed using an HPLC analyzer (Varian Microsorb-MV 100-3 C8). We used chl *a* concentration to estimate phytoplankton biomass. Pigment concentrations were calculated according to Kiliyas et al., (2013) and quality controlled according to Aiken et al., (2009) (Supplementary 1 File A). HPLC output data were analyzed using diagnostic pigments for the different taxa and phytoplankton functional types (PFTs) after Hirata et al., (2011) (Supplementary 1 File A, Table S1). This approach can be used to reveal dominant trends of the phytoplankton community and size structure at the regional and seasonal scales (Ras et al. 2008). Furthermore, diagnostic pigments were used to delineate three different size classes (pico-, nano-, and microplankton) according to Vidussi et al., (2001). The relative proportion of each phytoplankton size class (PSC) was calculated based on the linear regression model proposed by Uitz et al., (2006). We investigated the patterns of PSCs with a second-order polynomial fit.

Chapter 1 | 2.6 DNA analysis

Two liters of seawater from the shipboard underway system from each station were filtered through a 0.22 μm Sterivex® filter cartridge for DNA isolation, snap-frozen in liquid nitrogen, and stored at -80°C . DNA was extracted using a DNeasy® Plant Mini Kit (QIAGEN, Valencia, CA, USA, catalog no. or ID 69106) following the manufacturer's instructions. Sterivex cartridges were gently cracked open, and filters were removed and transferred into a new and sterile screw-cap tube. Approximately 0.3 g of pre-combusted glass beads (diameter 0.1 mm; 11079101 Bio Spec Products) and 400 μL buffer AP1 were added to the filter, followed by a bead beating step using a Precellys 24 tissue homogenizer (Bertin Technologies, France), with two times at 5500 rpm for

20 s with 2 min on ice in between and a final bead beating step at 5000 rpm for 15 s. DNA concentrations were quantified by the Quantus™ fluorometer and normalized to 2 ng μL^{-1} .

Amplicon 16S and 18S rRNA gene PCR and sequencing

Amplicons of the bacterial 16S rRNA gene and eukaryotic 18S rRNA gene (using primers from 27F–519R; (Parada et al. 2016), TA-Reuk454FWD1–TAREukREV3; (Stoeck et al. 2010), respectively) were generated following standard protocols of amplicon library preparation (16S Metagenomic Sequencing Library Preparation, Illumina, part no. 15044223 Rev. B). The 16S and 18S rRNA gene PCR products were sequenced using 250 bp paired-end sequencing with a MiSeq sequencer (Illumina) at the European Molecular Biology Laboratory (EMBL) in Heidelberg (Germany) and at the Leibniz Institute on Aging (FLI) in Jena (Germany), respectively.

Amplicon sequence data analysis

For both 16S rRNA gene and 18S rRNA gene amplicon sequences, we used the DADA2 R package, v1.15.1 (Callahan et al. 2016) to construct amplicon sequence variant (ASV) tables by the following steps: prefiltering “filterandtrim” function with truncLD50 and default parameters. Primer sequences were cut using the Cutadapt software implementation (v1.18) in the DADA2 pipeline, removing a fixed number of bases matching the 16S forward (515F-Y, 19 bp) and reverse (926R, 20 bp) primers, and the 18S forward (TAREuk454FWD1, 20 bp) and reverse (TAREukREV3, 21bp) primers. Primer trimmed fastq files were quality trimmed with a minimum sequence length of 50 bp and checked by inspection of the average sequence length distribution (for both the 16S rRNA gene and 18S rRNA gene sequences). Samples within forward and reverse fastq files were dereplicated and merged with a minimum overlap of 20 bp. ASV tables were constructed, and potential chimeras were identified de novo and removed using the “removeBimeraDenovo” command. Sequencing statistics for removed reads and sequences in each step can be found in Supplementary 1 Table S2. Taxonomic assignment was performed using the SilvaNGS (v1.4; Quast et al., 2013) pipeline for 16S rRNA gene data with the similarity threshold set to 1. Reads were aligned using SINA v1.2.10 (Pruesse et al. 2012) and classified using BLASTn (v2.2.30; Camacho et al., 2009) with the Silva database (v132) as a reference database. For taxonomic assignment of 18S rRNA gene amplicons, we used the plugin “feature-classifier” (from package “q2-feature-classifier”, v2019.7.0) in QIIME2 (Bokulich et al. 2018) and the pr2 database

(v4.12; Guillou et al., 2013). We removed ASVs annotated to mitochondria and chloroplasts from 16S rRNA gene ASV tables and ASVs annotated as metazoans from 18S rRNA gene ASV tables. ASV tables of 16S rRNA gene amplicons, and 18S rRNA gene amplicons were used for further statistical analyses.

Chapter 1 | 2.7 Ecological data and statistical analysis

A combination of temperature, salinity, dissolved oxygen concentrations, and dissolved inorganic nutrient concentrations (NO_3 , NO_2 , NH_4 , Si, and PO_4) were used to characterize the physical and biogeochemical environment of the study region.

Statistical documentation, package citations, and scripts are publicly archived under CoraHoerstmann/MD206_Microbes (<https://www.doi.org/10.5281/zenodo.5779517>). All statistical tests were performed in R version 3.6.3 (R Core Team, 2020).

Microbial alpha diversity was calculated with Hill numbers (richness, Shannon entropy, inverse Simpson, $q = 0-2$; Chao et al., 2014) using the iNEXT package v2.0.20 in R with confidence set to 0.95 and bootstrap = 100 (MD206_Microbes/alpha_diversity). Accordingly, rarefaction curves are shown in Supplementary 1 Fig. S3. Pearson correlations between microbial richness ($q = 0$), inverse Simpson diversity ($q = 2$), environmental parameters, and biological rates were calculated and plotted (ggplot2) (Supplementary 1 Fig. S4). The p values were adjusted for multiple testing using Holm adjustment (Holm 1979), and residuals were checked for normal distribution (Supplementary 1 Fig. S5). For comparability and statistical downstream analyses, we performed the following transformations to the ASV table and the environmental metadata: to account for the compositionality of sequencing data (see Gloor et al., 2017), we performed a centered log ratio (CLR)-transformation for redundancy analysis (RDA). We used Hellinger transformation (*decostand()* function in vegan) of the ASV pseudocount data (minimum pseudocount per ASV cutoff was 3) for PERMANOVA analyses. Environmental data were z-scored for comparable metadata analysis (S1_code_archive/transformations). For multivariate analyses of microbial beta diversity and environmental parameters, we performed redundancy analyses (RDA) of the CLR-

transformed ASV tables (MD206_Microbes /RDA). Differences of microbial beta diversity (based on Hellinger transformed ASV tables), phytoplankton community composition (based on pigment concentrations), and water masses were tested with permutational ANOVA (PERMANOVA; Anderson, 2001) using the *adonis2()* function in *vegan* along with a beta dispersion test to evaluate the homogeneity of the dispersion. To investigate where differences of environmental variables have an impact on microbial community dissimilarity, we performed a general dissimilarity model (GDM) of the community dissimilarity and environmental variables, and we checked for the influence of geographic distance based on spline magnitude (*gdm* package; MD206_Microbes/GDM). As differences in microbial beta diversity were significant in PERMANOVA between provinces and water masses, we performed a similarity percentage (SIMPER) analysis in R using the *vegan* package to assess which ASVs contribute most to the observed variance of microbial community composition (Supplementary 1 Table S3; MD206_Microbes/taxonomy_analyses). To determine the number of ASVs shared between provinces (or unique to certain provinces), we transformed ASV pseudocount tables into binary tables and calculated shared and unique ASVs using the *upsetR* package in R (v.4, Conway et al., 2017; MD206_Microbes/upsetR). We calculated the percentage of all within-sample-observed ASVs within the merged samples of a province (Supplementary 1 Table S4).

Chapter 1 / 3 Results

Chapter 1 | 3.1 Delimitation of regional water masses

Through our analysis of temperature, salinity, oxygen, and dissolved inorganic nutrient (N, P, Si) concentrations, we identified five distinct water masses, fronts, and frontal zones: the ISSG, STF, SAF, PFZ, and AZ, which broadly aligned with three oceanographic provinces (ISSG, SSTC, and SO; Fig. C1.1a). Within the Southern Ocean (SO), we identified four water masses in our transect including the Antarctic Zone (AZ) and three distinct frontal systems: (1) the Polar Front (PF), (2) the Subantarctic Front (SAF), and (3) the Subtropical Front (STF; Fig. C1.1). In our analysis,

stations 6, 7, and 9 were placed within the Polar Front Zone (PFZ), which is between the SAF and PF. Due to the bathymetrically driven convergence of the STF and SAF around Kerguelen island, we consider the SAF as part of the convergence zone between the SO and Indian Ocean (IO), i.e. the South Subtropical Convergence province (SSTC), rather than as a Southern Ocean frontal system. At 7 m depth, we noted clear shifts in temperature (SST), salinity, and dissolved inorganic nutrient (NO_3 , PO_4 , Si) concentrations when crossing the STF. The STF is described as a circumpolar frontal zone creating the boundary between our measurements of the warm (20–25 °C), saline (> 35), and oligotrophic ($\text{NO}_3 < 0.03 \mu\text{M}$; $\text{PO}_4 : 0.04\text{--}0.21 \mu\text{M}$) subtropical waters (STW) of the Indian South Subtropical Gyre (ISSG) and the cold (3–6 °C) macronutrient-rich SO ($\text{NO}_3 : 19.2\text{--}24.9 \mu\text{M}$; $\text{PO}_4 : 1.43\text{--}1.71 \mu\text{M}$) (Figs. C1.1, C1.2, Supplementary 1 Fig. S2). In the context of this study, STW and ISSG could be used interchangeably; we henceforth refer to it as ISSG.

Chapter 1 | 3.2 Primary productivity (PP)

Maximum primary productivity (PP) values within our dataset were measured near the Kerguelen Plateau in the Polar Front Zone (PFZ) at station 9 (3236.8 and 3553.3 $\mu\text{mol C L}^{-1} \text{d}^{-1}$, respectively) and station E (2212.4–2688.1 $\mu\text{mol C L}^{-1} \text{d}^{-1}$, $n = 6$). Comparing all PP measurements across water masses, we found relatively high PP in other stations of the PFZ (stations 6, 7; Fig. C1.3a; Table C1.1) and in the Subantarctic Front (SAF) (stations 4, 15). Lowest PP values (190.4–642.6 $\mu\text{mol C L}^{-1} \text{d}^{-1}$) were measured at the stations in the Indian South Subtropical Gyre (ISSG). While stations in the ISSG showed very little variations within one station (e.g., 226.09–371.07 $\mu\text{mol C L}^{-1} \text{d}^{-1}$, $n = 6$, station 18), variation within SO stations was relatively high (e.g., 587.42–1875.58 $\mu\text{mol C L}^{-1} \text{d}^{-1}$, $n = 6$, station 37; Table C1.1).

Overall, the variation of specific primary productivity (P^B) did not show great variations between provinces, with maximum rates at station 11 (Table C1.1; Fig. C1.3b). We did not find a significant correlation between mixed layer depth and P^B (Pearson correlation: $r = 0.21$, $p = 0.47$, $n = 12$).

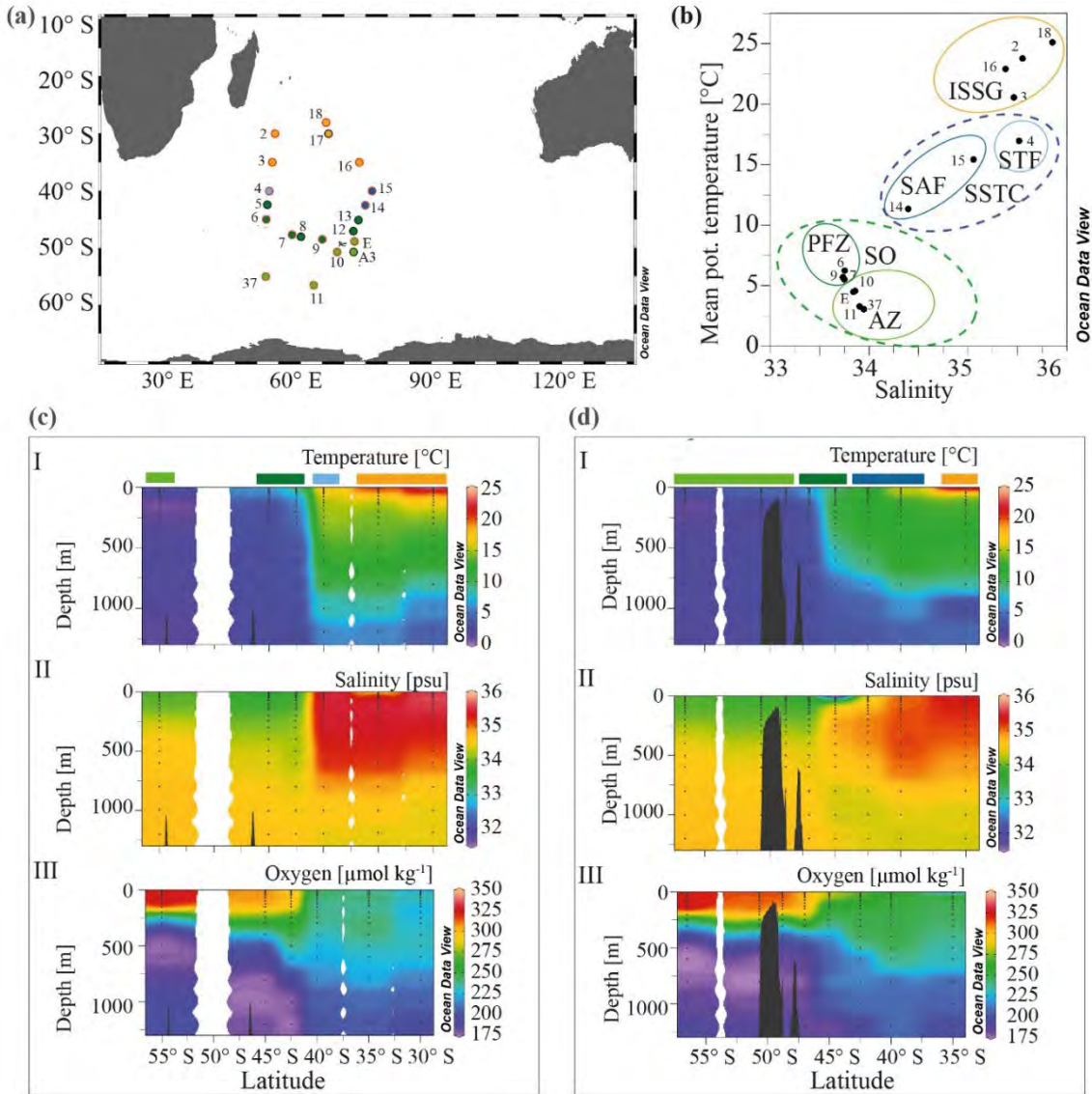


Figure C1.1. (a) The MD206 transect and OISO stations. Stations are colored according to water masses and encircled by sampling extent: black circles indicate stations where only CTD (conductivity, temperature, depth) data are provided, and stations encircled in red denote where additional samples for C, N, and community composition were taken. (b) A plot of potential temperature (in degrees Celsius (°C)) and salinity (in practical salinity units) using sea surface (7 m) data of the stations used in further microbial and C/N analyses. The yellow circle highlights the Indian South Subtropical Gyre (ISSG), light blue circle the Subtropical Front (STF), blue circle the Subantarctic Front (SAF), dark green circle the Polar Front Zone (PFZ) and the light green circle indicates the Antarctic Zone (AZ); dashed lines indicate water masses clustered within ocean provinces: the blue line marks the South Subtropical Convergence province (SSTC), and the green line marks the Southern Ocean (SO); panels (c) and (d) show depth profiles of temperature, oxygen, and salinity along two transects of the OISO stations. Colored bars indicate water masses according to (b). Panel (c) shows the western transect covering OISO stations 2, 3, 4, 5, 6, and 37 around 53±1°E longitude; panel (d) shows the eastern transect of OISO stations 10, 11, 12, 13, 14, 15, 16, and E around 68±5°E.

Table C1.1. Sampling stations visited during the OISO27 cruise, including chlorophyll *a* concentrations, primary productivity (PP), specific primary productivity (P^B), and N_2 fixation. Mixed layer depth (MLD) was calculated using $\Delta d = 0.03 \text{ kg m}^{-3}$ compared to a surface reference depth of 7 m. NA indicates no data. Ranges and mean for sample replicates of N_2 fixation and PP are given ($n = 3$ for stations 3, 9, 11, 15; $n = 6$ for stations E, 37, 2, 4, 6, 7, 14, 16, 18). Colored bars indicate water masses; yellow bar highlight the Indian Ocean gyre (ISSG), light blue bar the Subtropical Front (STF), blue bar the Subantarctic Front (SAF), green bar the Polar Front zone (PFZ), and dark green the Antarctic Zone (AZ).

Station	Longitude °E	Latitude °S	MLD [m]	chl <i>a</i> [$\mu\text{g L}^{-1}$]	Primary productivity (PP) [$\mu\text{mol C L}^{-1} \text{ d}^{-1}$]	specific PP (P^B) [$\mu\text{mol C } \mu\text{g chl } a^{-1} \text{ L}^{-1} \text{ d}^{-1}$]	N_2 fixation [$\text{nmol L}^{-1} \text{ d}^{-1}$]	MQR [$\text{nmol L}^{-1} \text{ d}^{-1}$]
37	52.003	55.004	52.5	0.38	587.42 - 1875.58; 1185.59	1562 - 4988; 3153	0.76 - 3.09; 1.97	1.2
11	63.006	56.499	49.5	0.18	1020.91 - 2065.12; 1541.95	5640 - 11409; 8519	0.23 - 2.20; 0.89	1.2
10	68.421	50.667	88.2	0.15	NA	NA	NA	NA
E	72.367	48.8	81.3	0.83	2212.37 - 3114.53; 2645.72	2662 - 3748; 3184	0.18 - 2.09; 0.92	0.7
7	58.004	47.667	49.6	0.58	942.99 - 4305.26; 2129.45	1634 - 7460; 3690	1.0 - 4.39; 1.75	1.2
9	64.999	48.501	69.4	0.34	3236.8 - 3553.33; 3395.07	9568 - 10503; 10035	0.19 - 2.15; 0.88	0.8
6	52.102	45.000	41.7	0.40	676.44 - 4242.33; 1977.6	1671 - 10481; 4886	0.17 - 3.25; 0.93	0.9
14	74.884	42.499	30.8	0.64	994.1 - 3847.07; 2635.94	1548 - 5994; 4107	0.0 - 2.26; 0.78	0.7
15	76.407	39.999	29.8	0.63	1579.92 - 2341.93; 1884.88	2514 - 3727; 2999	0.0 - 1.43; 0.24	1.2
4	52.79	40.001	54.6	0.42	524.32 - 1876.67; 1069.21	1258 - 4504; 2566	0.11 - 4.91; 2.01	3.5
3	53.499	35.000	15.9	0.25	350.33 - 845.86; 642.59	4787 - 11559; 8781	0.65 - 6.91; 2.81	5.4
16	73.466	35.001	19.9	0.05	170.05 - 537.91; 378.28	3483 - 11018; 7748	0.39 - 2.21; 1.05	1.3
2	54.1	30.001	12.9	0.06	63.24 - 324.72; 190.38	1003 - 5152; 3021	0.7 - 2.88; 1.58	2.6
18	65.832	28.0	16.9	0.04	226.09 - 371.07; 301.38	5406 - 8873; 7206	0.94 - 7.92; 4.04	5.0

Chapter 1 | 3.3 N_2 fixation

Di-nitrogen (N_2) fixation was above the minimum quantifiable rate (MQR) at all stations (Table C1.1). N_2 fixation measurements did not show a clear temperature-dependent trend (Fig. C1.3), and neither were they directly associated with low dissolved inorganic nutrient (DIN) values (Supplementary 1 Fig. S9). N_2 fixation in the warm oligotrophic waters of the Indian South Subtropical Gyre (ISSG) was up to $7.93 \text{ nmol N L}^{-1} \text{ d}^{-1}$ (station 18; Fig. C1.3c; Table C1.1). Lowest N_2 fixations were measured in the productive zone of the STF and SAF ($0.24\text{--}2.01 \text{ nmol N L}^{-1} \text{ d}^{-1}$, $n = 3$). In the AZ, N_2 fixation ranged between 0.89 and $1.97 \text{ nmol N L}^{-1} \text{ d}^{-1}$. The variation between replicates was high; e.g., rates ranged between 0.9 and $7.9 \text{ nmol N L}^{-1} \text{ d}^{-1}$ at station 18 (Table C1.1). Across provinces, we did not find notable differences in N_2 fixation.

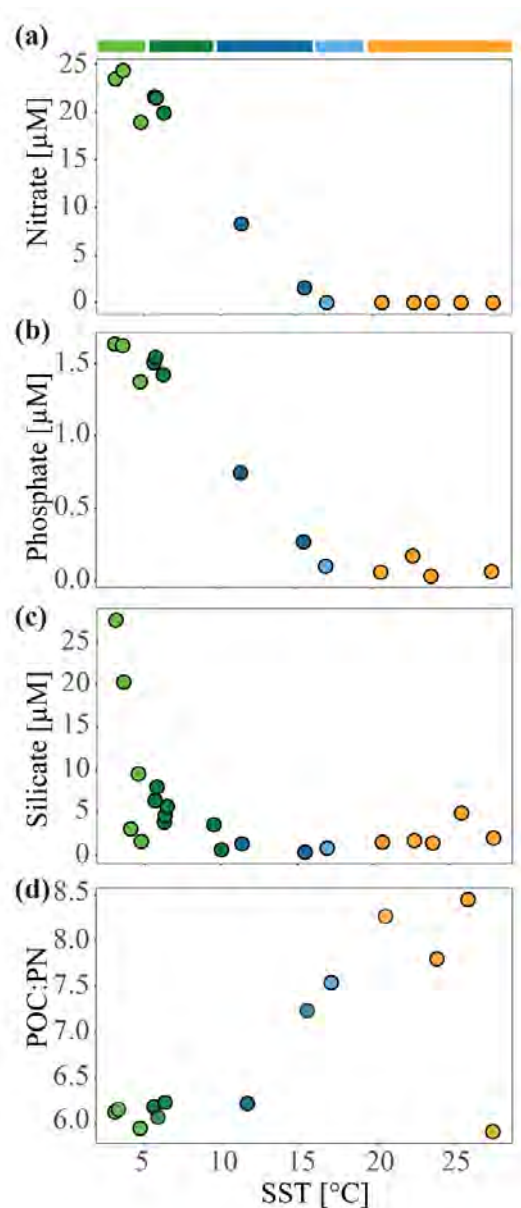


Figure C1.2. Nutrient concentrations ($\mu\text{mol L}^{-1}$) and molar ratios of particulate organic carbon (POC) to particulate nitrogen (PN) during the MD206 expedition against sea surface temperature ($^{\circ}\text{C}$): (a) nitrate, (b) phosphate, (c) silicate, and (d) POC : PN ratio. Colored bars indicate water masses according to their sea surface temperature: yellow bar highlights the Indian South Subtropical Gyre (ISSG), light blue bar highlights the Subtropical Front (STF), blue bar highlights the Subantarctic Front (SAF), dark green bar highlights the Polar Front Zone (PFZ), and light green bar highlights

Chapter 1 | 3.4 Phytoplankton pigment analyses

Photosynthetic pigment concentrations showed a clear separation between the oligotrophic ISSG and the nutrient-rich SO (Supplementary 1 Fig. S6). Chlorophyll *a* concentrations were relatively low in the warmer water stations of the ISSG than in the SSTC and SO (Table C1.1). The relative proportion of phytoplankton biomass to the total organic matter was estimated by calculating the

ratio of PN : chl *a* and showed a strong increase in the ISSG (11.5–29.7 PN : chl *a*, $n = 4$) in comparison to the SSTC (2.7–7.2 PN : chl *a*, $n = 3$) and SO (2.8–15.3 PN : chl *a*, $n = 6$; Supplementary 1 Fig. S7). The phytoplankton community composition was significantly and markedly different across provinces (PERMANOVA; permutations = 999, $R^2 = 0.76$, $p < 0.001$; $n = 14$) and water masses (PERMANOVA; permutations = 999, $R^2 = 0.81$, $p = 0.002$, $n = 14$). The pigment concentration of prokaryote-specific pigment zeaxanthin was high in the ISSG (0.03–0.06 mg m⁻³, $n = 4$; Supplementary 1 Fig. S6a). Zeaxanthin still occurred in the STF and SAF (0.03–0.04 mg m⁻³, $n = 3$) but disappeared in the SO (< 0.01 mg m⁻³, $n = 6$). *Prochlorococcus* was distinctly identified through its diagnostic pigment, divinyl chl *a*, and showed a relatively high pigment concentration in the ISSG (0.02–0.03 mg m⁻³, $n = 4$; Supplementary 1 Fig. S6a). We found concentrations of diatom-specific fucoxanthin (except station 18) ranging from 0.021 mg m⁻³ in the ISSG (station 16) to 0.34 mg m⁻³ in the SO (station 37; Supplementary 1 Fig. S6a). Across water masses, fucoxanthin concentration was slightly higher in the AZ (0.06–0.5 mg m⁻³, $n = 4$) than in all other water masses (0–0.13 mg m⁻³, $n = 10$). The distribution of potential phytoplankton size classes (PSCs; pico- nano- and microplankton), calculated from diagnostic pigments (Supplementary 1 File A), showed a clear pattern over temperature variations (Supplementary 1 Fig. S6b). The pigment data suggested that picoplankton dominated warm water in the ISSG, and picoplankton abundance sharply decreased (second-order polynomial fit: $R^2 = 0.98$, $p < 0.001$, $n = 14$) at lower values of SST. Pigment data also suggested that microplankton showed a contrary trend to the relative fraction of picoplankton, having high abundance in cold water and decreasing at higher values of SST, with a minimum at 20 °C SST and a slight increase (14% microplankton of all phytoplankton size classes) towards 25 °C SST (second-order polynomial fit: $R^2 = 0.77$, $p < 0.001$, $n = 14$). Nanoplankton showed a maximum at 12 °C SST and decreased both towards warmer and colder waters (second-order polynomial fit, $R^2 = 0.58$, $p < 0.01$, $n = 14$).

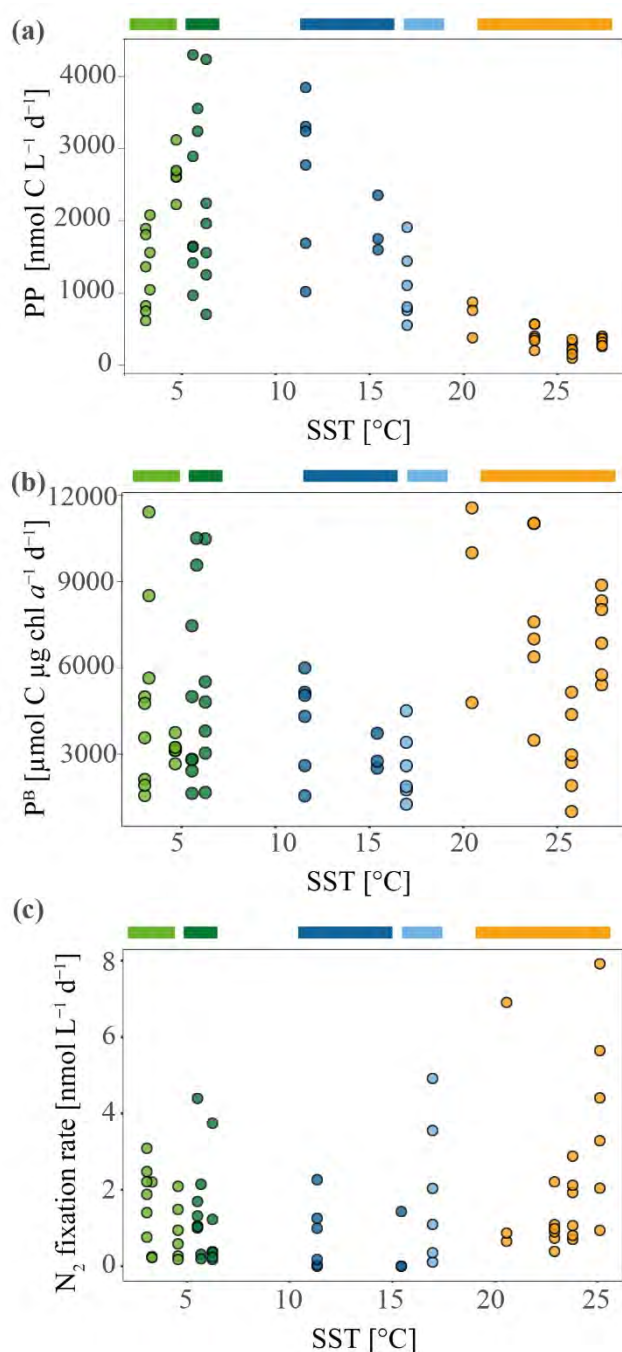


Figure C1.3. Primary productivity (PP) and specific primary productivity (P^B) measured during the MD206 cruise. (a) PP in micromole carbon per liter per day against sea surface temperature (SST) in degrees Celsius (°C). (b) P^B, normalized by chl *a* concentration. (c) Nitrogen fixation rates against sea surface temperature (SST) in degrees Celsius measured during the MD206 cruise. Rates are shown in nanomole nitrogen per liter per day. Colored bars indicate water masses: yellow bar highlights the Indian South Subtropical Gyre (ISSG), light blue bar highlights the Subtropical Front (STF), dark blue bar highlights the Subantarctic Front (SAF), dark green bar highlights the Polar Front Zone (PFZ), and light green bar marks the Antarctic Zone (AZ).

Chapter 1 | 3.5 Eukaryotic planktonic community composition

For each station, except station 4, the V4 region of the small subunit ribosomal RNA gene (18S rRNA) was amplified and sequenced to determine the community composition of micro-, nano-,

and pico-eukaryotes in all three oceanic provinces. We recovered a total of 2618 ASVs. After removing sequences annotated to metazoans, 2501 ASVs remained (4.4% of ASVs removed).

We found a strong correlation between both eukaryotic richness and diversity (inverse Simpson index) with SST (Pearson correlations: $r = 0.85$, $p < 0.001$ for richness and $r = 0.82$, $p = 0.001$ for inverse Simpson, $n = 12$; Supplementary 1 Fig. S4a, c). Overall, eukaryotic diversity was negatively correlated with PP ($r = -0.66$, $p = 0.02$, $n = 12$; Supplementary 1 Fig. S4e) and significantly and positively associated with N_2 fixation ($r = 0.74$, $p = 0.01$, $n = 12$; Supplementary 1 Fig. S4g). However, a strong correlation between rate measurements (PP, N_2 fixation) and eukaryotic diversity was only apparent in the ISSG, and no significant trend across other provinces (Pearson correlation after removal of ISSG samples from dataset: for PP $r = 0.47$, $p = 0.24$ and for N_2 fixation $r = -0.48$, $p = 0.23$, $n = 8$). Our RDA constrained 81% of the variance in the ASV table, with a p value of 0.095 (permutations = 999, $n = 12$). Sites were well separated between Longhurst provinces along the first two RDA axes (capturing 52.67% constrained variance, Fig. C1.4a). Our PERMANOVA, which tested the province-based separation, produced moderate but significant results (permutations = 999, $R^2 = 0.54$, $p = 0.001$, $n = 12$). An additional PERMANOVA grouping sites by water masses produced similar results (permutations = 999, $R^2 = 0.67$, $p = 0.001$, $n = 12$; Fig. C1.4a). We found that more ASVs only occurred in one province rather than in two or more provinces (Fig. C1.4e). Sites within the ISSG province were associated with SST and N_2 fixation. Sites in the SSTC were associated with high NH_4 concentrations. Sites belonging to the SO were associated with dissolved inorganic nutrients (NO_3 , PO_4 , Si), dissolved oxygen, and chl *a* concentrations as well as high PP. Linear relationships between beta diversity and rates were only weak for PP (PERMANOVA; permutations = 999, $R^2 = 0.27$, $p = 0.004$, $n = 12$) and both weak and insignificant between beta diversity and N_2 fixation (PERMANOVA; permutations = 999, $R^2 = 0.13$, $p = 0.14$, $n = 12$). Investigating whether and at which magnitude environmental parameters have an effect on microbial community dissimilarity, our general dissimilarity model (GDM) showed the expected curvilinear relationship between the predicted ecological distance and community dissimilarity (Fig. C1.4c I). Based on I-spline magnitudes of all tested environmental variables, geographic distance had little effect on community dissimilarity (Supplementary 1 Fig. S8a). Community dissimilarity changed most notably in response to variability in low magnitudes of PP (i.e., ISSG and STF; 17% of total community dissimilarity, $n = 12$) and plateaued with PP

above $1100 \mu\text{mol C L}^{-1} \text{ d}^{-1}$ (Fig. C1.4c III). A community dissimilarity change occurred most notably when N_2 fixation rates were above $2 \text{ nmol N L}^{-1} \text{ d}^{-1}$ (19% of change in total community dissimilarity associated with changes in N_2 fixation rates) (Fig. C1.4c IV). Among all tested environmental parameters, our I-spline results showed that community dissimilarity increased most in response to variability in MLD and PO_4 concentrations (49% of change in total community dissimilarity associated with MLD variability and 63% with PO_4 variability, $n = 12$; Supplementary 1 Fig. S8a). Significant differences in community dissimilarity structure between Longhurst provinces were associated with high-pseudocount taxa, dominated by dinoflagellates (Dinophyceae) and diatoms (Bacillariophyta; SIMPER analysis; Supplementary 1 Table S3). The pseudocount of ASVs belonging to the phylum Ochrophyta (Bacillariophyta_X) contributed to differences between ocean provinces (contributing to at least 9.51% of the differences in community dissimilarity between the SO and ISSG). Moreover, 4.79% of the differences in community dissimilarity between the SO and the SSTC were associated with a higher ASV count of Bacillariophyta_X ASVs in the SO. Further, we identified 10 ASVs belonging to the phylum Dinophyceae, contributing 2.1% to the community dissimilarity structure between the SO and ISSG and 5.79% to the community dissimilarity structure between the SSTC and ISSG. This was further supported by relatively high concentrations of the photosynthetic pigments chl c3 and peridinin (both indicative pigments for dinoflagellates) in the SO and SAF. We found a relatively high number of ASV94 and ASV23 (*Chloroparvula pacifica*) in the SSTC, contributing 3.07% to the community dissimilarity between the SSTC and the ISSG.

Chapter 1 | 3.6 Prokaryotic community composition

From each of the 14 stations, a fragment of the small subunit ribosomal RNA gene (16S rRNA) was amplified and sequenced to obtain insights into the diversity and community composition of prokaryotes. A total of 1308 ASVs were recovered from which we removed 267 ASVs annotated as chloroplasts and 68 ASVs annotated as mitochondria. Prokaryotic richness increased with increasing sea surface temperature (Pearson correlation: $r = 0.65$, $p = 0.03$, $n = 11$; Supplementary 1 Fig. S4a). Maximum alpha diversity (inverse Simpson) estimate was found in the SAF (81.92, station 15; Supplementary 1 Fig. S4d). Prokaryotic alpha diversity (inverse Simpson) was

positively (but not significantly) linked to primary productivity ($r = 0.36$, $p = 0.55$, $n = 11$; Supplementary 1 Fig. S4f) but showed a significant negative correlation with N_2 fixation ($r = -0.7$, $p = 0.05$, $n = 11$; Supplementary 1 Fig. S4h).

Our RDA of the prokaryotic ASV table captured 90% of the total variance with a p value of 0.06 (permutations = 999, $n = 11$). Sites clustered into Longhurst provinces along the first two RDA axes (62.48% of variance constrained; Fig. C1.4b). This was also shown in the PERMANOVA solution for Longhurst provinces (permutations = 999, $R^2 = 0.62$, $p < 0.001$, $n = 11$) and our PERMANOVA grouping into water masses (permutations = 999, $R^2 = 0.74$, $p < 0.001$, $n = 11$; Fig. C1.4b). We found more ASVs occurring in either the ISSG or the SO provinces rather than across all provinces (Fig. C1.4f). Further, the ISSG and the SO shared the least ASVs (Fig. C1.4f). In the RDA, sites within the ISSG province were positively associated with SST and N_2 fixation. Sites belonging to the SO were positively associated with dissolved inorganic nutrients (NO_3 , PO_4 , Si), dissolved oxygen, and chl a concentrations as well as high PP (Fig. C1.4b). The community composition within the SSTC (STF and SAF) was distinct from that of the ISSG and SO along the second RDA axis (21.67% variance constrained) and positively associated with NH_4 concentrations (Fig. C1.4b). Linear relationships between beta diversity and rates were weak for PP (PERMANOVA; permutations = 999, $R^2 = 0.31$, $p = 0.007$, $n = 11$) and N_2 fixation (PERMANOVA; permutations = 999, $R^2 = 0.2$, $p = 0.05$, $n = 11$).

Investigating whether and at which magnitude environmental parameters have an effect on prokaryotic microbial community dissimilarity, our general dissimilarity model (GDM) showed the expected curvilinear relationship (Fig. C1.4d I). Based on I-spline magnitude, geographic distance had little effect on community dissimilarity. The largest magnitude in community dissimilarity could be observed between 190–1200 $\mu\text{mol C L}^{-1} \text{ d}^{-1}$ (Fig. C1.4d III). Community dissimilarity changed most notably in response to variability in low magnitudes of N_2 fixation and did not change in samples with highest average N_2 fixation measurements (2.8 $\text{nmol N L}^{-1} \text{ d}^{-1}$ station 3, and 4.0 $\text{nmol N L}^{-1} \text{ d}^{-1}$ station 18). Largest magnitudes of community dissimilarity were associated with dissolved oxygen concentrations (Supplementary 1 Fig. S8b).

Taxonomically, based on analysis of the CLR-transformed ASV table, the prokaryotic community was dominated by Proteobacteria, Cyanobacteria, and Bacteroidetes, which are all typical clades for surface water samples (e.g., Biers et al., 2009). The greatest community differences occurred between stations of the Southern Ocean (SO) and the Indian South Subtropical Gyre (ISSG) provinces. Structure in community dissimilarity between the ISSG and SO were mostly associated with the number of Flavobacteriaceae (11.52% of total community dissimilarity, SIMPER analysis, Table S6) and Planktomarina (Alphaproteobacteria) (5.69% of the total difference in community dissimilarity, SIMPER analysis, Supplementary 1 Table S3). Further, the SO had distinct ASVs belonging to the SUP-05 cluster, contributing 2.56% (ASV_12) to the difference between SO and SSTC. The ISSG was characterized by a high number of Cyanobacteria and some Actinobacteria. The cyanobacterial fraction was dominated by *Prochlorococcus* and *Synechococcus*. Within the class level, all stations were dominated by Alpha- and Gammaproteobacteria, Bacteroidia, Oxyphotobacteria (Cyanobacteria), and Verrucomicrobia. Within the Alphaproteobacteria, we found a great dominance of ecotypes I, II, and IV of SAR11 clade throughout all samples. The relative number of pseudocounts of bacteria belonging to the phylum Bacteroidetes decreased towards warmer SST in the ISSG, with significant differences between the SO and ISSG (Welch two-sample t test $t = 4.58$, $p < 0.001$, $n_1 = 341$, $n_2 = 151$). The phylum Bacteroidetes was largely dominated by the order Flavobacteriales (90.98% of annotated ASVs). Cyanobacteria mainly occurred in the SSTC and in the ISSG, which were dominated by *Prochlorococcus* in the ISSG and *Synechococcus* in the SSTC. Cyanobacterial pseudocounts were significantly lower in the SO in comparison to the SSTC (Welch two sample t test, $t = -3.86$, $p < 0.001$, $n_1 = 17$, $n_2 = 31$) and to the ISSG (Welch two-sample t test, $t = 4.74$, $p < 0.001$, $n_1 = 17$, $n_2 = 45$). *Atelocyanobacteria* (UCYN-A) occurred in the SAF (station 14) and ISSG (stations 2, 3).

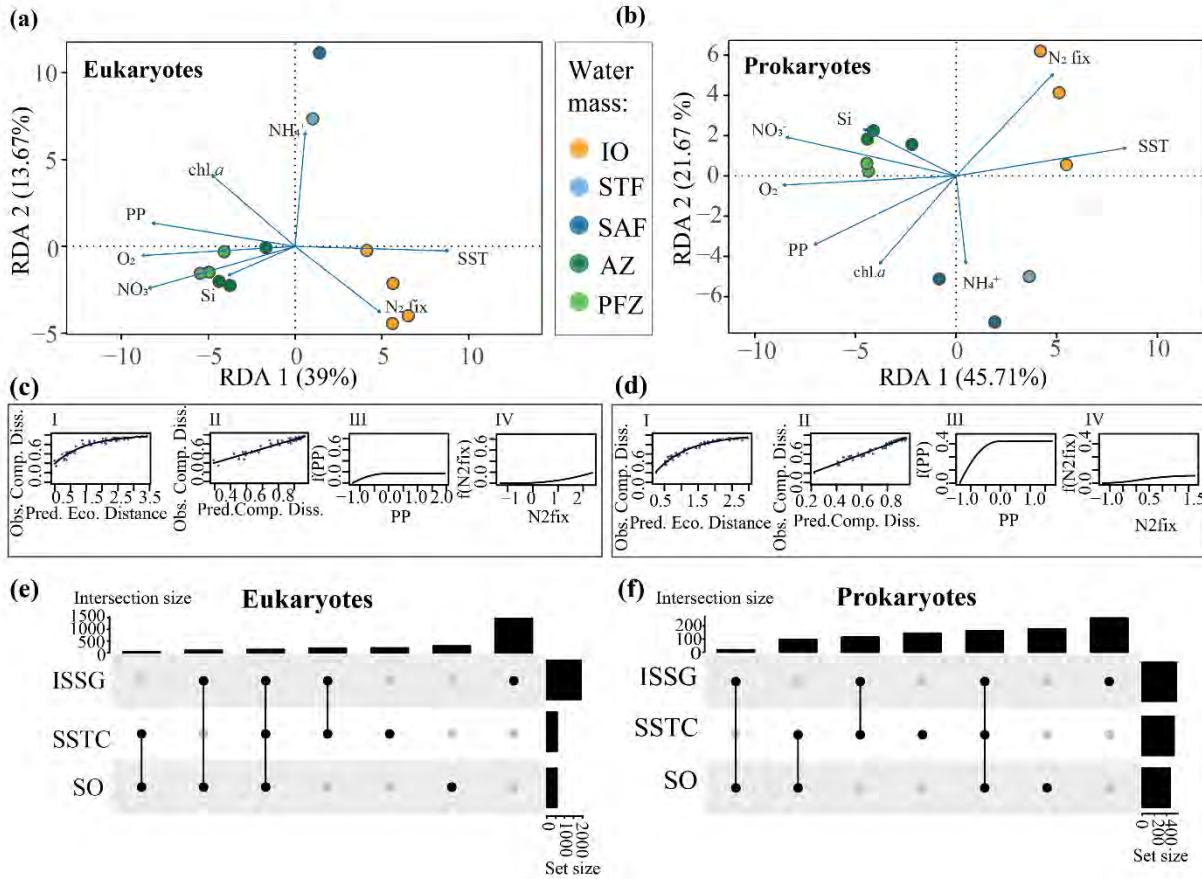


Figure C1.4. (a) Eukaryotic and (b) prokaryotic community structures of different water masses measured during the MD206 cruise. Redundancy analysis (RDA) of 18S and 16S rRNA gene ASV tables as response variables and environmental metadata as explanatory variables; environmental metadata are represented as arrows. Constrained analyses were performed by water mass. There were significant relationships between water masses and community dissimilarities (PERMANOVA, 999 permutations; $p < 0.001$, $R^2 = 0.67$ for eukaryotes and $p < 0.001$, $R^2 = 0.74$ for prokaryotes). Colors indicate major water masses according to the legend: yellow bar highlights the Indian South Subtropical Gyre (ISSG), light blue bar highlights the Subtropical Front (STF), blue bar highlights the Subantarctic Front (SAF), dark green bar highlights the Polar Front Zone (PFZ), and light green bar highlights the Antarctic Zone (AZ). Eukaryotic (c) and prokaryotic (d) general dissimilarity model (GDM) with (I) observed compositional dissimilarity against predicted ecological distance, calculated from temperature+dissolved oxygen + NO_3^- + NH_4^+ + Si + chl_a + PP + N_2 fixation; (II) observed compositional dissimilarity against predicted compositional dissimilarity to test the model fit; and contribution of (III) PP and (IV) N_2 fixation to community dissimilarity expressed as a function of the environmental parameter ($f(\text{PP})$ and $f(\text{N}_2\text{fix})$, respectively). For all functional plots of environmental data of the GDM analysis, see Supplementary 1 Fig. S8. Eukaryotic (e) and prokaryotic (f) UpSet plots of ASV intersections between Longhurst provinces. Analyses are based on binary tables (presence or absence) and the sum of all ASVs found across samples within one province. Intersection size shows number of ASVs shared between provinces (black dots, associated) and ASVs only found in one province (only black dot). Set size shows number of ASVs found in a specific Longhurst province.

Chapter 1 / 4 Discussion

Each water mass in our study had a distinct microbial fingerprint, including unique communities in frontal regions. We highlight clear relationships between microbial diversity, primary productivity, and N_2 fixation (high linear and nonlinear covariability) in the southern Indian Ocean Gyre (ISSG), the Southern Ocean (SO), and their frontal transition zone. Below, we discuss how this clear provincialism of microbial diversity is disconnected from regional gradients in primary productivity (PP) and N_2 fixation across our transect. This could suggest that microbial phylogenetic diversity is more strongly bounded by physical oceanographic boundaries, while microbial activity (and thus, perhaps, their functional diversity, not assessed here) responds more to chemical properties that changed more gradually between the low- and high-nutrient provinces we sampled.

 N_2 fixation and associated microbial diversity display distinct regional variations

Overall, our N_2 fixation (up to $4.4 \pm 2.5 \text{ nmol N L}^{-1} \text{ d}^{-1}$) was comparable to N_2 fixation measured by González et al. (2014) above the Kerguelen Plateau (up to $10.27 \pm 7.5 \text{ nmol N L}^{-1} \text{ d}^{-1}$) and showed a similar latitudinal trend as N_2 fixation further east in the Indian Ocean, although with around 10-fold lower absolute rates ($0.8\text{--}7$ vs. $34\text{--}113 \text{ nmol N L}^{-1} \text{ d}^{-1}$; Raes et al. 2014). We note that the localized rates reported by González et al. (2014) are to date the only published N_2 fixation measurements in this region, likely to be close to the annual maxima because of high irradiance; however, further investigations across seasonal changes within the study area are needed to confirm our observations. Our regional data are therefore important in closing the gaps in N_2 fixation measurements in the Southern Ocean, especially considering that large disagreements exist between models of high-latitude N_2 fixation rates (Tang et al. 2019). N_2 fixation measurements often show high basin-wide variability as well as high variability between samples at the same site, being sensitive to details of experimental design, incubation, and sea-state conditions (Mohr et al. 2010). In aggregate, these issues are best accounted for by calculating the minimum quantifiable rate (MQR; Gradoville et al., 2017). We observed high heterogeneity of biological samples taken from the underway flow-through system 5 min apart (separated by 15 km) within the same water

mass. Similar variability in absolute measurements of N_2 fixation ($2.6\text{--}10.3 \text{ nmol N L}^{-1} \text{ d}^{-1}$ to $7.5 \text{ nmol N L}^{-1} \text{ d}^{-1}$) were reported by González et al. (2014) close to our sampling site around Kerguelen island. This could imply a sub-mesoscale variability or influence of other unmeasured parameters. As oligotrophic gyres extend and displace southwards under climate change (Yang et al. 2020a), the biogeochemical and physical characteristics of the SO are changing (Caldeira and Wickett 2005; Swart et al. 2018), and biological regional N_2 fixation might become an important N source for productivity. Our data showed maximal N_2 fixation in the oligotrophic waters of the ISSG; however, notably, measurable N_2 fixation occurred well into the SO, to 56°S , suggesting that N_2 fixation contributes to the regional N pool, despite other available sources of N (Sipler et al. 2017; Shiozaki et al. 2018). Similarly, we found a negative N in the SO, which potentially indicates a P excess supporting N_2 fixation (Knapp 2012). Noteworthy is a slight increase in N_2 fixation in the Antarctic Zone (AZ). High-latitude measurements in northern polar regions (Bering Sea) reached $10\text{--}11 \text{ nmol N L}^{-1} \text{ d}^{-1}$ (Shiozaki et al. 2017), substantially higher than our measurements of the SO ($0.8\text{--}1.9 \text{ nmol N L}^{-1} \text{ d}^{-1}$), potentially supported by the close proximity to the coast or other factors such as day length, seasonality, diazotroph community, or trace metal concentrations. Our results suggest that regional N_2 fixation was not limited by the presence of other sources of bioavailable N (Supplementary 1 Fig. S9); this is a conclusion also reached in a number of studies including culture experiments (Knapp 2012; Eichner et al. 2014; Boatman et al. 2018), as well as in situ measurements in the South Pacific (Halm et al. 2012); off the coast of Chile and Peru with rates up to $190 \mu\text{mol N m}^{-2} \text{ d}^{-1}$ (Fernandez et al. 2011); and across the eastern Indian Ocean (Raes et al. 2015). This evidence counters the hypothesis of Breitbarth et al. (2007) that N_2 fixation occurs only when other sources of N are limited. The contribution of N_2 fixation to the N pool – and thus to productivity – varies strongly with ecosystem structure: in the SO, despite the local N_2 -fixation measurements, N_2 fixation remains likely a very minor contributor to the N required by the microbial community for primary productivity. Our results also strongly suggest that prokaryotic community structure and composition (beta diversity) were strongly impacted by the presence of biological N_2 fixation, which is itself a prokaryotic process (Karl et al. 2002). For example, the N_2 -fixing *Atelocyanobacteria* (UCYN-A) occurred in the SAF and ISSG; however, to gain a clear insight into the community and N_2 fixation, the diazotrophic community would need to be further resolved by amplicon analysis of functional (*nifH*) genes (Luo et al. 2012) as shown in other high-latitude studies (Fernández-Méndez et al. 2016; Raes et al. 2020).

Total and specific primary productivity differentially affect microbial diversity

We found PP was highest in the PFZ and decreased towards higher latitudes in the SO (Fig. C1.3a). Strass et al. (2002) showed that frontal maxima of PP are expected, and the observed decrease was probably due to Fe limitation in the SO (Blain et al. 2008). Primary productivity can also be limited by Si concentration and light availability when the mixed layer deepens (Boyd et al. 2000), but in our data Si concentrations were high in the surface water samples, and light levels were close to maximum in austral summer. The measured maximum PP above the Kerguelen Plateau (station E) was likely stimulated by Fe inputs (Blain et al. 2007). Our results did not support prior observations that frontal regions (SAF and STF) supported higher specific primary productivity (P^B ; as reported in the Antarctic Atlantic sector; Laubscher et al. 1993). While phytoplankton community composition, phytoplankton size distribution, and nutrient concentrations were strikingly different between the ISSG and SO, we found little difference in P^B , with some slightly lower values observed within the SSTC (Fig. C1.3b). Differences in P^B usually arise from physiological changes due to variabilities in irradiance (Geider 1987), nutrient concentrations (Chalup and Laws 1990; Behrenfeld et al. 2008), or differences in phytoplankton community structure, where cyanobacteria have the highest PP efficiency and diatoms the lowest (Talaber et al. 2018). Thus, our observations suggest that either (1) there is a lack of selective pressure on photosynthetic efficiency between provinces or (2) mechanisms driving P^B are different between provinces, and the sum of beneficial (e.g., increased nutrient concentrations in the SO) and detrimental mechanisms (e.g., low irradiance and photoinhibition through deep vertical mixing, reported from the Antarctic circumpolar current (ACC); Alderkamp et al. 2011) result in similar P^B . The slight variation around the frontal system is hard to interpret, as the complex interplay between factors may result in stochasticity. Primary productivity can be an important driver for (phylogenetic) microbial alpha diversity (Vallina et al. 2014), especially within ocean provinces (Raes et al. 2018). While our observational study only has a small number of samples within and between oceanic provinces ($n = 12$, $n_{ISSG} = 4$, $n_{SSTC} = 3$, $n_{SO} = 4$), it did suggest that further validation of this assumption is needed. We observed that PP changed gradually across the sampling region and that local variability in PP was high between samples taken 15 km apart within the SSTC and SO (Fig. C1.3a). These local variabilities can arise from complex physico-chemical interactions between the STF, SAF, and SO (Mongin et al. 2008). Counter to Vallina et al. (2014) and Raes et al. (2018), we found a significant negative correlation between eukaryotic alpha diversity and PP within the ISSG. Further, we found no correlation

between eukaryotic diversity and PP within the SSTC and SO and none between prokaryotic alpha diversity across all provinces (Supplementary 1 Fig. S7). In terms of beta diversity, we observed a structuring effect of PP for pigment, 16S rRNA gene, and 18S rRNA gene-derived diversity profiles (Fig. C1.4a, b, Supplementary 1 Fig. S8). Pigment analysis revealed that photosynthetic prokaryotic diversity is strongly impacted by the relative abundance of *Prochlorococcus*, which does not generally occur in cold high-latitude waters (> 40 S/N; Supplementary 1 Fig. S6) (Partensky et al. 1999) and, if so, only in low abundance (reviewed in Wilkins et al., 2013). Our 16S rRNA gene analyses confirm these observations showing that (1) picoplankton – and specifically *Prochlorococcus* – had relatively high proportions in the ISSG but very low in the SSTC, (2) *Synechococcus* dominated the Cyanobacterial fraction in the SSTC, and (3) both *Prochlorococcus* and *Synechococcus* were not detected in the SO (Supplementary 1 Fig. S6). In the SSTC and SO, phytoplankton communities had high proportions of dinoflagellates (Dinophyceae) and diatoms (Bacillariophyta) (up to 74% of diatom diagnostic pigment concentrations), which are known as essential contributors to marine PP and microbial diversity (Malviya et al. 2016) and known to dominate the phytoplankton fraction within the Polar Frontal Zone (PFZ), especially as the blooming season progresses (Brown and Landry 2001). Further, our results show that phytoplankton community structure appears to be tightly coupled to the occurrence of specific heterotrophic organisms (Supplementary 1 Table S3) and thus may mediate an indirect effect of PP through microbial food webs (as also noted in, e.g., Sarmento and Gasol 2012). For example, in areas of relatively high diatom concentrations, we found increased proportions of Flavobacteria. These bacteria specialize on successive decomposition of algal-derived organic matter (Teeling et al. 2012) and are known associates of diatoms (Pinhassi et al. 2004). Further, Planktomarina belonging to the Roseobacter clade affiliated (RCA) subgroup had relatively high proportions in the SO and is generally suggested to occur in colder environments (Giebel et al. 2009) and previously detected in the Polar Front (Wilkins et al. 2013). The RCA subgroup is known for dimethylsulfoniopropionate (DMSP) degradation in phytoplankton blooms (Han et al. 2020). In addition to bacteria known to be associated with phytoplankton, we also observed those which symbiose with other organisms (e.g., Georgieva et al. 2020), such as the sulfur oxidizing Thioglobaceae (SUP-05 cluster), previously found in symbiosis with Myctophidae fish near Kerguelen Islands (Gallet et al. 2019). While beyond the scope of this study, we encourage

further investigations of such trans-kingdom functional interactions as they themselves may offer regional insights.

Implications for microbial regionality

Microbial diversity was regionally constrained independent of geographical distance (GDM analysis), but it was partitioned into ocean provinces as repeatedly described for other ocean basins such as the Pacific (Raes et al., 2018) and the Atlantic Ocean (Milici et al. 2016). This supports the classical concept of microbial biogeography (Martiny et al. 2006). Further, we found that microbial beta diversity was even better resolved by individual water masses, highlighting the importance of including oceanographic boundaries that limit cross-front dispersal (Hanson et al. 2012; Wilkins et al. 2013; Hernando-Morales et al. 2017). Our beta diversity analysis confirmed the findings by Baltar and Arístegui (2017), who found unique environmental sorting and/or selection of microbial populations in the SAF and STF. Further, we were able to link these communities to high NH_4 concentrations. This suggests high recycling of nitrogen sources within the microbial loop and potentially favoring nitrification in this area (Sambrotto and Mace 2000). We also found increased dinoflagellate concentrations (PFT) which have been described to grow well under NH_4 conditions (Townsend and Pettigrew 1997). Despite our small sample size within the SAF and STF, we were able to detect these characteristics, supporting the call from Baltar et al. (2016) for better integrating frontal zones in our understanding of microbial biogeography. Different trade-offs such as nutrient limitation and grazing can shape the microbial seascape (Acevedo-Trejos et al. 2018). In our study, the deviation between $\text{PN} : \text{chl } a$ was large between the SO and IO with high $\text{PN} : \text{chl } a$ ratios in the ISSG (Supplementary 1 Fig. S7), which has been used as an indicator of a relatively high abundance of heterotrophic microbes and protists over autotrophic organisms (Hager et al. 1984; Waite et al. 2007; Crawford et al. 2015). This would suggest that grazers formed a higher fraction of total biomass in the ISSG than in the SO. However, we did not measure zooplankton biomass or grazing rates, so this remains speculative.

Chapter 1 / 5 Conclusion and outlook

Our study leads us to conclude that simultaneous assessment of microbial diversity, biogeochemical rates, and the physical partitioning of the ocean (provincialism) is central to the understanding of microbial oceanography. Each water mass in our study had a distinct microbial fingerprint, including unique communities in frontal regions. Microbial alpha diversity and community dissimilarity correlated with biogeochemical rate measurements; however, mechanisms driving this association need further investigation through high-resolution sampling across spatial and temporal scales. Our results also indicate that high-latitude N₂ fixation could meaningfully contribute to the global and regional N pool (as reported for Arctic N₂ fixation by Sipler et al. 2017), which may become especially significant as global stratification (and concomitant restrictions in deep water replenishment of nutrients) intensifies. While our sampling is too limited to conclude the point, our observations that phylogenetic diversity is constrained by hydrographic properties and province boundaries but that biogeochemical rates and nutrient concentrations are changing more gradually suggest that trans-province functional redundancy is present despite strong biogeographic separation in phylogenetic terms. As an outlook, we therefore encourage examining both phylogenetic and functional diversity to assess how functional groups and guilds contribute to the major biogeochemical (C, N) cycles across provinces and other biogeographic regions. Coordinated studies across ocean provinces are key to establishing the baselines we need to monitor the rapidly changing properties of the southern high latitudes in the face of rising temperature, acidification, and perturbations in regional currents.

Code availability.

All code is publicly archived under <https://www.doi.org/10.5281/zenodo.5779517> (Hörstmann, 2021) in the github repository CoraHoerstmann/MD206_Microbes.

Data availability.

All HPLC data; environmental and rate measurement data, including PN, MIMS data, PP, and N₂ fixation; and minimum quantification rate calculations are stored in the PANGAEA database (Hörstmann et al., 2018). All sequences are archived in the European Nucleotide Archive (primary accession: PRJEB29488).

CHAPTER 2: Microbial diversity through an oceanographic lens: refining the concept of ocean provinces through trophic-level analysis and productivity-specific length scales

Authors: Cora Hörstmann, Pier Luigi Buttigieg, Uwe John, Eric Raes, Dieter Wolf-Gladrow, Astrid Bracher, Anya M. Waite

Abstract

In the marine realm, microorganisms are responsible for the bulk of primary production, thereby sustaining marine life across all trophic levels. Longhurst provinces have distinct microbial fingerprints; however, little is known about how microbial diversity and primary productivity change at finer spatial scales. Here, we sampled the Atlantic Ocean from south to north ($\sim 50^{\circ}\text{S}$ – 50°N), every $\sim 0.5^{\circ}$ latitude. We conducted measurements of primary productivity, chlorophyll *a*, and relative abundance of 16S and 18S rRNA genes, alongside analyses of the physicochemical and hydrographic environment. We analyzed the diversity of autotrophs, mixotrophs and heterotrophs, and noted distinct patterns among these guilds across provinces with high- and low-chlorophyll *a* conditions. Eukaryotic autotrophs and prokaryotic heterotrophs showed shared inter-province diversity pattern, distinct from the diversity pattern shared by mixotrophs, cyanobacteria and eukaryotic heterotrophs. Additionally, we calculated samplewise productivity-specific length scales, the potential horizontal displacement of microbial communities by surface currents to an intrinsic biological rate (here, specific primary productivity). This scale provides key context for our trophically disaggregated diversity analysis that we could relate to underlying oceanographic features. We integrate this element to provide more nuanced insights into the mosaic-like nature of microbial provincialism, linking diversity patterns to oceanographic transport through primary production.

Chapter 2 / 1 Introduction

The continuous movement of seawater and turnover of microbial biomass and diversity in marine ecosystems form a time-varying mosaic of phylogenetic and functional biodiversity across ocean basins. This complicates efforts to map, understand and monitor key marine ecosystem attributes such as microbial growth, biodiversity, and carbon cycling (Stec et al. 2017).

Marine ecosystems comprise a wide array of microbial life: microbial (photo-)autotrophic, pro- and eukaryotes, form the base of the marine food web (Hutchins et al. 2015; Sunagawa et al. 2015), and sustain energy exchange, provision, and recycling of resources (Falkowski et al. 2008; Guidi et al. 2016) for higher trophic levels. Heterotrophs- and mixotrophs remineralize most of the carbon and nutrients from the primary production via the microbial loop, before these can be exported to the deep sea (Azam et al. 1983; Azam and Malfatti 2007).

In terrestrial systems, primary producer communities have been used to define major biomes (Woodward 1987; Woodward et al. 2004), biogeographic realms, and ecoregions (Olson et al. 2001) through their physical and functional structuring of ecosystems (Cardinale et al. 2011). In the ocean, Longhurst (Longhurst 2007) used chlorophyll *a* (chl *a*) concentrations as a proxy for phytoplankton biomass to delimit ocean provinces, alongside water temperature to distinguish water masses. Longhurst provinces are used to define oceanographic biogeographic subdivisions; however, static applications of these provinces typically overlook the dynamic interactions, life histories, endemism and/or vicariance within ecological assemblages, needed to truly map microbial biogeography. Thus, multiple variables including biomass, primary productivity (PP), and diversity need to be considered with carefully structured sampling across space and time (Kollmann et al. 2016). Here, we investigate how integrating multiple physical and biochemical variables – accounting for their horizontal displacement by surface currents – can improve our understanding of microbial provincialization across a high-resolution transect ($\sim 0.5^\circ$ latitude) of the Atlantic Ocean.

Microbial assemblages can act as “fingerprints” for water masses thanks to their high diversity, responsiveness, and (typically) fast generation time (hours to days; Martiny et al., 2006), aiding the definition of biogeographic boundaries in pelagic environments (Raes et al. 2011; Fuhrman et al. 2015). However, the dynamic nature of the oceans – with their fronts, currents, eddies, up- and downwellings, and other hydrographic features – has the potential to add additional complexity to the ecological variation between and within such regions (Oliver and Irwin 2008; Hernando-Morales et al. 2017).

Hydrographic features create structural variability in the ocean, which (through, e.g. modifying nutrient distributions) provides more diverse niche space for microbial communities to colonize. Unless mitigated, this would increase microbial diversity (e.g. Kemp and Mitsch, 1979; Cadotte, 2006) and should favor phytoplankton productivity (Legendre 1981). While basin-scale horizontal dispersal of organisms by major ocean currents is known to reduce microbial β -diversity (Raes et al. 2011; Richter et al. 2019; Sommeria-Klein et al. 2021), mesoscale and sub-mesoscale horizontal transport of microbial communities, and the impact on their activity, is rarely taken into account, except in a few frameworks that couple ecology and hydrography (e.g. the “dual-lens” approach; Oldham et al., 2013). Indeed, hydrographic dynamics within and between ocean provinces interact with both neutral and selective ecological processes, resulting in communities in different successional states (Zhou et al. 2014) and/or shaped by opportunistic responses (e.g. Duffy and Stachowicz, 2006; Hartmann et al., 2012; Fadeev et al., 2021).

Here, we assess the relative importance of regional water mass characteristics (physicochemical parameters and hydrography) on microbial diversity. Further, we resolve diversity responses along major trophic groups (auto-, mixo-, heterotrophs), an important but under-studied perspective in microbial ecology (reviewed in Seibold et al., 2018). We leverage a conceptual framework which asserts that microbial communities are distinct within oceanographic regions, separated by fronts and currents limiting microbial dispersal (Martiny et al. 2006; Milici et al. 2016; Raes et al. 2018). We then assess how ecosystem structure (i.e. hydrography), bottom-up, and top-down factors can qualify and advance traditional partitioning of the Atlantic into biogeographic provinces.

Chapter 2 / 2 Results

Delineation of oceanographic, ecosystemic provinces

Our analyses of remote sensing observations of geostrophic currents, sea surface temperature and chl *a* combined with in situ measurements of oceanographic features, chl *a*, and microbial diversity informed our determination of ecosystem boundaries shown in Figure C2.1. The ecosystem boundaries broadly overlapped with the Longhurst Provinces and we thus maintain the same naming conventions (Supplementary 2 Fig. S1). One exception was the South Atlantic Gyral (SATL) province, wherein sites were clearly ordinated into separate groups; one with cool waters (COLD; 20.4–23.3 °C) in the Argentine Basin, south of the Rio Grande rise, and northern, oligotrophic warmer waters (HOT; 25.1–27.5 °C) in the Brazil Basin (Supplementary 2 Fig. S2, S3). Further, Stations 24 and 92 were outliers with respect to the temperature-salinity plot used for identifying ecosystem boundaries (Supplementary 2 Fig. S2), and were thus not grouped in an ocean province.

On a broad scale, we could separate ocean provinces into provinces with low and high chl *a* with significant differences in chl *a* concentrations (Wilcoxon, $p = 9.112e^{-6}$, $n_1 = 38$, $n_2 = 39$), and significant differences in PP (Wilcoxon, $p = 3.686e^{-5}$, $n_1 = 38$, $n_2 = 39$; Table C2.1). In our principal component analysis (PCA) (Supplementary 2 Fig. S4), sites in the Southwest Atlantic Shelves province (FKLD), Brazil Current Coast province (BRAZ), Canary Current Coast province (CNRY), North Atlantic Subtropical Gyre province (NAST-E), and North Atlantic Drift province (NADR) (i.e. high-chl *a*) provinces were associated with high particulate organic matter and high dissolved inorganic nutrient concentrations. Further, the SATL-COLD, SATL-HOT, Western Tropical Atlantic province (WTRA), and North Atlantic Tropical Gyre province (NATR) (i.e. low-chl *a* provinces) provinces were associated with low temperatures and larger distance to coast (> 600 km; Supplementary 2 Fig. S4).

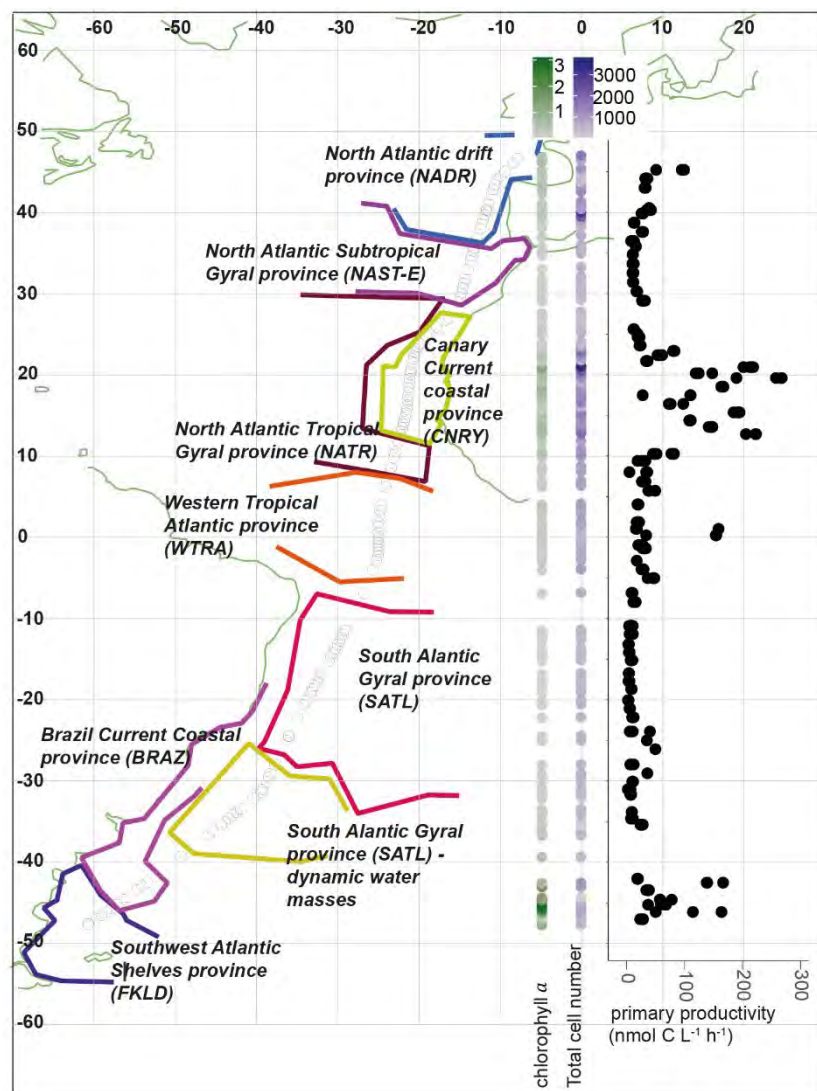


Figure C2.1. Map of the PS113 expedition showing chl *a* concentration gradients ranging from 0–3 mg chl *a* m⁻³, total nanoplanktonic cell number (ranging from 1000 - 3000 cells μl⁻¹) and primary productivity (PP) (nmol C l⁻¹ h⁻¹) against latitude. Stations are indicated by circles. Ocean provinces are indicated with colored lines.

Province PP and chl *a* concentrations correlate across the transect (Pearson correlation; $r = 0.48$, $p = 6.68 \times 10^{-6}$, $n = 80$), but did not always correlate within provinces (e.g. WTRA: $r = 0.08$, $p = 0.8$, $n = 11$). We measured the highest PP in the CNRY province, a peak in chl *a* concentrations relative to adjacent ocean provinces (Table C2.1), and total cell numbers of 1207 ± 737 cells μl⁻¹ (mean \pm SD, $n = 30$).

Table C2.1. Physico-chemical properties, primary productivity, chl *a* concentrations and productivity-specific length scale of ocean provinces during the PS113 expedition. Provinces are separated in H-CHL provinces and L-CHL provinces.

province	H-CHL / L-CHL	temperature	salinity	POC (nM)	PN (nM)	n	nitrate (μ M)	phosphat e (μ M)	n	PP (nmol C L ⁻¹ h ⁻¹)		Chl <i>a</i> (mg m ⁻³)		n	PL	n
										range	mean \pm sd	range	mean \pm sd			
BRAZ	H-CHL	11.26–18.62	33.98–35.28	4071–16953	608–2913	8	0.43–2.43	0.36–0.55	3	36.7–108.1	65.77	0.81–3.1	1.78	4	3.5–22.6	4
Station 24	H-CHL	19.89	36.14	3070	488.5	1	8.9	0.71	1	146	146	0.72		1	2.2	1
CNRY	H-CHL	19.96–22.58	35.87–37.05	3201–24922	250.1–4371.3	30	0–3.8	0.01–0.31	17	12.8–234.7	105.82 \pm 7	0.13–1.22	0.49 \pm 0.3	17	0.4–3.2	17
FKLD	H-CHL	8.19–9.42	33.76–33.99	4459–8246	497–1474	4	13.095	1.08	1	26.41	26.41	0.65	0.65	1	4.4	1
NADR	H-CHL	15.32–17.07	35.55–35.92	4904–12717	466.5–1875.9	17	0.01	0.04–0.08	8	13.3–81	36.6 \pm 20	0.14–0.44	0.32 \pm 0.1	8	0.6–4.5	8
NAST-E	H-CHL	17.55–25.14	36.14–36.93	2948–5731	227.2–779.8	12	<0.01	0.01–0.09	5	11.1–29.1	15.00 \pm 6	0.07–0.18	0.13 \pm 0.0	8	0.9–6.9	8
NATR	L-CHL	23.98–26.12	35.95–36.13	5356–9138	670.1–1237.8	6	0.01–0.3	0.02–0.19	4	24.9–79.3	44.76	0.14–0.43	0.3	4	0.8–1.8	4
SATL-COLD	L-CHL	20.39–23.28	35.83–36.30	2188–4357	257.9–561.6	16	<0.11	<0.1	9	4.7–35.7	14.63 \pm 10	0.08–0.32	0.18 \pm 0.1	9	1.2–12.2	9
SATL-HOT	L-CHL	25.14–27.54	36.93–37.41	1525–3042	58.2–377.3	20	<0.04	<0.1	15	3.6–49.3	12.3 \pm 13	0.04–0.13	0.07 \pm 0.0	15	0.4–4.7	15
WTRA	L-CHL	27.32–28.65	34.16–36.27	2609–4455	265.4–594.9	17	<0.2	<0.09	11	17.8–112.7	39.25 \pm 28	0.13–0.35	0.21 \pm 0.1	11	0.7–9.2	11
Station 90+92	L-CHL	28.33	36.33	2703	296.6	1	0.01	0.02–0.07	2	9.7–13.6	12.67	0.09–0.1	0.09	2	1.4–2.1	2

Biomass turnover and transport of pelagic microbial communities

In the following analyses, we used a quantitative comparison of biomass turnover (specific PP; P^B) and current speed (see Methods, Equations C2.3, C2.4) to estimate the distance a microbial community has traveled through passive advection before half of its biomass has been turned over: the productivity-specific length scale (Fig. C2.2a).

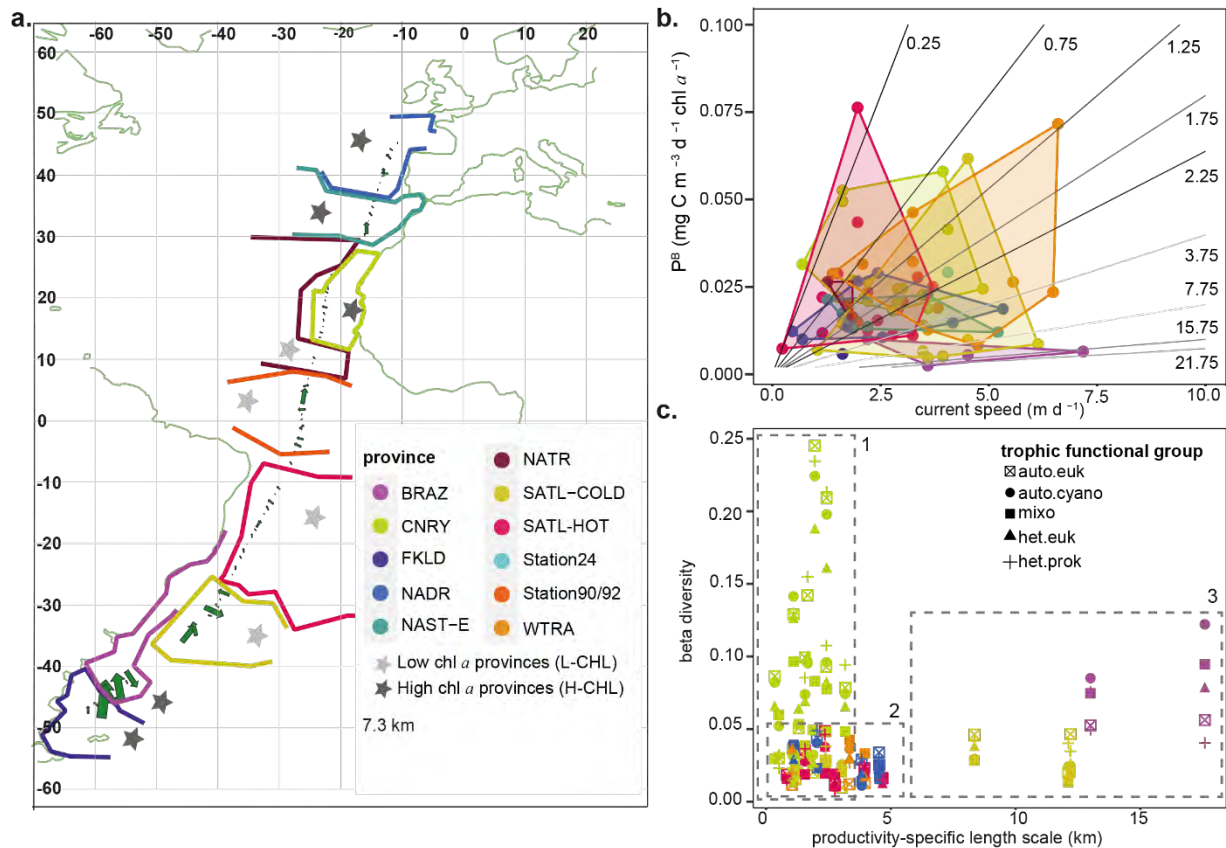


Figure C2.2. (a) map of the PS113 expedition indicating productivity specific length scale using a quantitative comparison of measured biomass turnover (half of standing stock phytoplankton biomass (chl a concentration) is replaced by new biomass (calculated from primary productivity rates) and measured horizontal current speed (ADCP horizontal velocity), see equation (3, 4) for more details. Direction is derived from ADCP measurements. Ocean provinces are indicated with colored lines. Light grey stars indicate low-chl a provinces, dark grey stars indicate high-chl a provinces (b) Specific primary productivity (P^B) against current speed (m d^{-1}) at each site. Sites are color coded according to ocean provinces. Productivity-specific length scales (km) are indicated by grey linears across the plot. (c) beta diversity to neighbouring sites based on ordination distances against productivity-specific length scales for each trophic functional group. Sites are color coded according to ocean provinces. Box 1 indicates high beta diversity with great variability within the CNRY province. Box 2 corresponds to beta diversity variation of most ocean provinces with an upper limit of 5 km productivity-specific length scale; box 3 indicates an increase in beta diversity with productivity-specific length scales > 5 km.

Based on the biological and physical component of the productivity-specific length scale, we observed strong sample variations in both variables (P^B based on PP per day (d) and chl *a*, Table C2.1, and surface current speed) within and between provinces (Fig. C2.2b). For example, the SATL-HOT province had high variations in P^B ($0.007\text{--}0.07\text{ mg C m}^{-3}\text{ d}^{-1}$, $n = 15$) with relatively constant horizontal current speed ($1\text{--}3\text{ m d}^{-1}$). In contrast, we measured a large range of horizontal current speed ($2\text{--}7\text{ m d}^{-1}$, $n = 4$) at relatively low P^B ($0.002\text{--}0.01\text{ mg C m}^{-3}\text{ d}^{-1}$, $n = 4$) within the BRAZ province resulting in the largest productivity-specific length scales (up to 22 km) of our dataset.

We used our calculations of sample-specific productivity-specific length scales to examine variabilities in beta diversity (β -diversity) between neighbouring sites (Fig. C2.2c). We noted three major regimes in our joint analysis of β -diversity and productivity-specific length scales (Fig. C2.2c): The first regime matches the CNRY province and shows markedly high variability in β -diversity, but always at low ($< 5\text{ km}$) productivity-specific length scales. The second regime shows a largely random scatter bounded by a well-defined upper threshold of β -diversity (0.05 ordination distance). Similar to the first regime, we observed this distribution of β -diversity to be restricted to productivity-specific length scales under 5 km. The final regime occurs at productivity-specific length scales between 5 km and 18 km. We observed an increase in β -diversity with increasing productivity-specific length scales. However, the variation between functional trophic groups increased with increasing productivity-specific length scales, too. Cyanobacteria, eukaryotic heterotrophs and mixotrophs were more similar to each other with a stronger increase in β -diversity compared to the β -diversity observed in prokaryotic heterotrophs and eukaryotic autotrophs.

Alpha diversity (α -diversity) across functional groups and oceanographic provinces

Rarefaction curves for both prokaryotes and eukaryotes saturated in all samples (Supplementary 2 Fig. S5). Prokaryotic Shannon diversity was highest (260.1) at Station 11 (BRAZ province) and lowest (59) at Station 97 (WTRA province; Supplementary 2 Table S2). Eukaryotic Shannon diversity was highest (766.8) at Station 45 (SATL-COLD province) and lowest (11.2) at Station 8 (BRAZ province).

We observed that α -diversity varied across trophic groups, provinces, and latitude (Fig. C2.3a). The α -diversity of both eukaryotic and prokaryotic autotrophs and eukaryotic mixotrophs was lower than that of heterotrophs. Relative to other functional groups, heterotrophic prokaryotes had higher α -diversity in the southernmost and northernmost provinces (FKLD, BRAZ, NADR). Shannon diversity increased towards the equator and was most pronounced for heterotrophic pro- and eukaryotes (Fig. C2.3a) with the highest α -diversity in the SATL-HOT, SATL-COLD, and WTRA provinces; the warmest and most oligotrophic provinces sampled. Similarly, the low-chl *a* provinces had significantly higher mixotrophic α -diversity than the high-chl *a* provinces (Wilcoxon, $p < 2.2e^{-16}$, $n_1 = 39$, $n_2 = 68$).

In relation to the physical-chemical parameters, α -diversity was positively and linearly correlated with physical variables (temperature and salinity), and negatively correlated with nutrient and biochemical concentrations (NO_3 , PO_4 , POC, PN, chl *a*) across ocean provinces (Fig. C2.3b). The correlation between temperature and α -diversity was significant for autotrophic eukaryotes, mixotrophs, and heterotrophic prokaryotes (Fig. C2.3b, Supplementary 2 Table S1). POC, PN, and chl *a* were significantly negatively correlated with the α -diversity of autotrophic eukaryotes. The α -diversity of heterotrophic prokaryotes was negatively correlated with NO_3 , PO_4 , POC, PN (Fig. C2.3b, Supplementary 2 Table S1). However, *within* provinces, we found both positive and negative correlations of functional groups with temperature, nutrients, and particulate matter (e.g. PN correlated positively with cyanobacterial diversity in the SATL-COLD province: $r = 0.8$, $p = 3.85e^{-4}$, $n = 15$, Fig. C2.3b).

Primary productivity had no significant correlations with picoplankton cell number or microbial α -diversity across any trophic functional group (Supplementary 2 Table S1). However, in the CNRY province, where productivity rates were significantly higher than in other provinces (Wilcoxon, $p = 5.75e^{-4}$, $n_1 = 17$, $n_2 = 77$), diversity of autotrophic eukaryotes and cyanobacteria were negatively correlated with PP ($r = -0.6$, $p = 0.04$, $n = 14$, and $r = -0.6$, $p = 0.01$, $n = 14$, respectively; Fig. C2.3b).

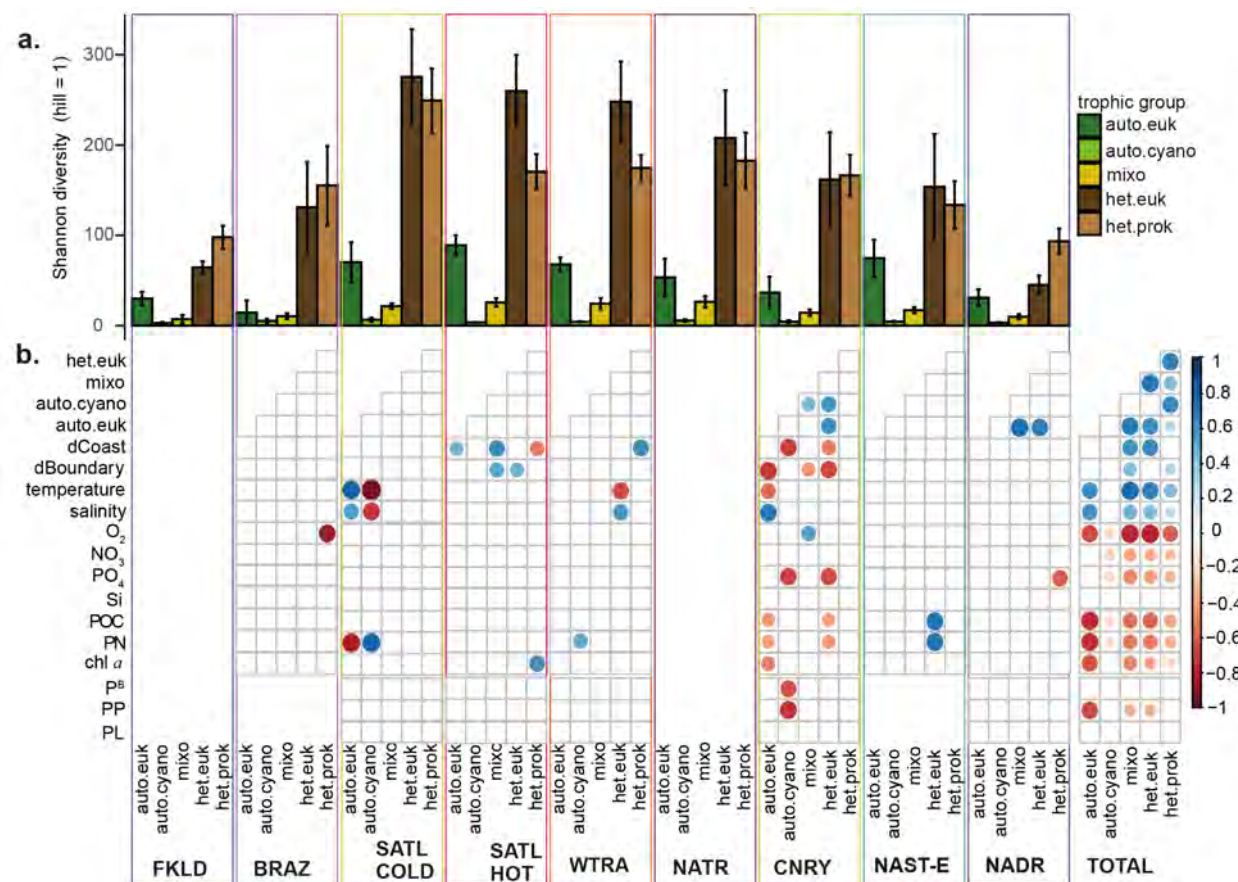


Figure C2.3. Alpha diversity across ocean provinces. (a) Mean microbial Shannon diversity of eukaryotic autotrophs, prokaryotic autotrophs (cyanobacteria), eukaryotic mixotrophs, heterotrophic eukaryotes and heterotrophic prokaryotes; Error bars indicate standard deviation of microbial Shannon diversity within each province. Sample size within each province are indicated below; (b) Pearson correlations of Shannon diversities with and environmental parameters within ocean provinces. Distance to coast (dCoast) calculated as distance of sample to next shore (in km), dBoundary is the distance to the province boundary of the sample identified to belong to, temperature is sea surface temperature (°C), salinity is sea surface salinity, oxygen (μM), nitrate (NO_3) in μM , phosphate (PO_4) in μM Silicate (Si) in μM , Particulate organic Carbon concentration (POC) in nM, particulate nitrogen (PN) concentration in nM, chl *a* in mg m^{-3} , specific primary productivity (P^{B}) in $\text{nmol C l}^{-1} \text{ chl } a^{-1}$, primary productivity (PP) in $\text{nmol l}^{-1} \text{ h}^{-1}$. Correlation plots with colors indicating gradient from negative (red) to positive (blue) correlation; correlations of not significant, i.e. $p > 0.05$, and with non-normal distribution of residuals in linear regression model were removed from the plot (for full correlation plot see Supplementary 2 Fig. S6, and Fig. S8–S15 for residual histogram plots of individual provinces and across the entire transect. Within provinces, correlations were calculated for eukaryotes and prokaryotes, respectively, and each trophic group (autotroph, mixotroph and heterotroph) against different environmental variables and against each other. Colored boxes indicate correlations within provinces: FKLD, Southwest Atlantic Shelves province ($n = 4$); BRAZ, Brazil Current Coastal province ($n = 8$); SATL, South Atlantic Subtropical Gyral province (COLD: $n = 15$; HOT: $n = 20$); WTRA, Western Tropical Atlantic province ($n = 17$); NATR, North Atlantic Tropical Gyral province ($n = 6$); CNRY, Canary Current Coastal province ($n = 26$), NAST, North Atlantic Subtropical Gyral province ($n = 10$); NADR, North Atlantic Drift province ($n = 13$). TOTAL indicates all samples across the entire transect ($n = 121$).

We noted that α -diversity of trophic functional groups are similarly structured in their correlations with environmental (physical + chemical) parameters; however, the magnitudes and significances of the correlations varied between groups (e.g. NAST-E PO_4 , auto.cyano; $r = 0.83$, $p = 0.01$; PO_4 , het.prok, $r = 0.39$ $p = 0.27$, Supplementary 2 Fig. S6, S7), as did the distribution of residuals (Supplementary 2 Fig. S8–S15). Between functional groups, we observed a significant positive correlation between the α -diversity of autotrophic eukaryotes and cyanobacteria in the CNRY provinces or between autotrophic eukaryotes and mixotrophs in the CNRY and NADR provinces as well as across the transect (Fig. C2.3b, Supplementary 2 Fig. S16).

We observed that each province has a distinct correlation pattern across the variables we examined (Fig. C2.3b). Correlations were more pronounced in the CNRY province (Fig. C2.3b). We could not identify pairs of parameters with correlations that were consistently repeated across provinces (e.g. temperature, auto.euk SATL-COLD; temperature, auto.euk SATL-HOT; Fig. C2.3b). Further, we noted that samples were not evenly spread along bivariate plots, especially in physically energetic regions (i.e. large scatter of temperature–salinity profile; Supplementary 2 Fig. S17).

Beta diversity (β -diversity) patterns of auto-, mixo-, and heterotrophs across provinces

Sites belonging to high-chl *a* and low-chl *a* provinces were well separated along the first axes of our redundancy analysis (RDA) plots for each trophic functional group (Fig. C2.4; between 40 and 61% of variance constrained). Sites belonging to high-chl *a* provinces were associated with higher dissolved inorganic nutrient concentrations, particulate organic matter, and chl *a* concentrations. Sites belonging to low-chl *a* provinces were associated with higher temperatures, salinity and larger distances to province boundaries and the coast.

We found that sites in low-chl *a* and high-chl *a* provinces were differentially ordinated in our RDA biplots. We observed the shortest distances between sites in low-chl *a* provinces when analyzing autotrophic eukaryotic diversity. In contrast, heterotrophic eukaryotic and mixotrophic diversity led to sites in low-chl *a* provinces being ordinated furthest apart from those in other provinces (Supplementary 2 Table S2). While autotrophic eukaryotic communities were distinctly separated between the high-chl *a* provinces, mixotrophic communities had short relative ordination distances or overlapped in these provinces (Fig. C2.4a, c; Supplementary 2 Table S2). Across all trophic

groups, the samples from the CNRY province showed the largest spread among the second RDA axes and partially overlapped with samples from the NAST-E province (Fig. C2.4).

Prokaryotic heterotrophic communities were more similar to autotrophic eukaryotes in their pattern of the biplots (procrustes analysis, Supplementary 2 Table S2), with shorter relative distances in low-chl *a* than in high-chl *a* provinces (Fig. C2.4a, e; Supplementary 2 Table S1). In contrast, cyanobacteria, heterotrophic and mixotrophic eukaryotes were more similar to each other (procrustes analysis, Supplementary 2 Table S2); shorter relative distances in high-chl *a* than in low-chl *a* provinces (Fig. C2.4; Supplementary 2 Table S1).

Analysis of the β -diversity among communities of heterotrophic pro- and eukaryotes, and mixotrophs ordinated station 24 between those that are characteristic of the BRAZ province and the southern, cold part of the SATL-COLD province (Fig. C2.4c, d, e). Similarly, analysis of β -diversity of the autotrophic eukaryotes and mixotrophs at Station 92 (located between the SATL-HOT and the WTRA province), ordinated the station between the SATL-HOT and WTRA (Fig. C2.4a, c).

The SATL province, as defined by Longhurst (2007), clustered into two distinct groups in microbial β -diversity (Fig. C2.4) which was also apparent in our biogeochemical inspection and division into a southern dynamic (COLD) part, and a northern, oligotrophic (HOT) part in the Brazil Basin (Fig. C2.4).

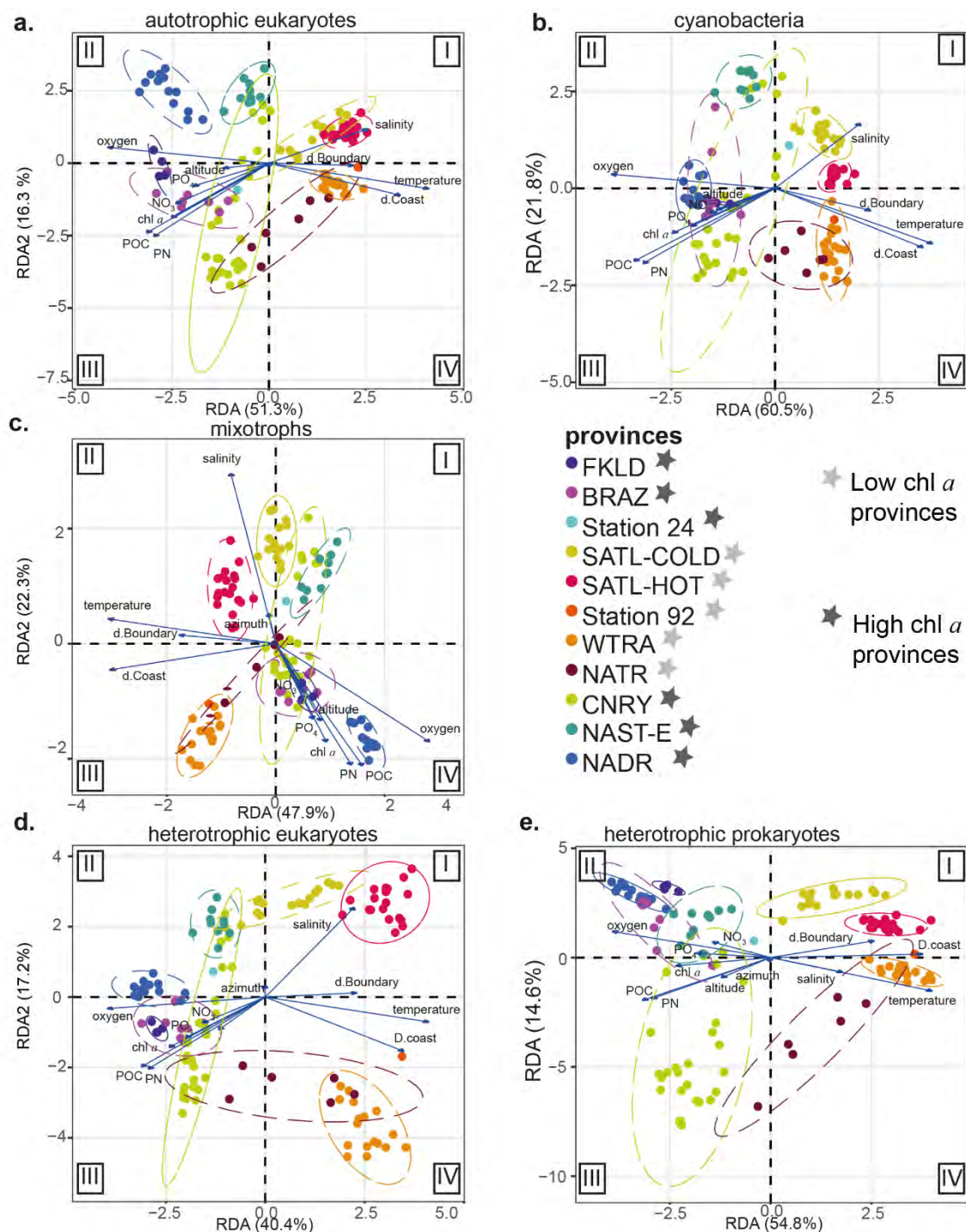


Figure C2.3. Redundancy analysis (RDA) of CLR-transformed ASV counts, subsequently partitioned into separate ASV tables for each functional group (a) autotrophic eukaryotes, (b) autotrophic prokaryotes (c) eukaryotic mixotrophs, (d) heterotrophic eukaryotes, and (e) heterotrophic prokaryotes. Contextual spatial and environmental data (z-scored) were used as explanatory variables, represented as arrows. Samples cluster according to ocean provinces. Samples separate between low-chl *a* provinces (SATL-COLD, SATL-HOT, WTRA, NATR) (indicated with light grey star) and high-chl *a* provinces (FKLD, BRAZ, CNRY, NAST-E, NADR) (indicated with dark grey star) along the first axis (RDA 1) across all functional groups.

Chapter 2 / 3 Discussion

Our study revealed that microbial trophic groups showed variable α - and β -diversity patterns in relation to physical and biogeochemical environmental parameters across our south-to-north transect. Below, we discuss how our integrative results can contribute to refining microbial, biogeographic provincialism. Our results provide an estimate of sample-based spatial scales of differences in biodiversity signals, suggesting that observations of ecosystem function and stability need regional high-resolution sampling.

In situ measurements of province characteristics deviate from Longhurst provinces

Broadly, traditional Longhurst province boundaries (Longhurst, 2007) are determined by clustering remote sensing data on chl *a* and temperature data (e.g., Devred et al., 2007; Hardman-Mountford et al., 2008) or in situ measurements of phytoplankton pigment composition (Barlow et al. 2007; Taylor et al. 2011; Bracher et al. 2020); Our delineation largely overlapped (for about 80% of the stations) with the provinces delineated by Bracher et al., (2020) using hierarchical cluster analysis on phytoplankton group composition data derived from the HPLC marker pigments (see Supplementary Material Table S6; Hörstmann et al. 2021). However, we observed regional divergence. Most pronounced was the separation of the SATL province into SATL-HOT and SATL-COLD; supported by our physical, chemical, and microbial observations. This difference could not be detected in phytoplankton pigment-based analyses (see Bracher et al. 2020). Such observations (if confirmed) exemplify how new provincialisation can arise from microbial oceanographic perspectives.

However, we also noted that some of our in situ measurements could lead to spurious distinctions within provinces. For example, we occasionally measured high nutrient and chl *a* concentrations in provinces known to be oligotrophic (Longhurst, 2007; e.g. up to 0.43 mg m^{-3} chl *a* in the NATR province). Given our sampling regime, this may simply be the result of sampling a transient patch of high-nutrient/high-chl *a* water. Our β -diversity supports this possibility, as NATR sites were largely ordinated with sites from other oligotrophic provinces along the first major axis in each of

our RDA triplots (Fig. C2.4). We note that this shows the value of microbial diversity data in holistically describing oceanographic provinces, especially when sampling dynamic regions.

Ocean currents can drive overlaps in β -diversity signals: For example, ordinations showed that sites from the NAST-E and CNRY province overlapped (Fig. C2.4), likely caused by transport and dispersal along the Azores Current which joins the Canary Current (Fedoseev, 1970; Supplementary 2 Fig. S2a). The complex dynamics in the CNRY province (driven, for example, by a weakening of the Canary Current through coastal upwelling and eddy formations as described in accompanying studies of the same sampling campaign; (Bracher et al. 2020; von Appen et al. 2020), likely contributed to increased microbial β -diversity, relative to most other provinces. Similarly, samples in physically energetic provinces (e.g. BRAZ province), defined by an increased scatter in their temperature-salinity profiles (Supplementary 2 Fig. S2, S17), were more distributed in their α -diversity suggesting that the habitat structure is more disrupted or patchy in these ecosystems and high-resolution sampling (< 50 km) is needed for stable biodiversity assessments as shown in other studies across ocean filaments (Fadееv et al., 2021) and fronts (Mousing et al. 2016).

Productivity-specific length scales and their role in microbial biogeography

The ratio of primary productivity rate and current speed gives us a productivity-specific length over which a biological community is carried during a single biomass turnover. This can be interpreted as a scale relating the potential for observed changes in β -diversity to horizontal transport. This may be seen as a time-sensitive version of earlier relationships described between chl *a* patchiness and current speed (Powell et al. 1975). Specifically, we estimated the primary productivity rate (normalized P^B) at each sampling site through ^{13}C stable isotope experiments and biomass assessment (chl *a* concentration). The advective environment in which this growth occurs was characterized through measuring surface current speed – in m d^{-1} – at each sampling point. Overall, the productivity-specific length scale provides a first order estimate of the spatial scale of ecosystem patchiness in epipelagic systems, ultimately controlled by a key biological rate (PP) and horizontal current speed.

Across the Atlantic, we identified three regimes of β -diversity associated with different magnitudes of productivity-specific length scales (high diversity at low productivity-specific length scales (Fig. C2.2c/box 1), low diversity at low productivity-specific length scales (Fig. C2.2c/box 2), and increasing diversity with increasing productivity-specific length scales (Fig. C2.2c/box 3)), affected by different magnitudes of P^B and current speed within different provinces (Fig. C2.2b).

For example, prokaryotic growth in oligotrophic regions with low current speed, such as the SATL-HOT province, has been shown to be controlled by heterotrophic grazing (Teira et al. 2019). In our data, the change in β -diversity within the SATL-HOT province occurred at intermediate productivity-specific length scales with higher P^B , higher proportional heterotrophic biomass, and higher eukaryotic heterotrophic richness relative to other sites (Fig. C2.2b, c/box 2). This can suggest that differences in productivity-specific length scales are a result of greater biological top-down control on local biodiversity than differences in horizontal current speed. The β -diversity signals in provinces with high horizontal current speed increased with increasing productivity-specific length scales (Fig. C2.2c/box 3 SATL-COLD, BRAZ), suggesting a possible shift in mechanisms for β -diversity at higher productivity-specific length scales. The observed increase in β -diversity was likely due to the variable local energetics of the Brazilian Coastal Current (De Souza and Robinson 2004); reflected in a large range of temperature-salinity signals in our dataset, Supplementary 2 Fig. S2). These energetics can result in a confluence zone of communities potentially far from their region of origin (Malvinas-Brazil Confluence; Clayton et al. 2013). However, at this point, we cannot differentiate between the physical and biological mechanisms as we also observed different magnitudes of ordination distances within different trophic functional groups. For most of our sites (Fig. C2.2c/box 2), the length scale was < 5 km and did not correlate with the relatively low β -diversity signals in our RDA analyses (< 0.05 ordination distance). In this case, β -diversity can be controlled by one or more of the multiple environmental variables shaping microbial communities such as vertical mixing (Cheng et al. 2020), atmospheric processes (Mayol et al. 2014), macroorganisms (Troussellier et al. 2017), and anthropogenic impact (Nogales et al. 2011).

It is important to note that at this point, we discuss surface dynamics only, and do not include any consideration of vertical mixing, itself clearly an important feature in generating diversity changes

(DeLong et al. 2006), and likely the source for the great heterogeneity of β -diversity in the CNRY province (Fig. C2.2c/box 1). Understanding the interplay of biological and physical controls would require more complete analyses of biological and physical interactions and is thus only a first-order estimate.

Essentially, the productivity-specific length scale can be seen as defining a relevant scale of the mosaic of the biodiversity signatures measured in this study, revealing where higher-resolution sampling ($< 5\text{km}$) would have added more value to our data set and thus where more careful interpretations of our β -diversity results are necessary. It also provides a perspective on the resolution future campaigns should adopt to develop biogeographic understandings in each region.

We are convinced that Lagrangian sampling – at scales informed by sample-based productivity-specific length scales – could enhance our understanding of microbial pelagic biogeography. This approach would help characterize how far microbial communities in a given water packet could disperse within provinces boundaries, giving shape to their internal mosaics of biodiversity. This is analogous to how the Damköhler number relates exposure timescale and processing timescale of a chemical reaction, to express how spatially and temporally extended the impacts of that reaction will be felt (Oldham et al., 2013).

Province-dependent correlation structures between α -diversity and environmental variables

Across our entire transect, correlations between microbial α -diversity and environmental variables followed well-described latitudinal temperature-diversity relationships (e.g. Fuhrman et al., 2008; Ibarbalz et al., 2019). However, we observed more faceted environment-diversity relationships within provinces, especially within physically energetic provinces (Fig. C2.3b). More measured variables, including temperature, correlated with α -diversity in energetic provinces relative to those with less energetic profiles. This suggests that the temperature-diversity relationship is not controlled by thermal energy alone (see, e.g., Giebel et al., 2009), but is nested in a more faceted microbial response to local water mass characteristics. For example, the high-temperature provinces we sampled were also oligotrophic (low chl a), which would confound any attempt to

independently assess temperature-diversity relationships without accounting for the impact of oligotrophic conditions on α -diversity (Fig. C2.3a; Santi et al., 2019).

Our results showed a nuanced relationship of PP to microbial α -diversity within provinces and do not corroborate previous observations, which detected significant positive correlations (Raes et al., 2018). We observed moderate positive (but not significant at $n = 20$) correlations of PP and prokaryotic heterotrophic α -diversity in low-chl *a* provinces. This may be due to more oligotrophic conditions supporting more even saturation of available niches, rather than boom-and-bust dynamics characteristic of eutrophic environments and events (Duffy and Stachowicz 2006). In contrast, we observed a negative and significant correlation between α -diversity and PP in the highly productive CNRY province, suggesting the rise and succession (via invasibility) of a few, opportunistic phylotypes (Steiner and Leibold 2004), due to input of limiting nutrients from Saharan dust (von Appen et al. 2020). Together, our observations in the Atlantic and those in the Pacific (Raes et al., 2018) suggest that the relationship between PP and microbial α -diversity is nuanced, and observed signals depend on province-specific characteristics, which drive competition.

Notably, the differences in sampling size ($n_{\min} = 8$, $n_{\max} = 20$) of different provinces impact individual correlation and correlation significance. Thus, we excluded the FKLD ($n = 4$) and NATR ($n = 6$) province from our correlation analysis. The great variability in significant correlations between environmental variables and microbial diversity showed that refined observations within provinces are needed to confirm the observed individual environment-diversity relationships.

Low and high chl *a* conditions correspond to contrasting diversity patterns between trophic groups. We observed contrasting β -diversity patterns in microbial functional groups associated with low- and high-chl *a* conditions, supporting the observations by (Irwin et al. 2006) of differences in the importance of environmental predictors for different phytoplankton functional types. Our results show that environmental conditions are important predictors not only on a phylogenetic level (e.g. diatoms vs dinoflagellates, Irwin et al., 2006) but also within trophic functional groups.

We observed that eukaryotic mixotrophic communities had relatively high α -diversity at each site and high β -diversity between sites in low-chl *a* provinces (vs. high-chl *a* provinces), suggesting that mixotrophy is supported under low-nutrient, low chl *a* conditions (see, e.g. Hartmann et al., 2012). Similarly, the β -diversity of cyanobacteria and eukaryotic heterotrophs was greater in sites under low chl *a* conditions, suggesting these functional groups are more susceptible to selective forces in these provinces (e.g. increased cyanobacterial diversity under low phosphate conditions; Thompson et al., 2013).

In contrast, we observed that there is greater β -diversity between sites across high-chl *a* provinces (vs. low-chl *a* provinces) when examining eukaryotic autotrophic and prokaryotic heterotrophic communities (Fig. C2.4a, e), which are known to structure one another (Seymour et al. 2017). However, we did not observe any proportional change in the α -diversity of these groups between low- and high-chl *a* provinces. This suggests that higher productivity is contributing more to species turnover between sites, rather than greater diversity within sites (Vallina et al. 2014). Observing a stable number of Hutchinsonian niches suggests that their occupancy is driven more by stochastic than deterministic processes in these provinces. These processes may include the dilution of slow-growing cells in dynamic systems (Irwin et al., 2006) as well as the prevalence and favoring of *r*-strategists in the community. However, further work is needed to explore these speculations.

Overall, our results extend previous findings (Legendre, 1981) of coupling between microbial diversity and productivity, showing trophic-specific diversity patterns between low-chl *a* and high-chl *a* conditions. Only more temporally and spatially expanded observations (factoring in functional diversity using metagenomics/-transcriptomics, improved hydrographic descriptions, and physicochemical/nutrient profiles in these regions would allow our observations to be more confidently linked to ecosystem states driven by productivity (Chase and Leibold 2002) and broader biogeographic descriptions.

Chapter 2 / 4 Conclusion

Here, we assessed how measures of microbial diversity and activity can better inform ecological partitioning of the ocean, granting new perspectives on functional microbial biogeography across, within, and between provinces. We showed that eukaryotic autotrophs and prokaryotic heterotrophs show similar cross-province β -diversity patterns, distinct from those shared by mixotrophs, cyanobacteria and eukaryotic heterotrophs. Our calculations of a productivity-specific length scale are, to our knowledge, the first attempt to quantify – on a per-sample basis – the influence of surface ocean currents on microbial communities coupled to primary productivity measurements. This provides a first-order estimate of how spatially extended a microbial diversity signature may be, and thus the scale of a recognizable patch in a larger biogeographic area. Our findings also show the value of exploring functional communities (i.e. guilds) of microorganisms to more holistically understand community ecology with phylogenetic diversity data. We conclude that highly resolved sampling of these factors along more Lagrangian designs would help microbial ecology coherently map the subdivisions of the ocean's biogeographic provinces.

Chapter 2 / 5 Materials and Methods

Chapter 2 | 5.1 Sample collection

Our sampling was part of the PS113 (ANT-XXXIII/4) campaign onboard *RV Polarstern* from Punta Arenas, Chile, to Bremerhaven, Germany, from 2018-05-08 to 2018-06-10 (Strass 2018). We took discrete measurements for biophysical analyses of sea surface water at 193 stations in the Atlantic Ocean from about 11 m depth through the ship's seawater system (Teflon® tubing with a membrane pump) at an interval of $\sim 0.5^\circ$ latitude.

Chapter 2 | 5.2 Province delineation after Longhurst

We defined ecological regions based on the variables suggested by Longhurst (Longhurst, 2007), which include gradients in sea surface temperature, salinity, chl *a*, and checked for matches in nutrient concentrations (Supplementary 2 section 1, Fig. S1, Fig. S2). We classified samples into provinces by identifying 1) clusters on a temperature-salinity plot (i.e. by water mass) where clusters were constrained by geographic proximity (Supplementary 2 Fig. S2), and 2) boundary currents that coincided with province boundaries (i.e., where surface velocity vectors (geostrophic currents) were strong at province boundaries as identified above; Supplementary 2 Fig. S3a; Copernicus, 2020). We also identified ocean ridges in bathymetry profiles as features potentially structuring the modified provinces (GEBCO 2019). We compared our classification into Longhurst provinces with delineations proposed by Bracher et al. (2020) (Supplementary Material Table S6; Hörstmann et al. 2021).

We classified provinces as high-chl *a* provinces if their mean chl *a* concentrations were above $0.3 \mu\text{g m}^{-3}$ chl *a*, or as low-chl *a* provinces if their mean concentrations were below $0.1 \mu\text{g m}^{-3}$ chl *a*. Provinces with mean concentrations between 0.1 and $0.3 \mu\text{g m}^{-3}$ were treated as ambiguous, and we defaulted to classifications in (Longhurst, 2007). For details see Supplementary 2 section S1.

We calculated the distances to the coast and to province boundaries in qgis 3.14 (QGIS Geographic Information System, 2020) by densifying the vector coastlines and extracting vertices, followed by extracting the distance from each site points to the nearest hub on the coast contours. We used the “110 m vector coastline” (v4.1.0, Natural Earth Data) and the province boundaries (delineated above) in our calculations.

Chapter 2 | 5.3 Biochemical and primary productivity profiling

At each station, we collected biochemical samples for dissolved inorganic nutrients (silicate, phosphate, nitrate, and nitrite) as well as particulate organic matter (POC and PN). For details see Supplementary 2 section S2.

We measured primary productivity (PP) with 200 $\mu\text{mol NaH}^{13}\text{CO}_3$ stable isotope incubations in triplicates over a time period of ~24 h. For details on experimental design see Supplementary 2 section S3 and incubation conditions are documented online in the Supplementary Material Table S7 (Hörstmann et al. 2021). Analysis of ^{13}C incorporated into organic matter was carried out on a PDZ Europa ANCA-GSL elemental analyzer interfaced to a PDZ Europa 20-20 isotope ratio mass spectrometer (Sercon Ltd., Cheshire, UK) by the Isotopic Laboratory at the UC Davis, California campus. PP was calculated as in equation (C2.1):

$$PP = \frac{\text{at\%(enriched)} - \text{at\%(NA)} * POC}{(100 * (DIC(13C) \div (DIC(13) + DIC(12C))) - \text{at\%(NA)} * Incubation\ time)} \quad (\text{C2.1})$$

with at%(enriched) as the atom percent derived from ^{13}C -spiked samples; at%(NA) is the atom percent of natural abundance derived from particulate organic carbon (POC) samples. Dissolved inorganic carbon (DIC) is assumed as 2000 $\mu\text{mol kg}^{-1}$ (after Zeebe and Wolf-Gladrow, 2001). Incubation time is the time between sample spike and filtration (~24 h).

We calculated specific primary productivity (P^B) based on PP and chl *a* concentration, as shown in equation (C2.2).

$$P^B = \frac{PP}{chl\ a} \quad (\text{C2.2})$$

Where PP is expressed in $\text{mg C m}^{-3} \text{ d}^{-1}$ and chl *a* concentration in $\text{mg chl } a \text{ m}^{-3}$, resulting in a P^B value in $\text{mg C m}^{-3} \text{ d}^{-1} \text{ chl } a^{-1}$.

Productivity-specific length scale

To link the biomass turnover of microbial communities to their pelagic advection, we calculate a length scale for biological-physical coupling (productivity-specific length scale) using a quantitative comparison of measured biomass turnover and measured current speed.

$$\text{productivity} - \text{specific length scale} = \ln(2) * \text{current speed} * \lambda^{-1} \quad (\text{C2.3})$$

Where the productivity-specific length scale is expressed in km. current speed reflects average horizontal current speed from the VM-ADCP which provided data for the depth range 20–50 m. The λ is derived as in equation (C2.4).

$$\lambda = \frac{PP}{\text{standing stock biomass}} \quad (\text{C2.4})$$

PP is expressed in $\text{mg C m}^{-3} \text{ d}^{-1}$ and the standing stock biomass corresponds to the chl *a* concentration (in mg m^{-3}) times $23 \text{ mg chl } a (\text{mg C})^{-1}$ (for conversion between chl *a* and carbon concentrations (Geider 1987); validated against in situ POC concentrations; Supplementary 2 Table S4).

Chapter 2 | 5.4 Microbial sampling, processing, and amplicon sequence analyses

Flow Cytometry

At each station, samples for cell counts were taken into 2 ml Eppendorf tubes, and incubated in the dark prior to fixation using 0.2% paraformaldehyde. Samples were snap-frozen and stored at -80°C while at sea. Samples were analyzed on a BD Accuri™ C6 Plus Flow Cytometer (BD Biosciences-US) according to Gasol and Moran, (2015). For details on measurements see Supplementary 2 section S4.

Microbial DNA sampling

For 16S and 18S rRNA gene sequencing analysis, 4 l of seawater were filtered through $0.2 \mu\text{m}$ Sterivex® filters. Filters were purged with air, tightly closed, snap-frozen in liquid nitrogen, and stored at -80°C until further analysis.

DNA extraction and amplicon sequencing

DNA was extracted using a DNeasy® PowerWater® DNA extraction kit (QIAGEN, Valencia, CA, USA, Catalog No./ID: 14900) following the manufacturer's instructions. After DNA extraction, DNA concentration was quantified using a Quantus™ Fluorometer and normalized to $2 \text{ ng } \mu\text{l}^{-1}$. Amplicons targeting the variable region 4 (V4) of the bacterial 16S rRNA gene (515F–806R) (Caporaso et al. 2016) and the variable region 4 (V4) of the eukaryotic 18S rRNA gene (TA-

Reuk454FWD1–TAREukREV3) (Stoeck et al. 2010) were generated following standard protocols of amplicon library preparation (16S Metagenomic Sequencing Library Preparation, Illumina, Part #15044223 Rev.B) and sequenced using a MiSeq Sequencer (Illumina). 16S and 18S rRNA gene amplicon reads were generated using 300-bp paired-end sequencing using Nextera XT Index Kit v2 Set A-B (Illumina) index primer. Samples were demultiplexed using bcl2fastq (Illumina) with barcode mismatches set to 1.

ASV tables for both 16S and 18S rRNA gene amplicon sequences were constructed using the DADA2 R package, v1.15.1 (Callahan et al. 2016). Samples were processed using the DADA2 pipeline (v.1.15) with an additional step where primers were trimmed using cutadapt v1.18 after pre-filtering of FASTQ files. Diagnostics of each filtering step are documented online in the Supplementary material Table S9 (Hörstmann et al. 2021) and number of reads plotted for each filtering step in density plots (Supplementary 2 Fig. S18).

Taxonomic assignment and functional grouping

Taxonomic assignment was performed outside DADA2 using SilvaNGS (v1.4) (Quast et al. 2013) pipeline for 16S rRNA gene data with the similarity threshold set to 1. Reads were aligned using SINA v1.2.10 (Pruesse et al. 2012), and classified using BLASTn (v2.2.30) (Camacho et al. 2009) with the Silva database (v132) as a reference database. For taxonomic assignment of 18S rRNA gene amplicons, we used the QIIME 2 Plugin 'feature-classifier' (v2019.7.0 from package 'q2-feature-classifier') in qiime 2 (Bokulich et al. 2018) and the pr2 database (v4.12) (Guillou et al. 2013).

We classified eukaryotic taxa as either autotroph, mixotroph, heterotroph, or unknown if no information of trophic status was available by performing an unstructured literature search to validate expert-led assignment by UJ (see outcomes in Supplementary 2 Table S5). Prokaryotes, except Cyanobacteria, were considered to be dominated by a heterotrophic lifestyle (*sensu* Herndl et al., 2008), see Supplementary 2 Table S5 for further details.

Chapter 2 | 5.5 Ecological data analyses

We performed our data analysis using R v4.0.3 (R Core Team, 2020) and RStudio v1.4.1103 (RStudio Team, 2021). Statistical analysis code is publicly archived on github (CoraHoerstmann/AtlanticMicrobiome) in the tagged release: <https://zenodo.org/badge/latestdoi/329689664>. Biophysical parameters, concentrations of particulate organic matter, dissolved inorganic nutrients, microbial cell counts, and PP measurements were used as environmental metadata for statistical analyses with microbial sequencing data.

Alpha diversity (α -diversity) was determined by calculating Hill numbers (Chao et al. 2014) of richness, Shannon entropy, and Simpson concentration ($q = 0$, $q = 1$, and $q = 2$), using the iNEXT package (v2.0.20) repeating calculations 100 times (AtlanticMicrobiome/alpha_diversity/alpha_diversity.R). Rarefaction curves of all Hill numbers were plotted using *fortify()* of the ggplot2 package (v3.3.3). Further analyses were performed using Shannon diversity measures of each trophic functional group, as it reflects true diversity (richness + evenness) and is less susceptible to fluctuations in rarer phylotypes.

Pearson correlations between microbial Shannon diversity and environmental parameters (salinity, temperature, PO_4 , H_4SiO_4 , NO_3 , dissolved oxygen, distance to province boundary, distance to coast, PN, POC, chl *a*) were calculated and plotted using the *corrplot()* function of the corrplot package (v0.84) and adjusted for multiple testing using the Holm-Bonferroni method (AtlanticMicrobiome/correlation_analyses/general_corrplot.R). Residuals of all correlations were screened for notable departures from Gaussian distributions, in which case these correlations were excluded from our results (AtlanticMicrobiome/correlation_analyses/Residuals_corrplot.R; Supplementary 2 Fig. S8–S15).

ASV tables and environmental metadata were transformed for comparability and statistical downstream analyses (general_clr_hellinger_transformations.R, z-scoring_subset.R in AtlanticMicrobiome/data_transformations/). Before transformations, we removed all ASVs with ≤ 3 instances across all samples. Further, before CLR transformation, we performed Bayesian-

multiplicative treatments of zeros in the ASV tables using *cmultRepl()* function of the *zComposition* package (v1.3.4): this uses sample-wise totals to convert zero counts (which will lead to errors in log-ratios) into near-zero estimates, assuming undersampling rather than absence. To account for compositionality effects in our ASV tables (see Gloor et al., 2017), we performed a CLR-transformation for RDA. Prior to PERMANOVA analyses using the *decostand()* function in *vegan* (v2.5.6), Environmental variables were checked for normal distribution (Supplementary 2 Fig. S19), and z-scored for scale-independent intercomparability.

We examined the distribution of sites among environmental gradients using a PCA on our microbial diversity metadata. We tested differences of ocean provinces using two-sided Wilcoxon rank sum tests (hereafter: Wilcoxon) (*AtlanticMicrobiome/environmental_analyses/PP_plots.R*).

To examine microbial beta diversity (β -diversity) and its relation to environmental and contextual variables, we performed a set of RDA using the CLR-transformed ASV tables as response matrices and tables of environmental variables as explanatory matrices. We calculated sun azimuth and altitude based on sampling time and location using the *suncalc* (v.0.5.0) package in R. We performed stepwise model to identify significant environmental variables using *ordiR2step()* function in *vegan* (*AtlanticMicrobiome/multivariate_analysis_ordistep.R*). Differences of microbial dissimilarity between ocean provinces were tested with a permutational MANOVA (PERMANOVA) (Anderson 2001) on Aitchinson distance of the CLR-transformed ASV tables using the *adonis2()* function along with a beta dispersion test to evaluate the homogeneity of dispersion using the *betadisper()* function in *vegan* (*AtlanticMicrobiome/multivariate_analysis/RDA_functions.R*).

Data availability.

Chlorophyll *a* (chl *a*) data was derived from (<https://doi.org/10.1594/PANGAEA.913514>) and published in (Bracher et al. 2020). Dissolved inorganic nutrients, Particulate organic matter and primary productivity measurements are publicly archived (<https://doi.org/10.1594/PANGAEA.926458>, <https://doi.org/10.1594/PANGAEA.926460>, and <https://doi.org/10.1594/PANGAEA.926462>, respectively). Sequence data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number

PRJEB42499 (<https://www.ebi.ac.uk/ena/browser/text-search?query=PRJEB42499>), using the data brokerage service of the German Federation for Biological Data (GFBio, (Diepenbroek and Glöckner 2014), in compliance with the Minimal Information about any (X) Sequence (MIxS) standard (Yilmaz et al. 2011). The analysis based on satellite data was generated using E.U. Copernicus Marine Service Information and are available at: <http://marine.copernicus.eu> (Copernicus, 2020) using the GLOBAL_ANALYSISFORECAST_PHY_CPL_001_015 data (Lea et al. 2015). The 110 m vector coastline (ESRI shapefile; v4.1.0) is available at <https://www.naturalearthdata.com/>.

Code availability.

Statistical analysis code is publicly archived on github in the tagged release: <https://github.com/CoraHoerstmann/AtlanticMicrobiome>

CHAPTER 3: Marine-terminating glaciers structure Arctic and sub-Arctic picoplankton diversity

Authors: Cora Hörstmann, Tore Hattermann, Pauline Thomé, Pier Luigi Buttigieg, Isidora Morel, Anya M. Waite, Uwe John

Abstract

Arctic fjords are complex marine ecosystems at the interface between terrestrial and marine systems experiencing the highest air and water temperature increase globally. These pristine ecosystems are some of the most productive high-latitude systems and are sustained by a diverse microbial community at the base of the marine food web. Due to the unique features of Arctic fjords, most studies concern a single or few fjords, and comparative studies incorporating connectivity paths and community networks are lacking. Furthermore, it is unknown how marine-terminating glaciers impact key ecosystem attributes such as species co-occurrence structure and microbial community productivity. Here, we investigated the eukaryotic and prokaryotic picoplankton alpha and beta diversity structures of 93 surface water samples collected from 21 Arctic and sub-Arctic fjords between 2012 and 2019. We modeled oceanographic connectivity between individual fjords and detected increasing dissimilarities in microbial community structures between fjords with greater hydrodynamic separation. Additionally, a multivariate analysis revealed that the microbial community structure was highly reflective of the presence or absence of marine-terminating glaciers. Furthermore, we observed a more prokaryotic-dominated co-occurrence network within fjords without marine-terminating glaciers. This suggests that further loss of marine-terminating glaciers in fjord systems could induce a baseline community shift from more eukaryotic- to prokaryotic-dominated community structures and corresponding changes in

ecosystem functions. Our findings also suggest that these potential community changes could be further amplified by glacier melt and oceanographic transport of taxonomic groups. These effects are likely to be more pronounced for prokaryotes, whose dispersal is currently more restricted by temperature barriers than is the dispersal of eukaryotes. The present study refines understanding of the microbial picoplankton community structure, with potential implications for future scenarios and shifts at the base of the marine food web. Understanding these mechanisms will require more refined microbial observations along with time-resolved oceanographic models.

Chapter 3 / 1 Introduction

Arctic fjords are among the most productive regions of high-latitude ocean systems (Smetacek and Nicol 2005). Due to increasing anthropogenic climate change, fjords have been subject to severe ecosystem changes such as glacial retreat, changes in freshwater input, and altered matter exchange between terrestrial and coastal ocean systems (Hopwood et al. 2020). Glacial retreat and associated environmental changes in fjord systems pose a serious threat to local biodiversity and ecosystem function (Szeligowska et al. 2021).

Microorganisms, the key players in marine primary productivity, form unique, regional communities in sub-Antarctic fjords (Maturana-Martínez et al. 2021). However, the structural impact of marine-terminating Arctic glaciers on microbial diversity is incompletely understood, limiting the ability to project how glacial retreat will impact fjord ecosystems.

Arctic fjords are characterized by steep spatial and temporal environmental gradients, including strong seasonality (Marquardt et al. 2016) and strong physical gradients along the fjord length, which are driven by dynamic inputs of freshwater and nutrient release from glaciers and river runoff (Müller et al. 2018; Kim et al. 2020b). Additionally, anthropogenic climate change has regionally increased heat transport through poleward extension of Atlantic water, referred to as Atlantification (Polyakov et al. 2018), altering the ‘trails of life and death’ of species into the Arctic Ocean

(Wassmann et al. 2015). Increased Atlantification can cause northward shifts of key species in marine biogeochemical cycles such as *Emiliania huxleyi* (Neukermans et al. 2018; Oziel et al. 2020), and increased dominance of *Phaeocystis pouchetii* (Nöthig et al. 2015) with potential alterations in energy transfer to higher trophic levels. Moreover, warmer surface water temperatures can increase the metabolic rates of picoplankton, ultimately increasing primary productivity (Frey et al. 2018) and altering the local carbon pool (Azzaro et al. 2021). At the community level, the microbial food web of Arctic estuaries is dominated by parasites during the winter season, and the co-occurrence of parasites with microbial species defines the majority of species interactions in genomic network analyses, shifting in summer to detritus-dominated microbial communities (Kellogg et al. 2019). Detailed ecological studies at multiple sites with different oceanographic, glacial, and environmental conditions are needed to elucidate the complex climate-related changes in microbial diversity and associated changes in food webs and carbon-cycling.

Analysis of microbes and their role in microbial food webs requires trait-based approaches, especially considering the large size range and differential ecological roles of microbes (Massana and Logares 2013). Among typical planktonic size classes (micro-, nano-, and picoplankton), pico-eukaryotes, (here defined by a cell size of 0.2–3 μm), and prokaryotes, share several traits and make up a large fraction of the photoautotrophic biomass (Li et al. 2013). Furthermore, picoplankton are likely to play an increasingly important role in the Arctic Ocean in the future due to their high nutrient affinity (Li et al., 2009), large effective population size, and high genetic diversity (Elferink et al. 2017).

Here, we investigated the impact of marine-terminating glaciers on picoplankton assemblages of multiple Arctic and sub-Arctic fjords, which included samples from 93 sites in 21 glacial- and non-glacial-influenced fjords in Sweden/south Norway, north Norway, Svalbard, Iceland, east Greenland, and west Greenland. We modeled oceanographic connectivity between these regions, hypothesizing that microbial beta diversity changes would be related to regional and pan-Arctic oceanographic transport of microorganisms, while being limited by local environmental selection processes (Wassmann et al. 2015). We applied both multidimensional analyses of prokaryotic and eukaryotic 16S and 18S rRNA gene sequences, together with environmental variables, and co-

occurrence network analyses of prokaryotic and eukaryotic taxa. We hypothesize that fjords with and without glaciers have distinct beta diversity patterns and co-occurrence networks in both prokaryotes and eukaryotes, with microbial communities in glacier-influenced regions being more sensitive to freshwater input and the potential upwelling of deep communities (Maturana-Martínez et al. 2021). Further, similar to seasonal (Kellogg et al. 2019) or regional sea ice-influenced water masses (Fadeev et al. 2018), differences in dominating taxa and co-occurrence patterns could repeat regionally due to different glacier-influenced environmental conditions (Seifert et al. 2019). Our analysis contributes to an improved understanding of the microbial ecosystem structure on a pan-Arctic scale in an ecologically-sensitive region subject to rapid and dramatic climate-driven changes.

Chapter 3 / 2 Materials and Methods

Chapter 3 | 2.1 Sampling

Data analyzed in this study include datasets from five expeditions in Arctic and sub-Arctic coastal waters and fjord waters in the northern hemisphere during spring and summer between 2014–2019. The expeditions took place on the research vessels *RV Maria S. Merian* (MSM) and *RV Heincke* (HE). Specifically, this dataset consists of environmental (temperature, salinity, fluorescence, PO₄, NO₃, and Si) and picoplankton DNA data from the following cruises: HE431 in north Norway, south Norway, and Sweden in 2014; HE492 in Svalbard in 2017, HE533 in north Norway in 2019; MSM21-3 in Iceland and west Greenland in 2012 and MSM56 in Svalbard and east Greenland in 2016. Sampling and processing of all environmental and sequence data were previously published elsewhere (Elferink et al. 2017, 2020; Seifert et al. 2019) or are described in Supplementary Section A (HE533, HE492). We sub-sampled one replicate (replicate A) for sites with multiple replicates. Additionally, we excluded clear outliers from analysis, which were due to a *Crysocromulina* spp. bloom (John et al. *in prep.*) during the HE533 expedition.

Chapter 3 | 2.2 Microbial DNA processing, amplicon sequencing, amplicon sequence analyses, and taxonomic assignment

16S and 18S rRNA metabarcoding

To extract DNA from the 0.2 µm polycarbonate filters, the Genomic DNA from soil (NucleoSpin® Soil) kit was used following the manufacturer's protocol with a minor modification: the sample lysis was conducted with a bead beater (MagNA Lyser, Roche) for 2×30 seconds at 55,000 rpm.

The variable region 4 (V4) of the small subunit ribosomal RNA gene (16S for prokaryotes and 18S for eukaryotes) was used as a molecular marker to determine the taxonomic community composition. Primers were selected in accordance with the Earth Microbiome Project (<http://www.earthmicrobiome.org/protocols-and-standards/>) using prokaryotic primers (515F–806R) (Caporaso et al. 2016) and eukaryotic primers (TA-Reuk454FWD1–TAREukREV3) (Stoeck et al. 2010) with overhanging Illumina adapters. The library preparation preceding the sequencing followed standard protocols (16S Metagenomic Sequencing Library Preparation, Illumina, Part #15044223 Rev.B; Illumina Technology).

The amplicon libraries were paired-end sequenced on the Illumina MiSeq sequencing platform. The prokaryotic samples were sequenced at the Alfred Wegener Institute in Bremerhaven, Germany, and the eukaryotic samples were sequenced at the Leibniz Institute on Aging (FLI) in Jena, Germany, using 300-bp paired-end sequencing on a MiSeq Sequencer (Illumina) with a MiSeq Reagent Kit v3 (600-cycle). All samples were demultiplexed using bcl2fastq (Illumina) with barcode mismatches set to 1.

16S/ 18S rRNA sequence processing and taxonomic assignment

Amplicon sequence variants (ASVs) were obtained by processing the resulting raw paired-end reads with R (R Core Team, 2013) package DADA2 v1.16.0 (Callahan et al. 2016) following a modified version of the DADA2 Pipeline Tutorial (v1.8). Processing of 18S rRNA gene reads and 16S rRNA gene reads were performed separately. Reads were pre-filtered by length (*minLen* = 50) and quality (*minQ* = 2), followed by removal of the primers. The pre-filtered reads were further filtered by the expected length of the amplicon (240–160 bp for 16S rRNA V4 and 270–220 bp for

18S rRNA V4) and quality, for which the maximum number of expected errors (*maxEE*) was set to 2.7 for forward reads and 2.2 for reverse reads. De-replication, error learning, and sample inference were then performed on the filtered reads. To obtain the full denoised sequences, the paired-end reads were merged with a minimum overlap (*minOverlap*) of 20 bp. Finally, chimeras were removed and the ASV tables were built. Prokaryotic and eukaryotic taxonomy were assigned against the Silva v132 database and the pr2 database v4.12, respectively. We removed all ASVs that were annotated as mitochondria or chloroplasts from the 16S rRNA ASV tables, and ASVs annotated as metazoans from the 18S rRNA ASV table.

Oceanographic dispersal

Numerical simulations were performed using the output of a 3D, high-resolution, general ocean circulation model of the Arctic Mediterranean to compute large ensembles of Lagrangian drifter trajectories and assess the potential dispersal of microorganisms and map oceanographic connectivity between the sampling sites.

The hydrodynamic model used to represent the ocean currents in the study area was based on metROMS (<https://doi.org/10.5281/zenodo.290667>), which couples the state-of-the-art Regional Ocean Modeling System (ROMS, <http://myroms.org>), a free-surface, hydrostatic, primitive equation ocean general circulation model (Shchepetkin and McWilliams 2005), with the comprehensive dynamic-thermodynamic sea ice model CICE (<https://zenodo.org/record/1205674>). metROMS was run with a horizontal resolution of 4×4 km in an orthogonal, curvilinear grid covering the entire Arctic Mediterranean over 2005–2017, referred to as the “A4-setup” (Isachsen 2015; Hattermann et al. 2016). A4's initial state and boundary conditions were derived from monthly-averaged global reanalyses (Storkey et al. 2010) and additional forcing along its open boundaries using the global TPXO tidal model (Egbert and Erofeeva 2002). Atmospheric forcing was conducted via 6-h ERA-Interim reanalysis (Dee et al. 2011). Riverine freshwater input was provided in the form of monthly climatologies from 26 major rivers encircling Northern Europe and the Arctic Mediterranean and additional glacial runoff estimates from Greenland and Svalbard. Output from the A4-setup contained velocity fields in 32 terrains following vertical layers, and a temporal resolution of 24 h. The first two model years were

discarded as spin-up, and validation against available observations confirmed that the model satisfactorily reproduced the general features of currents, hydrography, and sea ice.

Synthetic floats were introduced to A4 and tracked using TRACMASS (www.tracmass.org; Döös 1995) while they were freely advected by the model's daily averaged velocity field. Cohorts of approximately 1,000 floats were seeded every 10 days, evenly distributed between 5 m and 20 m depth in an area of 3×3 model grid cells centered on each sampling site and tracked over a lifetime of 5 years. To assess the statistical dispersal due to time-varying ocean currents, for each sampling site, time series of ensemble-mean drifter concentrations were computed on the A4 grid by counting the number of drifters of a given age per grid cell from all cohorts, divided by the total number of drifters of that age and all cohorts within the model domain (Supplementary 3 Fig. S1). Based on these calculations, the matrices of oceanographic connectivity between release and receiving sampling sites were obtained (Supplementary 3 Fig. S2–S7) by averaging the drifter concentration of a given release site within a 15×15 cell area centered on the receiving site. Station HE533.F02.28A (Tanafjord) and HE431.F02.19 (Sognefjord) were excluded from this analysis, as these stations were located close to land and could not be resolved by the A4 coastline geometry. Because drifter counts generally decrease exponentially as drifters disperse away from their origin, a logarithmic scale was used to compare drifter concentrations.

Statistical analyses

We performed all data analysis using R v4.1.0 (R Core Team, 2021) and RStudio v1.4.1106 (RStudio Team, 2021). All code publicly archived AGJohnAWI/ArcticPicos under <https://doi.org/10.5281/zenodo.5781579>.

ASV tables and environmental metadata were transformed for comparability and downstream statistical analyses. Before transformations, we removed all ASVs with ≤ 1 instances for eukaryotes and prokaryotes, and performed Bayesian-multiplicative treatments of zeros using the *cmultRepl()* function of the zComposition package (v1.3.4). We performed CLR-transformation of the ASV tables for redundancy analysis (RDA). Further, we calculated the Aitchinson distance of the ASV tables for distance analysis (ArcticPicos/data_transformations).

Hill numbers for alpha diversity sample richness and Shannon entropy (Chao et al., 2014) were calculated using the iNEXT package (v2.0.20) with 100 iterations. Pielou evenness was calculated by dividing Shannon evenness by sample richness. Differences in alpha diversity scores were tested with the F-test for sample variances followed by a two sample t-test with `var.equal = FALSE` (ArcticPicos/alpha_div).

We tested the effect of hydrodynamic connectivity on microbial beta diversity changes through correlation analysis. Specifically, we turned the oceanographic connectivity matrices for each temporal bin into hydrodynamic distance matrices by calculating the inverse of the \log_{10} of the synthetic particles between each site pair. The total number of observed site pairs was 210. To test the effect of oceanographic connection on microbial beta diversity, we removed sites that were not oceanographically connected from analysis. We then normalized the data to the range between 0 and 1, with values close to “0” indicating small hydrodynamic distance and values close to “1” indicating large hydrodynamic distance. We calculated the Pearson correlations between the normalized hydrodynamic distances and 16S rRNA and 18S rRNA Aitchinson distances using the *cor.test()* function of the stats package (v4.1.0) (ArcticPicos/analysis_distance).

We identified significant explanatory variables for microbial beta diversity structure using the permutational analysis (*ordiR2step()* function with `perm.max` set to 200), and investigated microbial beta diversity and contextual environmental data using RDA (ArcticPicos/analysis_multivariate). Residuals of RDAs were checked for normal distribution (Supplementary 3 Fig. S8). We tested differences between sites based on ASV tables between geographic regions, fjords, and regions with and without marine-terminating glaciers with a permutational ANOVA (PERMANOVA) using the *decostand()* function in *vegan* (v2.5.7). Environmental variables were checked for normal distribution and z-scored for scale-independent intercompatibility. Further, we produced binary (presence/ absence) ASV tables and calculated the number of overlapping ASVs between sites with marine-terminating glaciers and without marine-terminating glaciers using the *upsetR* package (v1.4.0).

We created cross-domain microbial networks of glacial- and non-glacial-influenced regions of merged eukaryotic and prokaryotic CLR-transformed ASV tables (*sensu* Tipton et al. 2018). We

imported ASV tables, contextual metadata, and taxonomy tables into a phyloseq object using the phyloseq package (v1.36.0) (ArcticPicos/analysis_network). ASVs were agglomerated at the genus level using the *tax_glom()* function in phyloseq and split into glacier- and non-glacial-influenced regions using the *phyloseq_sep_variable()* function in metagMisc (v0.0.4). If taxonomic assignment at the genus level was missing, we added a class level denoted by “f__” and an individual number. Microbial networks were calculated and visualized using the NetCoMi package in R (v1.0.2) (Peschel et al. 2020). Conditional dependencies of ASVs were calculated based on SPIEC-EASI (Sparse InverseE Covariance estimation for Ecological Association and Statistical Inference) associations using the *netConstruct()* function with sparsification set to 0.3 and using the “multRepl” option for zero handling. Networks were analyzed using the *netAnalyze()* function with the cluster method set to “cluster_fast_greedy”, hubs set to nodes with simultaneously characterized by highest degree, betweenness, and closeness centrality (*hubPar* = c(“degree”, “between”, “closeness”)), and nodes identified as hubs if the node’s centrality value was above the 90% quantile (*hubQuant* = 0.9).

Chapter 3 / 3 Results

Chapter 3 | 3.1 Geographic location and large-scale oceanographic-related community dissimilarity

We analyzed a total of 93 surface ocean samples from Arctic and sub-Arctic fjords of five expeditions between 2012 and 2019 in the northern hemisphere spring and summer. Geographically, the samples were clustered together in six geographically distinct regions: north Norway, Sweden/south Norway, Svalbard, Iceland, east Greenland, and west Greenland (Fig. C3.1a, Supplementary 3 Table S1). These regions are oceanographically influenced by cold and low-salinity polar water and warmer and more saline Atlantic water, as indicated by their varying temperature-salinity profiles (Supplementary 3 Fig. S9).

Dissolved inorganic nutrient (NO_3 , PO_4 , and Si) concentrations had high regional variability (Fig. C3.1b). High internal variability was detected within fjords in north Norway, the Orust-Tjörn Fjord (Sweden), and Disco Bay (west Greenland), primarily due to relatively higher concentrations in the fjord tips (Supplementary 3 Table S1). Due to the high internal variability, we could not identify common differences between regions with and without marine-terminating glaciers.

We simulated the trajectories of synthetic Lagrangian drifters to estimate the potential transport of microbial communities with ocean currents between 2007 and 2018 (Fig. C3.1c). Drifter concentrations were averaged and color-coded based on groups of release sites, and their occurrence at other fjords was recorded if this resulted in oceanographic time-dependent connectivity matrices (Supplementary 3 Fig. S3–S8). Drifters released for ≤ 3 months exhibited strong regionality (e.g., only fjords within northern Norway were interconnected). The primary trajectory pathways at > 6 months were (1) from southern Norway to northern Norway; (2) from northern Norway along the Barents Sea towards west of Svalbard, and from Svalbard (and Iceland) along the east Greenland Current towards west Greenland (Fig. C3.1c, Supplementary 3 Fig. S9). At the maximum simulation time of 5 years, drifters exhibited a widely dispersed distribution across the Arctic Ocean with Svalbard, east and west Greenland fjords receiving drifters from southern Norway/Sweden, northern Norway, and Iceland (Supplementary 3 Fig. S9, S10).

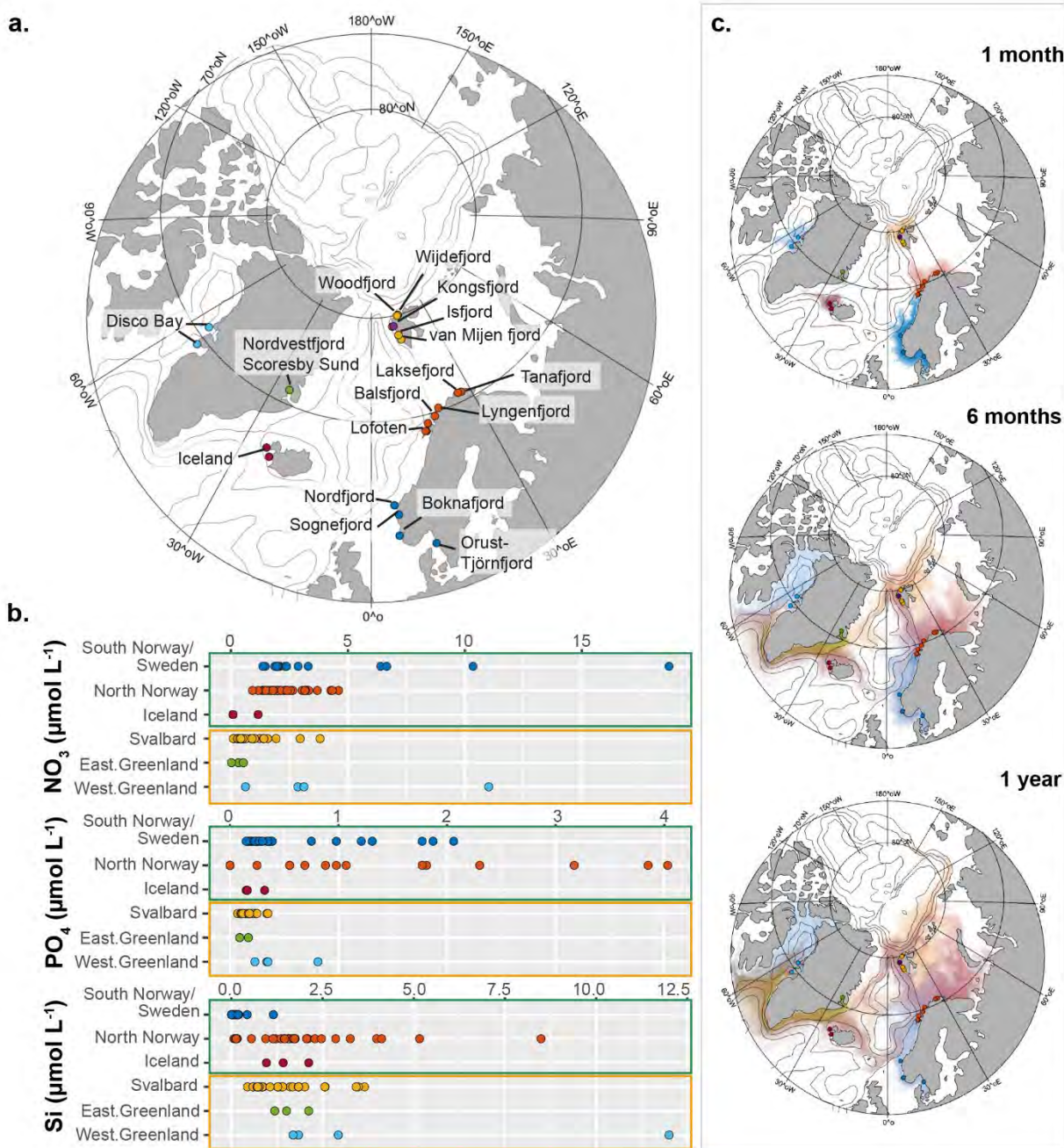


Figure C3.1. **a.** Map of fjords. Sample sites, color-coded to different regions. **b.** Dissolved inorganic nutrient concentrations of nitrate (NO_3), phosphate (PO_4), and silicate (Si) in $\mu\text{mol L}^{-1}$. Sites are color-coded as in **a.**, dark green boxes indicate regions without marine-terminating glaciers, orange boxes indicate regions with marine-terminating glaciers. **c.** Trajectories of modeled drifters mapping oceanographic connectivity of individual fjords after 1 month, 6 months, and one year (see Figure S2 - S7 for connectivity matrices). See Fig. S10 for full tracking (up to 5 years). Sites and drifters are color-coded according to geographic regions of release where darker colors indicate higher drifter concentrations.

Chapter 3 | 3.2 Impact of marine-terminating glaciers on picoplankton beta diversity structure and oceanographic connectivity

We observed both regional and intrafjord microbial beta diversity signals in our redundancy analyses (RDA) (Fig. C3.2a, c), which were also significant in permutational ANOVA analyses (PERMANOVA of geographic regions: prokaryotes (permutations = 999, $R^2 = 0.5$, $p = 0.001$); eukaryotes (permutations = 999, $R^2 = 0.6$, $p = 0.001$); PERMANOVA of individual fjords: prokaryotes (permutations = 999, $R^2 = 0.8$, $p = 0.001$); eukaryotes (permutations = 999, $R^2 = 0.8$, $p = 0.001$)). Notably, within the eukaryotic RDA sites from east Greenland ordinated farther away than other samples along the second RDA axis (Fig. C3.2a). Within the eukaryotic RDA, we found one outlier in the regional groupings from Disko Bay (Station MSM21.F02.514) that ordinated closer to samples from northern Norway than other sites from west Greenland. Within prokaryotes, we found an outlier (MSM56.F02.553) that ordinated closer to samples from Iceland than to other samples from Kongsfjorden (Svalbard).

We found that sites grouped into a cluster of sites with and without glacial influence in our RDA for both prokaryotes and eukaryotes (Fig. C3.2). Similarly, the PERMANOVA, which tested the glacial influence on beta diversity, produced significant results (eukaryotes: permutations = 999, $R^2 = 0.2$, $p = 0.001$, $n = 90$; prokaryotes: permutations = 999, $R^2 = 0.2$, $p = 0.001$, $n = 95$). Among eukaryotes, sites separated into regions with and without glacial influence along the first RDA axis, which captured 46.5% of variation. Among prokaryotes, sites with and without glacial influence separated along the first and second RDA axis. Ordination of prokaryotic beta diversity sites further separated along the first RDA axis into the different geographic regions (Sweden/southern Norway, northern Norway, Iceland), while regions with glacier influence primarily overlapped with each other (Fig. C3.2c).

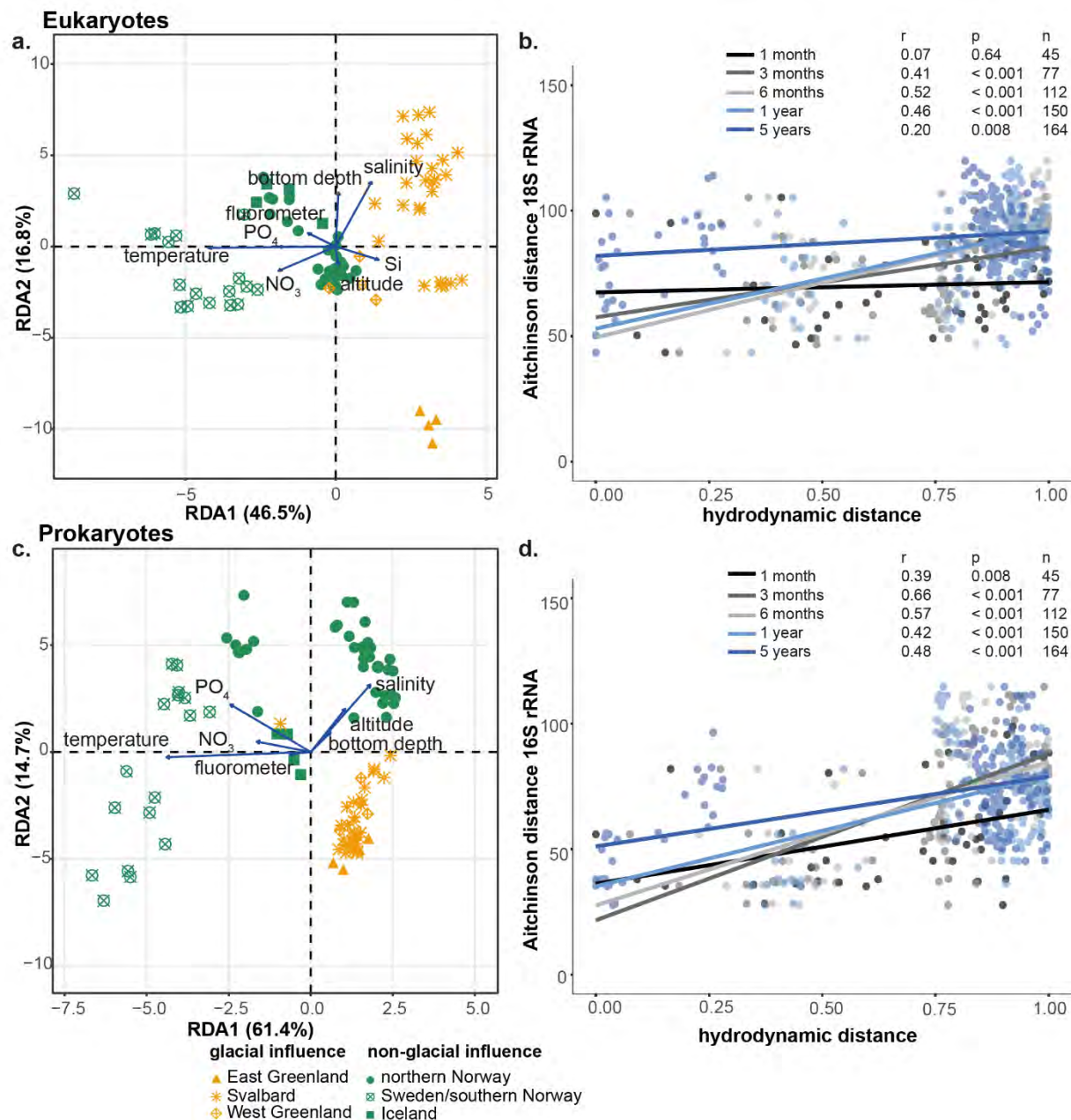


Figure. C3.2 **a.** Redundancy analysis (RDA) of CLR-transformed eukaryotic ASV table; **b.** Aitchinson distance analysis of eukaryotic microbial communities (18S rRNA sequences) against hydrodynamic distance defined as the inverse of the \log_{10} of the synthetic drifter concentration normalized to the range from 0 to 1. Correlations are color-coded and Pearson correlations calculated for each temporal bin. Sites without oceanographic connection are not included; **c.** RDA of CLR-transformed prokaryotic ASV table. Colors indicate regions with glaciers and shape corresponds with geographic regions; **d.** Aitchinson distance analysis of prokaryotic microbial communities (16S rRNA sequences) against hydrodynamic distance defined as the inverse of the \log_{10} of the synthetic drifter concentration normalized to the range from 0 to 1. Correlations are color-coded and Pearson correlations calculated for each temporal bin. Sites without oceanographic connection are not included.

Almost all environmental variables, except salinity and presence of silicate in eukaryotic RDA, were positively associated with sites without glacial influence (Fig. C3.2a, c). Silicate was positively associated with the first RDA axis in the eukaryotic analysis, and thus predominately positively associated with glacial-influenced sites (Fig. C3.2a). In our prokaryotic RDA, silicate had no significant association (stepwise permutation model analysis). Sites belonging to Sweden/southern Norway, a relatively warm and nutrient-rich region in the dataset, were strongly associated with temperature, nitrate concentrations, and phosphate concentrations.

Despite the different salinities of Atlantic and Arctic water masses and their different proportions in glacial and non-glacial sites (glacial sites more influenced by Arctic water; Supplementary 3 Fig. S1), salinity was orthogonal to separation into glacial and non-glacial sites and was thus independent of their ordination in eukaryotic and prokaryotic RDA.

For each fjord, picoplankton beta diversity distance (Aitchinson distance) of the sites closest to the fjord mouth was compared to the respective hydrodynamic distance (defined as the normalized inverse of the \log_{10} of the synthetic particle concentrations) of connected site pairs. Hydrodynamic distance was positively correlated with pro- and eukaryotic Aitchinson distances across all temporal bins (Fig. C3.2b, c; Pearson correlation). Sites that were revisited over multiple years (Lofoten/Vesterålen (northern Norway) in 2014 and 2019; Kongsfjorden (Svalbard) in 2016 and 2017) were clear outliers relative to other sites with similar hydrodynamic distance, particularly in the 1-, 3-, and 6-month bins (Supplementary 3 Fig. S11).

Generally, we observed a slightly stronger correlation for prokaryotes with hydrodynamic distance than for eukaryotes (Fig. C3.2b, c). Eukaryotes exhibited high internal beta diversity variability (e.g., high Aitchison distance) at sites with a comparable degree of hydrodynamic distance (Fig. C3.2b). This variability also increased with increasing temporal bins. Among prokaryotes, the beta diversity distance between sites from Sweden/southern Norway to northern Norway were more pronounced than other sites with comparable hydrodynamic distances. Contrastingly, the beta diversity distance was comparably small between sites from Svalbard and east Greenland, as well as between sites from Svalbard and west Greenland.

Chapter 3 | 3.3 Microbial picoplankton diversity and co-occurrence networks in glacial and non-glacial-influenced regions

The interactions between prokaryotic and eukaryotic taxa were analyzed using a Sparse Inverse Covariance estimation for Ecological Association and Statistical Inference (SPIEC-EASI) of 108 intersecting taxa in the CLR-transformed ASV tables on the genus level at glacial- and non-glacial-influenced sites ($n = 34$ and $n = 50$, respectively). We found that the microbial subnetworks of glacial- and non-glacial-influenced sites included 37 associations in 70 components in non-glacial-influenced regions and 37 associations in 64 components in glacial-influenced regions (Fig. C3.3a, b; Supplementary 3 Table S2).

The subnetworks of glacial- and non-glacial-influenced sites were markedly different in both structure and taxonomic composition. Modularity, an indication of dense connection within groups and sparse connection between groups, was higher in the glacial-influenced network (0.83) than in the non-glacial-influenced network (0.74) (Supplementary 3 Table S2). Similarly, the glacier-influenced subnetwork consisted of multiple clusters with only two or three nodes, while the non-glacial-influenced subnetwork was predominately one cluster with few small component clusters (Fig. C3.3a, b). Taxonomically, the non-glacial-influenced network was dominated by prokaryotes (33 bacterial and archaeal nodes vs. 15 eukaryotic nodes) while the glacial-influenced subnetwork was relatively even (30 prokaryotic nodes vs. 29 eukaryotic nodes) (Fig. C3.3a, b).

Central nodes in the non-glacial-influenced network were exclusively prokaryotic, while central nodes in the glacial-influenced network were both prokaryotic and eukaryotic. Specifically, f__6(Arenicellaceae), Candidatus Actinomarina, Synechococcus CC9902, NS5 marine group, and f__11(SAR116 clade) appeared as central nodes in the non-glacial-influenced network; and f__1(Nitrospiraceae), f__92(Mamiellales), MAST-3_XX, Clade 1a (SAR11), and f__108(MAST-8) were central nodes in the glacial-influenced network. Node size corresponded to CLR-transformed relative abundance of ASV sequences. Notably, central nodes in the non-glacial-influenced regions were proportionally larger than in the glacial-influenced regions, as these taxa were also the most abundant in non-glacial fjords (Supplementary 3 Table S3).

The sum of each CLR-transformed ASV occurring in glacial and non-glacial fjords revealed that only a few prokaryotic ASVs in non-glacial fjords were strongly enriched, while many ASVs were relatively depleted (CLR-transformed count < 0 ; Fig. C3.3c). Most prokaryotic ASVs in glacial-influenced fjords were either enriched (CLR-transformed count > 0) and occurred in all fjords (Supplementary 3 Fig. S12) or were not detected (770 of 1433 total ASVs = 0). Eukaryotic ASV counts were relatively uniform between the glacial and non-glacial groupings (Fig. C3.3c), but individual glacial-influenced fjords had more unique ASVs than overlapping ASVs in contrast to non-glacial-influenced fjords, where most ASVs occurred in all fjords (Supplementary 3 Fig. S12). Microbial richness was similar for prokaryotes and eukaryotes in the non-glacial-influenced regions (Fig. C3.3d). While eukaryotic richness in glacial-influenced sites was not significantly different from eukaryotic richness in non-glacial-influenced sites, prokaryotic richness was significantly lower in glacial-influenced sites than in non-glacial-influenced sites (Welch Two Sample t-test, $p < 0.001$, $n_1 = 43$, $n_2 = 55$). Sample evenness of both prokaryotes and eukaryotes was significantly higher in the non-glacial-influenced sites (t-test, prokaryotes: $p < 0.001$, $n_1 = 43$, $n_2 = 55$; eukaryotes: $p < 0.001$, $n_1 = 47$, $n_2 = 58$).

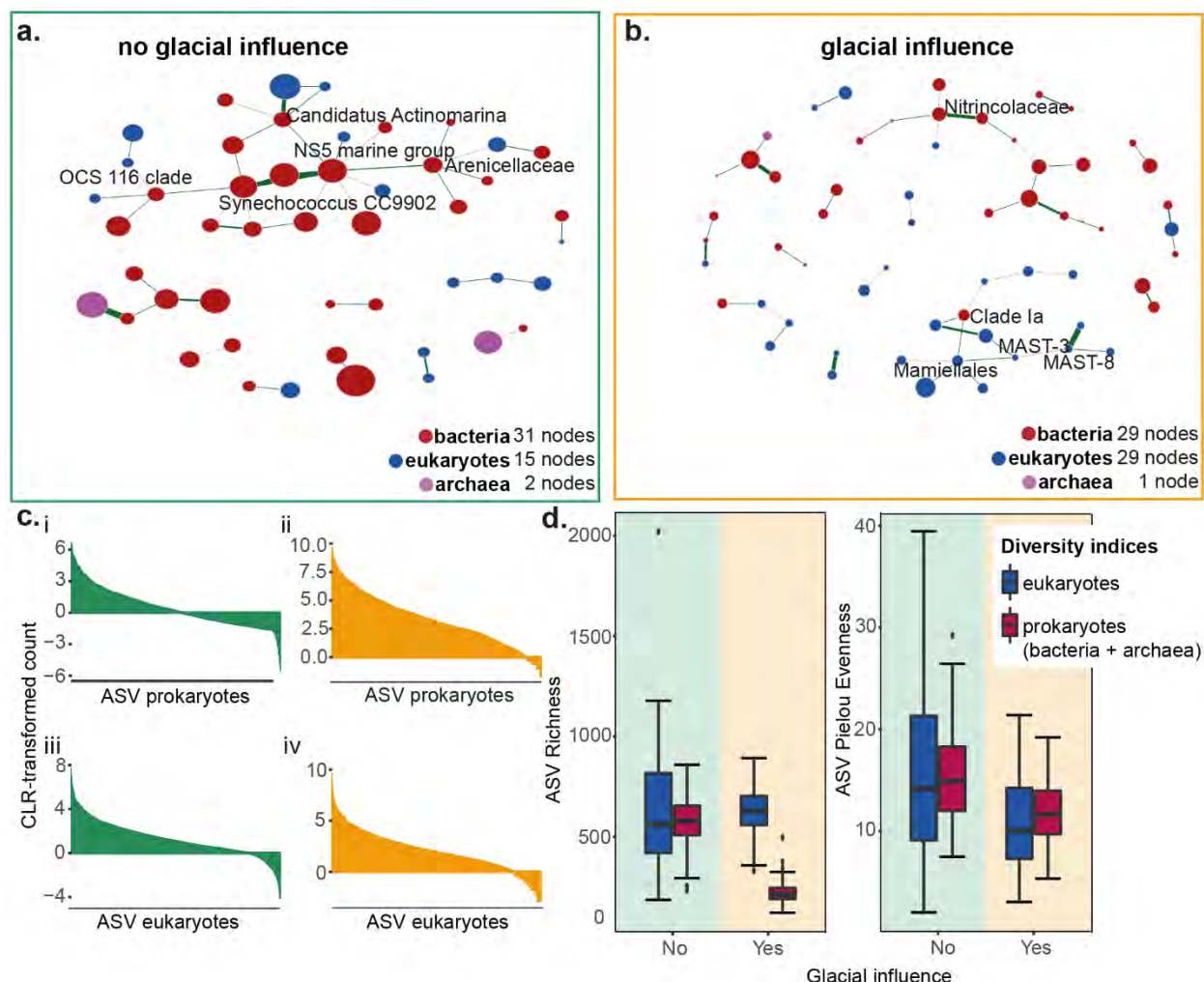


Figure. C3.3. co-occurrence network analysis of a. glacier and b. non-glacier microbial communities using spieeceasi association on CLR-transformed ASV tables and melted on species level for melted ASV tables of pro- and eukaryotes. C. CLR-transformed ASV sequence count of prokaryotes between sites without (i) and with (ii) marine-terminating glaciers and of eukaryotes between sites without (iii) and with (iv) marine-terminating glaciers. Values above 0 reflect relatively enriched ASV within samples and below 0 reflect relatively depleted within samples. ASVs that did not occur in the respective group (presence of marine terminating glaciers) were removed from the plots. d. Boxplots of picoplankton ASV richness and Pielou Evenness of Sites with and without marine-terminating glaciers.

Chapter 3 / 4 Discussion

The present study revealed a positive relationship between hydrodynamic distance and diverging microbial beta diversity signals. However, this relationship weakened between 3 and 6 months, particularly among prokaryotes (Sweden/southern Norway and northern Norway; drifter time 3–6 months) that were also distinct in their environmental conditions, such as nutrient profiles and sea surface temperature. Further, we demonstrated that the beta diversity of picoplankton communities separated statistically into fjords with and without marine-terminating glaciers. We documented a heterotrophic/parasitic-dominated co-occurrence network in glacier-influenced sites and a prokaryotic, decomposer-dominated network in glacier-free regions. Statistically, eukaryotes dominated the central nodes (i.e., cores of individual covariance clusters) in the glacial-influenced picoplankton co-occurrence network, while prokaryotes exclusively formed the central nodes in the glacier-free network. This suggests a potentially fundamental ecosystem baseline shift towards prokaryotes, particularly decomposers, under continuing glacial melt and Atlantification of the Arctic Ocean.

Oceanographic connectivity shapes microbial beta diversity change

Advection is a key determinant of microbial survival in the Arctic Ocean, but is yet inadequately quantified at scales relevant for biogeochemical cycles and species distribution (Wassmann et al. 2015). Our drifter analysis revealed strong regionality among shorter timeframes (< 6 months) and a nearly pan-Arctic distribution across multiple years, (but no recirculation of northern sites into southern fjords in southern Norway or Iceland, which requires time scales > 10 years). Microbial beta diversity change (Aitchinson distance) positively correlated with hydrodynamic distance (Fig. C3.2b, d), supporting previous observations of structuring importance of hydrodynamics for microbial diversity assemblies (e.g., subsurface chlorophyll maxima communities; Monier et al. 2015). Consistent with these findings, the Nordvestfjord Scoresby Sund (East Greenland), a region strongly affected by the cold East Greenland Current (low temperature/low salinity; Supplementary 3 Fig. S9), clearly ordinated from other sites along the second RDA axis in eukaryotic picoplankton.

Prokaryotes were more influenced by hydrodynamic distance than were eukaryotes (Fig. C3.2b, d), suggesting that prokaryotes were more restricted in their geographic dispersal, potentially through temperature selection (Logares et al. 2018). In support of strong temperature selection, we found lower prokaryotic enrichment in fjords with marine-terminating glaciers. Further, temperature selection could explain the strong microbial beta diversity distance between Sweden/southern Norway and northern Norway, despite their relatively strong oceanographic connectivity. However, prokaryotic dispersal could increase in the future due to increasing Atlantification of the Arctic Ocean, which occurs due to factors such as increased heatflux and climate change-induced glacial melt (Holmes et al. 2019)). These shifts could lead to an expansion of suitable habitat to thrive, as observed for the eukaryotic nanoflagellate *Emiliania huxleyi* over the past 24 years (Oziel et al. 2020) and the shift of diatom-dominated spring communities in the Fram Strait to the domination of *Phaeocystis pouchetii* (Nöthig et al. 2015) and their strong competition in the Arctic coastal region (Hegseth and Tverberg 2013).

Temporal effects

Although observations across temporal scales were not the primary objective of our analysis, we detected both inter-annual and seasonal effects in beta diversity signals, likely because the dataset spanned across multiple years (2012–2019) and ranged seasonally from the end of May to mid-August. For example, the Lofoten/Vesterålen (northern Norway) and Kongsfjorden (Svalbard) sites that were repeatedly sampled in our dataset exhibited strong temporal changes in the picoplankton diversity signals despite their regionality, suggesting inter-annual and/or seasonal variabilities. To exclude the effect caused by different times of sampling, campaigns such as the Ocean Sampling Day (Kopf et al. 2015) combined with temporal back-tracking of community transport with ocean hydrodynamics could help to establish a more unbiased view of oceanographic connectivity between sites than our microbial dataset offers.

Seasonality is a primary structuring temporal factor in Arctic marine ecosystems (Marquardt et al. 2016). We observed a significant seasonality effect (here approximated by sun altitude) on both prokaryotic and eukaryotic beta diversity (Fig. C3.2a, c). In summer, glacial-influenced regions can be shaped by strong stratification through differences in ice cover and freshwater influence from thawing glaciers and permafrost (Cottier et al. 2010). Similarly, seasonal shifts in microbial

relationships have been observed between a parasitic network during the winter and a detritus-based food web in summer along the Alaskan Beaufort Sea coast (Kellogg et al. 2019) and alterations within the food web such as episodic events of rapidly sinking ice algae (Wiedmann et al. 2020). At present, we cannot exclude a seasonal effect induced through the different geographic (latitudinal) locations on the observed pattern in our data, which could be related to diverging seasonal signals, such as different timing of ice melt and river discharge (McGovern et al. 2020). To gain insights into the impact of seasonality on (re-)occurrence of microbial communities, other seasons and samples throughout the year would need to be considered, such as in the Isfjorden Adventfjorden Time Series (IsA) <https://research.unis.no/isa/>, combined with an analysis of geographic connectivity.

Distinct picoplankton beta diversity signals of fjords with and without marine-terminating glaciers

Fjord systems form unique and productive habitats for marine life, and their distinct microbial biodiversity pattern in relation to the presence of marine-terminating glaciers have been recently highlighted in the Southern Hemisphere across the Patagonian fjords (Maturana-Martínez et al. 2021). Our observations confirmed that these distinct microbial beta diversity patterns are also present in the picoplankton size fraction across glacial- and non-glacial-influenced fjords in the Arctic (Fig. C3.2a, c). The beta diversity patterns between glacial- and non-glacial-influenced sites could be related to recently observed breakpoints in beta diversity patterns of temperate vs. polar algal microbiomes (Martin et al. 2021), potentially driven by distinct metabolic adaptations to changes in temperature and salinity (Elferink et al. 2020). To determine if similar mechanisms are affecting our sample sites, microbial activity measures, for example, metatranscriptomics, across all fjord microbiomes would be required.

Eukaryotic glacial-influenced beta diversity was positively associated with Si concentration, a key element for diatoms, which can potentially change with glacial melt and increased acidification, resulting in cascading effects on the microbial community (Cantoni et al. 2020). Other than Si, we could not determine any environmental parameters, for example, increased organic matter and nutrient input from glacier runoffs into the fjords (Müller et al. 2018; Kim et al. 2020b), which we could relate to our observations. Furthermore, we did not detect a significant increase of prokaryotic taxa involved in iron and sulfur cycling (*Sulfitobacter* spp., Thiotrichales,

Thiomicrospirales), which have previously been associated with glacial-influenced fjord systems (Van Keulenfjorden, Svalbard (Buongiorno et al. 2019); Kongsfjorden, Lilliehöökfjorden, Dicksonfjorden, Svalbard; (Laufer-Meiser et al. 2021)). This lack of iron and sulfur cycling taxa could potentially be due to seasonal variations in the magnitude of iron and sulfur flux and location (Herbert et al. 2021) and/or comparable input through enriched sediments in non-glacial-influenced fjords, such as in the Norwegian shelf (Hamilton-Taylor and Price 1983; Wehrmann et al. 2009). We did observe an outlier in the eukaryotic RDA, in which the station closest to the coast ordinated closer to stations from northern Norway and had much higher N and P concentrations than other samples from Disko Bay, potentially arising from glacier melt and associated nutrient input (Hansen et al. 2012). This suggests a more complex interplay of microbial diversity and environmental conditions such as nutrient input, changes in light availability, and stratification, which compromises the ability to conclude a general shift in ecosystem structure, elemental (C, S, Fe) cycling, and food web processes on a pan-Arctic scale (Hopwood et al. 2020). This also highlights the need for further studies that look at microbial indicators of turnover of key elemental cycles, at a high spatial resolution and over multiple fjord systems, in relation to the presence of marine-terminating glaciers.

Co-occurrence network analysis reveals different scales of spatial patchiness between glacial-influenced and glacier-free regions

Higher modularity, an indication for dense network clusters with loose interconnections of the glacial-influenced picoplankton network in comparison to the non-glacial-influenced network (Fig. C3.3a, b) allows us to infer higher environmental heterogeneity in the glacial-influenced network (*sensu* Röttgers and Faust 2018). This observation is further supported by a high number of relatively enriched ASVs in all glacial-influenced sites (Fig. C3.3c), despite significantly lower sample evenness (Fig. C3.3d), likely a result of multiple ASVs with different high abundance within individual fjords, as also indicated by the heterogeneity of the ten most abundant eukaryotic ASVs in glacial-influenced fjords (Supplementary 3 Table S3).

Increased network modularity increases the susceptibility to disruption by invasive species (Frost et al. 2019), such as the increasing presence of harmful algae (Bruhn et al. 2021; McKenzie et al. 2021) and ship-borne invasive microorganisms in the Arctic (Lawson 2019). However, depending

on where disruptions occur, modular communities can either return to equilibrium compositions (Hernandez et al. 2021), or can be affected by significant cascading effects in the food web (Kadoya et al. 2018). Therefore, to make conclusions from this general trend, this observation would need to be supported by an analysis of both temporal dynamics such as seasonal and inter-annual changes and geographic-independent signals. For example, a comparative study of co-occurrence patterns in Arctic and Patagonian fjords could provide more robust, geographic-independent insights into the structuring role of marine-terminating glaciers in the microbial community.

Prokaryotes dominate glacier-free co-occurrence networks

Glacial- and non-glacial-influenced co-occurrence networks were also taxonomically distinct, indicating potential implications for food web processes and fjord biogeochemistry. Prokaryotes (Arenicellaceae, NS5 marine group, Candidatus Actinomarina), previously associated with phytoplankton blooms and organic matter degradation (Krüger et al. 2019; Kopprio et al. 2021), formed the central nodes in the co-occurrence network of glacier-free fjords, and numerically dominated the taxonomic representation of connected nodes. If confirmed across temporal scales and supported by in situ rate measurements, this could suggest a baseline shift of the ecosystem towards higher prokaryotic activity and thus increased organic matter degradation within the microbial loop. By comparison, the central nodes in the glacial-influenced co-occurrence network were formed by a diverse assembly of prokaryotic degraders (Nitrospiraceae, SAR11 clade 1a), eukaryotic parasites (MAST), and eukaryotic phototrophs (Mamiellales). This is consistent with prior findings (Fadeev et al. 2018) that identified plankton associations in ice-covered versus ice-free ecosystems that matched general pre- and post-bloom microbial communities in the Fram strait.

Maranger et al. (2015) previously described a proportionally higher increase of bacterial production relative to viral lysis with higher temperatures, resulting in more efficient cycling of bacterial carbon within the microbial loop in the Arctic Ocean. Our observations support and extend these observations: We observed proportionally reduced centrality of eukaryotic parasitism towards a more detritus-dominated non-glacial-influenced and relatively warmer network. This could, however, as discussed above, potentially shift seasonally (Yang et al. 2020). Our analysis highlights the need for co-analysis of prokaryotic and eukaryotic microbial communities

considering a potential baseline shift of influencing processes within the microbial loop such as predation, parasitism, and organic matter degradation. This is particularly important in light of previously described shifts in phytoplankton from diatom- to dinoflagellate-dominated communities (Camarena-Gómez et al. 2018). Moreover, less predation could decrease carbon-cycling, also considering the importance of glacier runoff for carbon-export (Wiedmann et al. 2016), despite an observed climate-driven increase in productivity (Frey et al. 2018). These potential baseline shifts provide novel insights into previously described increasing resilience of a fjord food web (Griffith et al. 2019), which extends to factors such as sustaining important fish stocks (Meire et al. 2017).

Chapter 3 / 5 Conclusion

Arctic and sub-Arctic fjords are highly seasonally productive ecosystems, but it is unclear how they will be affected by ongoing glacial retreat and sea ice decline. Our co-occurrence network analysis, which included multiple fjord systems in the Atlantic sector of the Arctic, revealed that fjords with marine-terminating glaciers harbor prokaryotic degraders and autotrophic and heterotrophic-parasitic eukaryotes as central nodes of the network, while glacier-free fjords are primarily centered around prokaryotic detritus. This observation has also been previously demonstrated on a temporal scale (Kellogg et al. 2019). Such baseline changes within the food web can imply changes in carbon-cycling efficiency and resilience of existing ecosystem structures and across the entire food web. Specifically, oceanographic dispersal of marine microorganisms, a major structuring variable in our analysis, can contribute to accelerated change, which could also be affected by the expansion of cross-Arctic ship routes and subsequent introduction of alien species (Seebens et al. 2018; Frost et al. 2019). Our multivariate analysis further highlights that fjord ecosystems are unique in both their environmental and microbial profiles. Therefore, modeling approaches to differentiate between spatial and temporal (seasonal and long-term) components are urgently needed to make appropriate predictions regarding ecosystem state and change. These can be used to inform fisheries and identify the risk of biome collapse with ongoing

changes in active and passive transport of species and/or environmental changes such as sea ice decline, ocean warming, and extreme weather events. We demonstrate that a network-based approach of simultaneously evaluating prokaryotes and eukaryotes provides insight into the susceptibility of the base of fjord food webs. However, temporal observations in addition to spatial coverage are needed for effective monitoring and establishing sustainable solutions for these highly vulnerable ecosystems that are undergoing significant environmental changes.

Code Availability.

All code is archived and publicly available in github AGJohnAWI/ArcticPicos under the release v1.0 (<https://doi.org/10.5281/zenodo.5781579>)

Data availability.

Physical oceanographic data is publicly available for all expeditions: HE533, <https://doi.org/10.1594/PANGAEA.903511> (John and Wisotzki 2019) HE431, <https://doi.org/10.1594/PANGAEA.863438> (John and Rohardt 2016); HE492, <https://doi.org/10.1594/PANGAEA.881306> (John and Wisotzki 2017); MSM56, <https://doi.org/10.1594/PANGAEA.871015> (Friedrichs et al. 2017); MSM21-3, <https://doi.org/10.1594/PANGAEA.819731> (Zielinski et al. 2013).

Genomic sequences are publicly archived in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB49358 for of the HE492, PRJEB50059 for HE533, PRJEB50592 for HE431, PRJEB50596 for MSM56, and PRJEB50593 for MSM21-3 using the data brokerage service of the German Federation for Biological Data (GFBio, Diepenbroek et al. 2014), in compliance with the Minimal Information about any (X) Sequence (MIxS) standard (Yilmaz et al., 2011). Sequences data from the HE513, HE533, MSM26, MSM21-3 are submitted to GFBIO.

CHAPTER 4: Perspectives on Documenting Methods to Create Ocean Best Practices

Authors: Cora Hörstmann, Pier Luigi Buttigieg, Pauline Simpson, Jay Pearlman,
Any M Waite

Abstract

This perspective outlines how authors of ocean methods, guides, and standards can harmonize their work across the scientific community. We reflect on how documentation practices can be linked to modern information technologies to improve discoverability, interlinkages, and thus the evolution of distributed methods into common best practices within the ocean community. To show how our perspectives can be turned into action, we link them to guidance on using the IOC-UNESCO Ocean Best Practice System to support increased collaboration and reproducibility during and beyond the UN Decade of Ocean Sciences for Sustainable Development.

Chapter 4 / 1 Introduction

The ocean community is facing ever-increasing methodological complexity in its efforts to meet global challenges and to further discovery. Fragmentation of methods and data across regions, nations, and disciplines inhibits effective collaboration and ocean observation (Brett et al. 2020). Fortunately, a new culture of open sharing of knowledge on the web is merging along with the FAIR principles (Wilkinson et al. 2016), leading to new forms of dialog in the science community. While their scope of application is broad, the FAIR principles are strongly focused on data, but can equally apply to methods, standards, and protocols. Indeed, if data is to be reproducible, reusable, and interoperable, then methods themselves have to be FAIR. Simultaneously, the community is recognizing that there is a reproducibility crisis in science (Lithgow et al. 2017) and marine science is no exception to this rule (Lowndes et al. 2017). However, multiple initiatives are promoting new, structured, and persistent ways to share methods online (Teytelman 2018).

The path to digitizing methods and making them FAIR is not trivial, but concrete steps towards more structured and machine-readable documentation are already being taken. Developing the skill to write high-quality methods can be especially helpful for those early in their career or new to a field (Bell 2014). Further, method documentation itself should be harmonized to make it transparent and easily understandable across multiple practitioners; and lastly, methods should be made available in online, accessible, and machine-readable ways. Guidelines, templates, and (e)protocols exist in other fields such as medical research (Weissgerber et al. 2016; Aerts 2018), accompanied with emerging possibilities in knowledge sharing and training (Nurhas et al. 2018; Nti et al. 2020).

To help address this for the ocean community, Hörstmann et al. (2020) recently published an Intergovernmental Oceanographic Commission (IOC) Manual and Guide on enhancing method documentation for increased FAIRness. The guidance in Hörstmann et al. (2020) is acted upon by the deployment of technologies underpinning the IOC Ocean Best Practices System (OBPS). The OBPS was developed to host and interlink methodological documents of any kind (protocols, guidelines, standard specifications, etc.); and seeks to support continuous convergence of methods

as they undergo community refinement to become “best practices.” In this way, the OBPS has laid a foundation for global, cross-community federation of method archiving and system development.

However, the effective use of any system of this kind strongly depends on how it is used. That is, the way in which we – as an interdisciplinary ocean community – write and structure our methodological documents will have a direct effect on the effectiveness of the design, and technical capacities, of archiving and dissemination systems. It is our responsibility to document our methods in a reproducible and transparent way, to efficiently use available technologies. From this perspective, we discuss the value and meaning of a “quality” method document, and how we use better documentation to share our methods within the community. Further, we reflect on the current state and future directions of handling diverse methodological content for a global community, increasingly making use of advanced digital resources in an inclusive and ethical way.

Chapter 4 / 2 What does it mean for a method to become a best practice (and why should we care)?

A method can be considered a best practice when it consistently produces superior results over other methods with the same objective (Simpson et al. 2018; Pearlman et al. 2019). Of course, context is key; what is best given a particular level of resourcing or environment of application (among many other factors), may not be so when circumstances change. Further, having a potential best practice in itself does not guarantee utility to the wider ocean community – only when shared, cross-validated, and used in the creation of better practices beyond its original scope, can we truly become excited about high-quality methods. On the aspect of context-dependence, the perspective on what makes a method “best” can rapidly change across communities, disciplines, and resource levels. This is especially true, as most communities see their methods as the “best” for their own needs. Creating “best practices” across these scales therefore needs an inclusive, transparent, and sustained set of processes to communicate why, when, and for whom a method is a best practice.

What practical steps can we take to systematically identify localized best practices and explore their value in the broader global community? At the very least, methods must be sustainably archived, discoverable and accessible online to play a role in contemporary systems. In addition, technologies that interlink methods, and the communities that champion them, are needed to cope with the scale of the challenge (Buttigieg et al. 2019). Once the link is made, communities have the opportunity to test and advance methods in order to determine under what circumstances they are “best”. The persistence, visibility, and version control described above support widespread testing and cross-validation, thus allowing multiple practitioners to determine where, when, and how a method can be a best practice. To accomplish this, those that are developing methods and those that are archiving and interlinking them with new technologies must enter a more intense dialog. Through this, systems serving methods to the ocean community can draw closer to their end-users. This exchange encourages method evolution including adaptation for new technologies.

Chapter 4 / 3 How can we support the ocean community in creating best practices?

To bridge key communities, the first, feasible, and particularly impactful step in this process is simply the use of more structured methodological documentation to synchronize novel technologies with users’ needs, allowing more rapid and transparent improvements of ocean practices. We contend that structured templates, clear and complete metadata, version control, as well as mechanisms to support convergence and interdisciplinary exchange are foremost among the community’s needs. Templated documentation promotes well-structured, reproducible, and, in some cases, machine readable best practices. However, to effectively support the user experience, templates – and the sometimes required additional workload which comes with populating the templates – need to be introduced in a systematic way (Alwazae et al. 2014). Developers can be supported through additional guidance and explanations of the rationale behind the conventions suggested. In accordance with this principle, the ocean community called for guidance on how to effectively use the OBPS templates and what would be a “best practice for best practices” (BP4BP; Hörstmann et al., 2020) in method documentation (Simpson et al. 2020). With this, we hope that

submitters are supported in archiving and sharing their methods to enable community refinement and global harmonization more successfully.

A key component of method documentation is the provision of relevant metadata. Pearlman et al. (2019) to access information about method maturity, relevant societal values/goals, and accompanying datasets, while the BP4BP document provides guidance on how users and developers can leverage them. This guidance reinforces the completeness of a document's metadata, which promotes machine access and further interaction with the system's technologies, and, consequently, document discovery and reuse by the community. With organized and trusted online method documentation in place, systems are able to couple secure archiving and traceable versioning. This ensures that community refinement, extensions, and hard- or software changes are traced through the evolution of a method. In addition to being self-contained and complete, methodological documents created using the same templates can be more rapidly – and in some cases automatically – compared. Rapid intercomparison can greatly enhance our community's capacity to rigorously assess a large body of methods and choose those that are most appropriate to their mission and capacities. As improved guidance on methodological development emerges, we are increasingly able to globally coordinate ocean observation and action, which have not yet converged. However, the technologies in methodology management systems (like the OBPS) rely on detailed, well-structured documentation to provide researchers with useful interfaces and functions. We note that developing such documentation is a continuous process and thus welcome the communities' critique of the BP4BP. Harmonized documentation supports interdisciplinary understanding of methods and data, and offers the possibility for better coordination in ocean observation. Methods then progress from stand-alone elements into more globally standardized pieces of the puzzle for improving global understanding of ocean systems.

Chapter 4 / 4 Why are ethics and inclusiveness central to the documentation of methods and best practices?

The ocean is in crisis yet offers humanity immense resources for sustainable development; in this complex reality, the publicly funded scientific community has an ethical obligation to rapidly share and collaboratively evolve the best of its methods to secure a truly sustainable and healthy relationship with the planet's oceans. This is especially true in the face of a scientific reproducibility crisis (Lithgow et al. 2017), data leaks (Gibney and Van Noorden 2013), and a publication bias towards positive results, resulting in wasted time and resources (Rothstein et al. 2005). Further, consideration of how new methods directly or indirectly impact communities – and including those communities in the development of such methods in a co-design process, – is often overlooked, despite being essential to build and maintain trust at the science-society interface (Achterberg et al. 2015). A prerequisite for this trust is the (re)use, recognition, and official/institutional crediting of open methods as valued scientific outputs to support those making necessary changes in our culture of marine science. However, the ethics of sharing know-how grow in complexity as more stakeholders – outside of science are engaged to investigate ever broader and more multifaceted ocean and societal phenomena.

Across marine science, there is growing multilateral motivation to contribute scientific insight and methodology to societal missions through local, national, regional, and global programs. Simultaneously, science is integrating more methods from other sectors to interface with policy, the private sector, local communities, and other stakeholder groups. Consequently, inter-sectoral methodological harmonization and exchange has become a priority. To this end, scientific missions will need to include multi-sectoral perspectives and interests in solution-oriented research, and not only on the level of post-mission data transfer (Weichselgartner and Kasperson 2010). Including such partners early, during method development itself, and working with them to balance the interests in and impacts of marine science will be a key strategy to promote regional sensitivity and trusted engagement across societal groups. Even if such groups cannot be immediately engaged, Hörstmann et al. (2020) describe how a document's content and prose may be selected, structured, and written to allow authors to engage with global stakeholders, which will be especially important

to furthering the goals of the UN Decade of Ocean Science for Sustainable Development (henceforth, the “Ocean Decade”; Ryabinin et al. 2019).

An inclusive confluence of stakeholders and interests is a complex space; while it brings much-needed diversity to science, it also brings together different ethical norms that may not always align. In methodological development, this is often expressed through debates on what can be considered a “best practice” or even an appropriate way to conduct research, where, when, and by whom (Barbier et al. 2018). Discussion of these ethical challenges is broader than this article’s scope, but we can take a concrete step towards improvement by proactively accounting for the ethical dimensions of any scientific method in its documentation. This is particularly important when a method impacts those who typically lack a voice in the scientific enterprise, to prevent the scientific discourse on standards, methods, and best practices being dominated by the larger or well-resourced global campaigns and actors, and high-level decision-makers.

In practical terms, we must co-develop a more consistent and transparent way to document the ethical dimensions of our methodologies. First, this will necessitate that ethical codes and guides (e.g., EU ethic appraisal procedure) relevant to the marine sciences are digitized and made available online, following the same FAIR-aligned practices that systems like the OBPS use for other documents (e.g., DOIs, semantic indexing, searchability). Second, the sections of those methods where ethical concerns arise should consistently reference and/or link to such digitized ethical guidance enabling both humans and machines to readily understand when and where it is applicable. While the above will provide a path to build new networks of trust, ethical inclusion extends well beyond the inclusion of ethics in scientific practices. Rather, it sets a precedent whereby the co-development of methodology in marine science includes dedicated sections of methodological documents co-authored by other stakeholders (e.g., policy makers, citizen scientists, private sector interests, and conservation groups). This is a natural act of reciprocity to the same stakeholders who include scientific content and insights in their working documents, and will help address the persistent distrust in science across multiple sectors (Pechar et al. 2018). Further, such co-authored sections will be more broadly understandable to all groups involved, and connect them to resources and initiatives, which are of immediate relevance to their collective decision-making processes. The links and conventions which emerge from such inclusive

interaction are a precursor to the application of technologies (referenced above) that will accelerate innovation, responsiveness, and cross-boundary collaboration in best practice development for the marine sciences.

Chapter 4 / 5 Concluding remarks and the path ahead

The core of this perspective is that improved communication of methods, within and beyond marine science, is essential to interdisciplinary, community-led efforts to create a global collection of best practices. This evolves our responsibility to document our methodologies to support reproducibility into an opportunity to create an inclusive, global community of practitioners testing, adopting, and refining scientific methods in more diverse contexts.

Of course, the scale of this challenge needs a range of technologies to help our community coherently develop their methodologies. This perspective explored key aspects of increasing human-human and human-machine interaction, cross-linking documents, guidance, and information to support sustainable ways for truly inclusive and ethical method documentation. To move forward, our community will be guided by a number of open questions that we collaboratively need to address:

- *What are the major steps we need to take to document our methods for more universal understanding?* We believe that structured templates and guidance documents, such as the BP4BP, are important steps in this direction, but the marine science community will need broader and continuous discussion of this theme to scope our needs and track advancements.
- *Is our marine science community ready to accept more prescribed ways of documenting their methods?* The diversity of needs within the community is high, with many overlaps and redundancies. Addressing this diversity should not compete with more standardized ways of documenting methods. Rather, the freedom of scientific exploration should be

expanded through the improved communication structured documentation can bring, especially when combined with technologies able to enhance transparency and inclusivity. Balancing standards and innovation will be a persistent challenge in this space.

- *How can we connect communities with similar methods and promote?* Interoperable documentation is, again, a first step, but harmonization is a much more faceted, community-led process. We need a more systematic understanding of what motivates and incentivizes marine scientists to harmonize their methods, so that systems like the OBPS can more effectively accelerate it.
- *How can we federate best practice development in marine science?* We believe that systems that allow projects, programs, and other short-term initiatives to preserve, organize, and track the re-use of their methodological know-how can be federated to create a sustainable, global pillar of marine science. Interoperably cross-linking the holdings of regional and global platforms will unite distributed and independent systems already moving towards open, version-controlled, and transparent method sharing (e.g., Marine Sampling Field Manuals and the OBPS). Once again, how these systems – and the community of practitioners contributing to them – can be incentivized to build a common vision for a truly global solution is still to be discovered.

We invite the ocean community to join us in addressing these questions to continuously and collaboratively improve our collective methodological capacity. Through this, we will be able to unite and observe the ocean on larger scales than ever before. Together, we can evolve marine science into a planetary-scale and multi-stakeholder enterprise contributing to societal goals – particularly the pursuit of a healthy marine ecosystem – with unparalleled coordination.

4. Synthesis

The extent, connectedness, and presence of hydrographic barriers is crucial to map marine ecosystems. The patterns of microbial beta diversity can reflect such delineations of marine ecosystems. Furthermore, measurements of microbial functional traits and activity can give rise to the mechanisms of ecosystem function within and across these defined regions as marine microbes mediate key biological processes of global carbon (C) and nitrogen (N) cycling, such as C and N₂ fixation, that can be differentially altered by the traits and trade-offs within a microbial community at one site. In this thesis, I explored microbial alpha and beta diversity patterns in conjunction with microbially-mediated processes (primary productivity and N₂ fixation) through an oceanographic lens. My results revealed that phylogenetic and functional diversity is statistically different in regard to (a) external physical and chemical processes, and (b) internal community structure of trophic functional groups. A key structuring feature for microbial beta diversity was the horizontal transport of microbial communities leading to both oceanographic connectivity between sites, but also to a displacement and mixing of microbial communities from different sites of origin. In the following I synthesize the key outcomes of my thesis, providing new insights into spatial scales of microbial diversity and their activity in the context of physical and chemical processes.

4.1 *Alpha diversity, temperature and Primary Productivity*

4.1.1 Multiple facets of temperature correlating with microbial alpha diversity

Temperature has been shown to be a main predictor for biodiversity on land (e.g., Zhou et al. 2016) and in the ocean (Faure et al. 2021). However, temperature can take on different roles such as thermoregulation (ambient energy hypothesis; Turner 2004) as well as a water mass indicator in the ocean. Temperature can also correlate with sunlight radiation across latitudinal scales and other water-mass characteristics such as oxygen concentration and salinity (that also appear as a significant correlate with microbial alpha diversity; Zorz et al. 2019). In **Chapter 1**, the correlation was positive between both prokaryotic and eukaryotic diversity measures and temperature (Supplementary 1 Fig. S4) but the correlation structures were different in their variation and intensity, suggesting different relative importance of temperature (and/or associated other variables). This also suggests that the observed correlation of temperature with microbial diversity can sometimes result from hydrodynamics, an indicator for a water mass, and sometimes due to other environmental characteristics, and/or a combination of both.

The study region in **Chapter 1** was marked by strong temperature gradients (3–25°C), and associated steep nutrient gradients between the oligotrophic South Indian Ocean Gyre and the high-nutrient Southern Ocean. Our data also exhibited strong variability of temperature–biodiversity correlations between pro- and eukaryotes in these regions. This could be due to the destabilizing effect of nutrient enrichment on temperature–biodiversity relationships as previously shown in mesocosm experiments (Wang et al. 2016). Moreover, the environmental variables likely act to different degrees on microbial traits (e.g. potential trade-offs among traits and direct vs. indirect trait selection; reviewed in Litchman et al. 2021) of functional groups that blur existing correlations of individual groups with temperature and/or nutrients.

Indeed, Shannon diversity of functional groups (auto-, mixo-, and heterotroph) and temperature were markedly different. For example, analyses showed that the relationship with temperature varied between trophic groups; encompassing negative, positive, or no relationships with Shannon

diversity. Moreover, the correlation structure of temperature with microbial diversity varied between different provinces and across the entire transect in the Atlantic Ocean (~50°S–50°N). This potentially implies that the correlation of microbial diversity with temperature is a matter of scale: regionally, the temperature component of water masses could reflect environmental profiles. Globally, thermal kinetics could be linked to latitudinal gradients in microbial diversity due to the microbes' thermal tolerance spectra (Ibarbalz et al. 2019). Disentangling these different mechanisms and scales of temperature-related biodiversity change can also ultimately help to understand the potential impact of climate change-induced ocean warming on the adaptation, interaction, and resilience of microorganisms (Cavicchioli et al. 2019).

Key outcomes

- I observed a positive relationship between temperature and microbial diversity across large-scale latitudinal transect in the Atlantic Ocean, confirming **H1.1** of this thesis.
- Correlation structure varied regionally between ocean provinces and across different trophic groups.

The variability at regional scales and across functional groups highlights that the different roles of temperature on microbial community structure need to be taken into account when investigating microbial alpha diversity patterns across temperature gradients (**O1**).

4.1.2 Trait-based approaches are needed to understand the link between microbial alpha diversity and primary productivity

Primary production is the fundamental food source for all life in the ocean and depends on both the quantity and quality of the primary production. Theoretically, the more diverse primary producers can co-exist and produce, the more diverse a consumer community can be (Fuhrman et al. 2008). In the ocean, the relationship between phytoplankton and photoautotrophic primary productivity is hump-shaped with a positive correlation when grazers control phytoplankton populations through selective grazing and negative when blooming of individual species occurs (Vallina et al. 2014). I found a negative PP correlation with cyanobacteria in the CNRY province (Fig. C2.2), likely due to the blooming of opportunistic species and thus reflecting the negative correlations as described

in Vallina et al. (2014). The large spread and sometimes even negative correlation between, e.g., autotrophic eukaryotes and primary productivity, and cyanobacteria and primary productivity (Supplementary 2 Fig. S6) suggests different strategies of adaptation and niche occupation. Diversity measures will probably be more precise with refined trait-based approaches (as applied in **Chapter 2**), and functional gene analyses (e.g., RubisCO gene involved in primary productivity), as different kinds of microbial functional groups have distinct effects on biogeochemical cycles such as carbon fixation and export (reviewed in Litchman et al. 2015).

Most correlation analyses between primary productivity and microbial diversity were not significant (**Chapter 1, 2**) likely due to strong regional habitat filtering and the effect of co-correlations with other variables such as nutrients (Cardinale et al. 2009). Additionally, as reviewed in Gamfeldt and Hillebrand (2008), dispersal is a key determinant for the primary productivity-biodiversity relationship through resource efficiency at different dispersal rates. The effect of horizontal dispersal on the spatial scales of ecosystem patchiness, as shown in **Chapter 2** of this thesis, could potentially lead to different spatial scales of primary productivity-biodiversity relationships than the province-scale as observed in the Pacific Ocean (Raes et al. 2018). To conclude confidently on environmental mechanisms, higher spatial resolution sampling at scales that correspond to horizontal dispersal of phytoplankton communities in relation to their internal biomass turnover rates (0–10km; as calculated in **Chapter 2**) is needed to identify mechanisms of these alpha diversity patterns.

Key outcomes

- Microbial alpha diversity was mostly independent of primary productivity except in a few cases: a decrease in alpha diversity at high primary productivity measures, partially confirming **H1.2**.
- Trophic functional groups exhibited distinct correlation structures (magnitude and significance) with primary productivity, highlighting the need for trait-based categorization when studying primary productivity-biodiversity relationships (**O1**).

4.2 Mapping Microbes

4.2.1 Microbial beta diversity reflect Longhurst provinces

Longhurst provinces (Longhurst 2007) reflect well microbial beta diversity patterns in the Indian Ocean (**Chapter 1**), Atlantic Ocean (**Chapter 2**), and Pacific Ocean (Raes et al. 2018). The physical oceanographic profiles likely contribute to microbial diversity patterns as ocean currents passively disperse microorganisms. For example, microbial diversity signals can reflect the microbial community from their water masses of origin (e.g., Bolaños et al. 2020), which result in ecosystem patches, as also observed through remote plankton observations (Mackas et al. 1985; D'Ovidio et al. 2010). Similarly, in **Chapter 2**, samples in those provinces that had a relatively high variation in their temperature-salinity profile (e.g., Brazil Current Coastal province) in comparison to other provinces also had a higher scatter of the alpha diversity signals. This suggests a physically-driven disruption of the biological habitat.

However, this does not mean that fronts and currents that form the boundaries between two provinces show an additive and mixed signal of the two source water masses (as suggested by Longhurst (2007)): the South Atlantic Front and Polar Front had a biodiversity signal distinct from the South Indian Ocean and Southern Ocean (**Chapter 1**; Baltar and Arístegui, 2017). This can also be due to the fact of different water mass origin (i.e., the strong and fast east-ward flow of the South Atlantic Front and Polar Front in comparison to the much slower current speed in the South Indian Ocean and the Southern Ocean; Supplementary 1 Fig. S2), but also due to the combination of high temperature from the South Indian Ocean with high nutrients from the Southern Ocean creating a unique physico-chemical environment. For example, this could allow flourishing of microbial recycling (reflected by increased ammonium concentrations in **Chapter 1**) and potentially increased rates of nitrification (Sambrotto and Mace 2000). The relationship between increased nutrient recycling and high temperature- high nutrient conditions was also recently shown in a freshwater mesocosm experiment by Ren et al. (2021).

Key outcomes

- Microbial beta diversity patterns broadly followed Longhurst provinces, confirming **H2.1**, but were more scattered in physically energetic provinces, suggesting that depending on physical structure (and sampling scale), the patterns of ocean provinces can collapse in regional patchiness.

Including the mechanisms of environmental sorting and ecosystem patchiness in relation to different water masses of origin can help better grasp the structure and function of microbiomes (**O2**).

4.2.2 Microbial phylogenetic diversity and activity

Microbial phylogenetic diversity and activity can decouple across strong physical gradients

I found that primary productivity and N₂ fixation changed more gradually along the transect between the South Indian Ocean and the Southern Ocean, largely following gradients of N, P, and Si concentrations, while microbial diversity grouped according to different water masses (**Chapter 1**). This suggests that microbial phylogenetic diversity is more physically driven, but their activity (and potentially to an unknown degree also their final phylogenetic community composition) results from internal effects through strong habitat filtering (Fig. 4.1). Furthermore, this also suggests – but would be premature to conclude without further functional gene analysis – the presence of functional redundancy.

Through a trait-based lens (**Chapter 2**), I observed clear signals that support the idea of partial redundancy within the microbial community that overlap in some functions but differ in other ecological traits (*sensu* Galand et al. 2018). Specifically, prokaryotic beta diversity was strongly impacted by biological N₂ fixation, which is itself a prokaryotic process (Karl et al. 2002). Such trait-based approaches are critical to understanding the microbial functional potential, as N₂ fixation could become more relevant especially in high latitudes, and with increasing global stratification of the water column (Hood et al. 2004; Breitbarth et al. 2007) and associated nutrient

limitations (Boyd et al. 2010; Garcia et al. 2011). The extent and impact of partial functional redundancy through functional gene analyses (e.g. *nifH* gene for N₂ fixation; RubisCO gene for primary productivity) and different functional traits and trade-offs need to be further resolved concerning C and N₂ fixation and recycling efficiency (e.g. cell size as a master trait for light absorption, nutrient uptake, sinking rate; Litchman et al. 2015). Additionally, traits and trade-offs regarding species survival in relation to environmental changes such as increasing sea surface temperatures are needed to understand the physical and biological (physiological) boundaries that may induce whole-ecosystem functional changes in the future.

Activity strategies can overcome dispersal barriers

Environmental selection acts differently on prokaryotes and single-cell eukaryotes due to their life strategies, such as the ability of some eukaryotic cells to form cysts (reviewed in Falkowski et al. 2004). In **Chapter 3**, I found that prokaryotes were more restricted by oceanographic connectivity, suggesting that prokaryotes are more impacted by environmental filtering such as temperature selection (Logares et al. 2018). However, prokaryotes dominated the co-occurrence networks in fjords without marine-terminating glaciers. Supposedly, low temperature is currently the limiting factor for prokaryotic dispersal, then prokaryotes and prokaryotic production might more and more dominate the future Arctic Ocean, as also modeled by Kim et al. (2020a) and suggested by Maranger et al. (2015). Further investigations of traits and trade-offs, as well as functional genes linked to physiological adaptation to abiotic conditions such as temperature resistance (e.g., presence of glycosyltransferases and glycosylsynthetases; Verde et al., 2016) would help to understand the mechanisms driving the establishment and dispersal limitations of individual taxonomic and functional groups.

Different trophic groups

Categorizing prokaryotes and eukaryotes in trophic functional groups revealed that trophic groups were more distinct in their beta diversity signal in relation to high and low chl *a* conditions (**Chapter 2**). This relationship reflects which trophic group can potentially better diversify under certain environmental conditions. For example, beta diversity of mixotrophs was markedly higher in low chl *a* provinces, supporting the argument by Hartmann et al. (2012) that a mixotrophic lifestyle is favorable under low nutrient conditions (“you take what you get”).

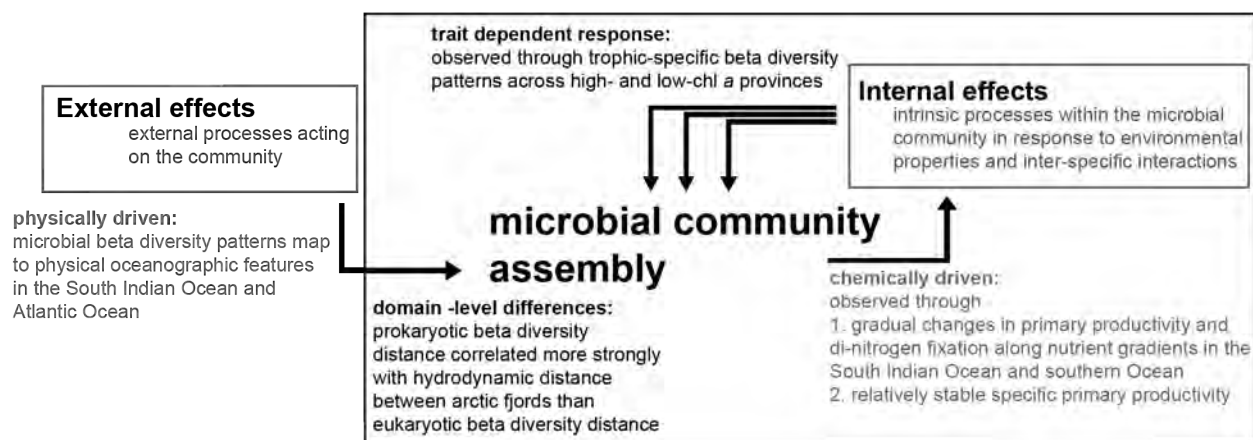


Figure 4.1 Observed microbial beta diversity patterns nested in theoretical concepts of structuring microbial community assembly mechanisms, including external and internal effects. External effects are more physically driven, with domain-level differences of beta diversity signals. Internal effects appeared to be more chemically driven with trait-dependent differences of beta diversity patterns.

Key outcome

- Domain-level physiological differences between prokaryotes and eukaryotes, and functional traits (here, differentiation between trophic groups) can bridge the link between microbial phylogeny and their activity (**O2.1**), but the structuring mechanisms of phylogenetic diversity can be distinct from functional patterns and rate measurements.
- Generally, across my data, microbial phylogeny mapped according to physical, oceanographic features, while microbial activity appeared to be more chemically driven (**O2**).

4.2.3 Beta diversity patterns allow refined mapping of microbial ocean provinces

Although microbial beta diversity largely mapped to Longhurst province, I observed regional divergence of microbial biodiversity patterns (**Chapter 2**). This divergence was also shown in the physical and chemical profile but not in phytoplankton pigment-based observations (Bracher et al. 2020). This shows how microbial observations can provide a more nuanced (biological) picture of


Longhurst provinces and – if confirmed across temporal scales – can help to refine boundaries for better ocean ecosystem management and observation.

Refining microbial sampling is also becoming more important as physical oceanographic observations will become more spatially resolved (e.g., through gliders; Testor et al. 2019; Whitt et al. 2020). In a recent publication, we showed that small-scale processes at the kilometer scale, specifically in dynamically active regions, can considerably change productivity, and thus global C cycling considering the ubiquity of these smaller-scale processes (von Appen et al. 2020). Additionally, variations of ecosystem dynamics across temporal scales are needed to monitor ecosystem states with anthropogenic pressures. For instance, Sarmiento et al. (2004) defined biomes focussing on up-and downwelling processes impacting primary productivity increase under climate change and noted large regional variations that contribute to high uncertainty of the predictions (between 0.7 and 8.1%). Understanding microbial responses and associated changes in their activity (i.e., primary productivity) can help elucidate and refine regional scales of ecosystem variability and ultimately integrate these in global understanding.

Key outcomes

- Microbial beta diversity patterns were more refined than previous province delineations based on phytoplankton pigment observations.

The results of this thesis highlight that microbial beta diversity patterns can be used as indicators for spatial ecosystem boundaries (**O2.2**). However, samples along temporal scales are needed to confirm the persistence of these features.



4.3 Different levels of spatial scales: from a microbial community's perspective to regional and global scales

4.3.1 Scaling the impact of ocean dynamics on microbial biodiversity through bio-physical coupling

In my work, I identified multiple spatial scales of microbial beta diversity patterns within open ocean systems (**Chapter 1** and **2**) and coastal fjords (**Chapter 3**) (Fig. 4.2). These systems could be subdivided into smaller spatial levels, namely ocean provinces, geographic regions, individual currents, and Arctic and sub-Arctic fjords. Additionally, I identified structuring mechanisms: microbial beta diversity separated into groups of ocean provinces with high and low chl *a* concentrations in the Atlantic Ocean (**Chapter 2**), and microbial beta diversity sites formed clusters in relation to the presence or absence of marine-terminating glaciers (**Chapter 3**) (colored circles Fig. 4.2). Furthermore, oceanographic connectivity between sites was a major structuring variable in my analysis and connected all spatial dimensions through the horizontal transport of microbial communities. Of course, microbial beta diversity can be further influenced by many dimensions not included in my analysis, such as temporal scales of dynamic community succession processes (Zhou et al. 2014) or vertical processes such as up- and downwelling (Gregoracci et al. 2015). In that context, I am currently investigating the effect of microbial beta diversity on carbon export through vertical profiles compared to underwater vision profiles (UVP; Rogge et al. *in prep*). The scales identified in this thesis help to map microbial biodiversity and support a more dynamic understanding of microbial biodiversity patterns by including physical oceanography and hydrodynamic horizontal scales.

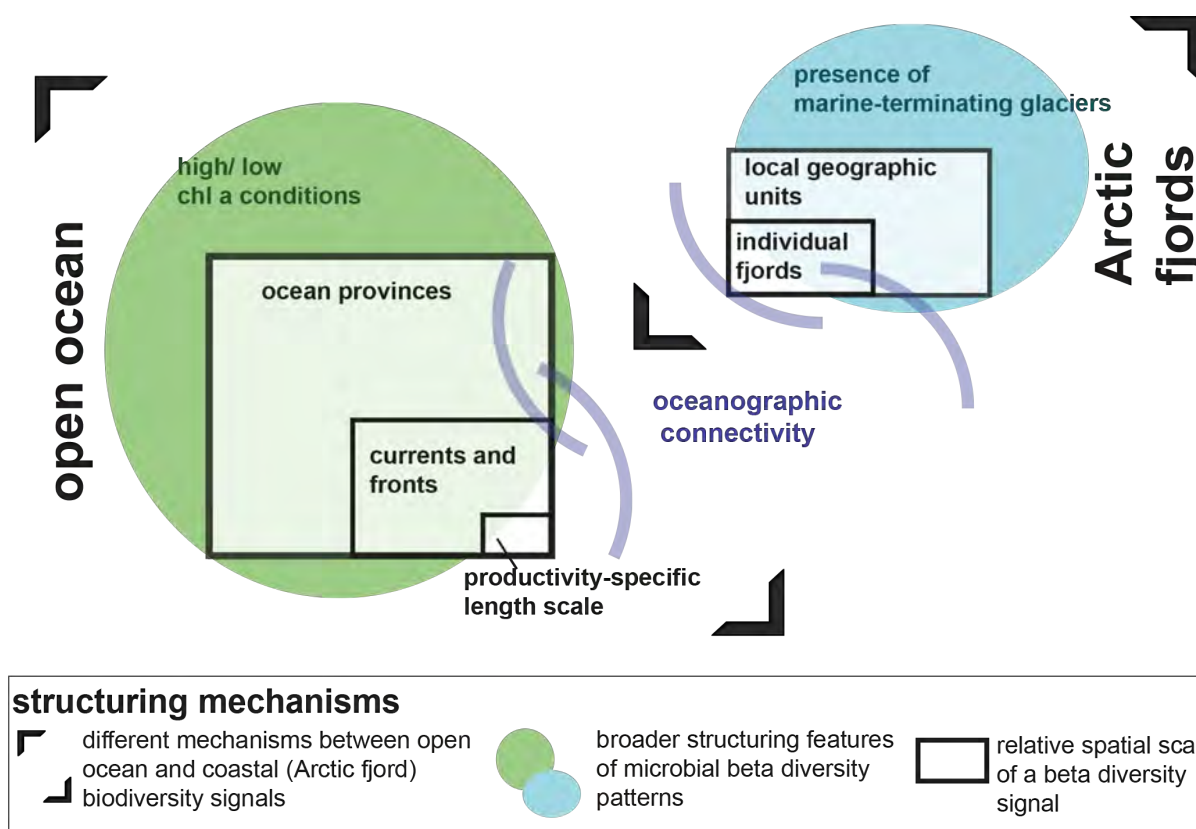


Figure 4.2 Nested spatial microbial beta diversity patterns that correspond to oceanographic/geographic features investigated in this thesis. Spatial level can be considered as individual tiles of microbial beta diversity pattern that form small unique patterns that can be nested within larger-scale patterns. For example, one ocean province can consist of multiple unique microbial diversity patterns within different ocean currents but that are distinct from other ocean provinces. Yet, notably, all these oceanographic features are to a certain degree interconnected via oceanographic connectivity, i.e. transport of microbial communities through surface ocean currents. The individual scales and spatial patterns are set in comparison to total dataset scales (open ocean/Arctic fjords).

4.3.2 A microbial community's perspective reveals environmental patchiness

Particularly, small-scale (0.1 m–100 km; see **box 3** for different oceanographic features) oceanographic heterogeneity can lead to patchiness that significantly relates to regional productivity and elemental cycles (von Appen et al. 2020). As highlighted by Powell et al. (1975), a crucial unknown is the smallest scale of microbial beta diversity change. In their study, Powell and colleagues linked spectral analysis of chl *a* to current speed and found that around 100 m length scale biological variance significantly contributes to observed patterns in chl *a* concentrations

rather than physically-driven distribution patterns alone. In **Chapter 2**, I presented the calculations of a productivity-specific length scale that includes a temporal component of the microbial diversity change in addition to the spatial scale presented by Powell et al. (1975). This scale is a first-order estimate of biodiversity change considering external and internal processes of microbial beta diversity patterns. To refine this concept, one would need to consider the different traits of cell growth, division and the ability to form dormant stages until suitable environmental conditions, which is a key survival strategy linked to global distribution of microorganisms with the “microbial conveyor belt” (Mestre and Höfer 2021). Further, the calculation of a productivity-specific length scale complements the concept of microbial residence time, which represents evolutionary processes of microbial speciation (Mansfeldt et al. 2019).

In **Chapter 3**, the timescales of advection (1month–5 years) considerably decreased in their correlation with microbial beta diversity distance over time, suggesting that other internal and external processes contributing to microbial community composition increased over time. These align with the observations by Louca (2021), who noted that dispersal and dispersal limitations might be primarily relevant at short (ecological) time scales, while most observed differences in microbial beta diversity are likely due to local abiotic and biotic conditions. For example, temperature selection, known to be a driving mechanism for prokaryotic abundance (Logares et al. 2018), could explain the distinct prokaryotic beta diversity signals between Sweden/southern Norway and northern Norway, despite their relatively strong oceanographic connectivity (**Chapter 3**). Further, this could also explain the different relationships between productivity-specific length scales and beta diversity within different oceanographic features such as coastal upwelling in the Canary Coastal province (**Chapter 2**). Observations taken at spatial intervals equivalent to productivity-specific length scales are needed to reveal at which scale advection via surface currents is the primary structuring variable for microbial beta diversity. Additionally, to understand the spatial dynamics and temporal persistence of these biodiversity signals, a set of samples may be taken at the same time point, scattered along a spatial resolution of interest, and their data analysed in numerical particle trajectory models to identify water masses of origin and time-sensitive oceanographic connectivity between sites (e.g. as applied for fish larvae and origin of spawning area; Falcini et al., 2020).

Key outcomes

- The productivity-specific length scale is a first-order quantification of the effect of regional advection via surface ocean currents (**O3**) and provides a new perspective on spatial patchiness of pelagic microbial ecology.
- Hydrodynamic distance significantly correlated with microbial beta diversity change, confirming **H3.1**, however the magnitude of correlation decreased over time suggesting other structuring mechanisms increase in their relative importance of microbial beta diversity.

4.3.3 Coastal scales: regional signals of microbial beta diversity, with the presence of marine-terminating glaciers as an overarching structuring mechanism

Coastal areas are particularly challenging to map as beta diversity patterns are less recurrent (Chafee et al. 2018; Lemonnier et al. 2020), and as hydrodynamic features are more heterogeneous through, e.g. tidal events or land-sea exchange (Hardman-Mountford et al. 2008). I found geographic and within-fjord beta diversity signals despite the hydrodynamic heterogeneity in the Arctic (**Chapter 3**), supporting previous studies that identified horizontal transport being critical to community assemblies of planktonic species (Wassmann et al. 2015). Yet, notably, the spatial scales of these regional clusters are smaller than the ones observed in the open ocean (**Chapter 1, 2**) (Fig. 4.2), conceptually supporting the approach by Hardman-Mountford et al. (2008) that a more spatially fine-grained classification of ocean [micro]biomes is needed in coastal areas.

While fine-grained classifications provide insights into regional patches of beta diversity, large-scale comparative analyses are needed to identify structuring mechanisms and persistence of signals in relation to local environmental filtering. In the comparative study in **Chapter 3**, I found that the presence of marine-terminating glaciers was a main structuring mechanism of microbial community assembly. However, it is also an indication of regional persistence of this feature and allows insights into community structure and co-occurrence patterns, such as the observed domination of prokaryotes as central nodes in the co-occurrence network without glacial-influence,

confirming similar observations across seasonal cycles (Kellogg et al. 2019). Of course, the limitation of my analysis was a geographic bias, which requires a geographic-independent analysis through, e.g., a cross-hemisphere comparative study to confirm the persistence of structuring mechanisms.

Key outcomes

- Microbial beta diversity patterns showed multiple levels of spatial scales (geographic regional and within-fjord specific clusters) in coastal fjords (**O4**).
- Microbial beta diversity patterns statistically grouped according to the presence of marine-terminating glaciers, confirming **H4.1**.

4.3.4 Collaboration and method sharing is a key to global-scale microbial diversity observations

The field of marine microbial biogeography combines fields of evolutionary biology, ecology, geology, oceanography and geography, which requires interdisciplinary work that is also often outside one's discipline comfort zone (Gillespie 2013). In order to improve understanding and communication in interdisciplinary collaboration and research we need precise, complete and understandable method documentation. The perspectives provided in **Chapter 4** are meant to contribute to this process within the ocean observing community, to help scaling the science through more robust method intercompatibility. An important contribution in marine microbial observation efforts is the development of Minimum Information for an Omic Protocol (MIOP) (Samuel et al. 2021). The perspectives provided are directed towards inspiring and guiding communities to co-develop and standardize documentations and workflows.

Key outcomes

The perspectives provided in this thesis highlight that a cultural change of method documentation and method sharing can advance science and scale intercomparisons and multidisciplinary at global scale (**O5**).

4.4 Future directions

Forming the base of the marine food web, microbial (functional) diversity is crucial for understanding, monitoring and managing whole ocean ecosystems. Yet, many studies focus on abiotic ecosystem attributes, such as net primary productivity, without considering the biotic components of the system (reviewed in Kollmann et al. 2016), and therefore lack a mechanistic understanding through biological traits and trade-offs of ecosystem function. For example, different phytoplankton species have distinct effects on carbon fixation, recycling and export (reviewed in Litchman et al. 2015). Considering all biological, chemical and physical components in a system is also fundamental to the concept of resilience in conservation efforts (Curtin and Parker 2014), which is particularly crucial under the current global biodiversity crisis (O'Connor et al. 2020). To better manage and protect marine life, we need defined categorization of marine ecosystems. The impact the loss of diversity can have on ecosystem productivity strongly depends on functional redundancy within the community, such as presented in **Chapter 1** where stable specific primary productivity remained the same across different Longhurst provinces. In contrast, shifts at the base of the food web, as observed in **Chapter 3**, can have cascading impacts on whole ecosystem productivity and need to be further monitored at temporal and spatial scales. The integrative analyses presented in this thesis offer a framework of spatial scales of microbial biodiversity that can indicate boundaries of marine ecosystems. However, further understanding the mechanisms of microbial diversity dynamics requires both sampling and methodological advancements.

Adapting spatial resolution of measurements to productivity-specific length scales

Linking microorganisms' beta diversity to their activity (i.e., specific primary productivity) was a key approach of this thesis, which was used to identify spatial patterns of microbial biodiversity in relation to their internal turnover rate (**Chapter 2**). While this concept offers a perspective of the productivity-turnover, it does not allow a mechanistic understanding regarding population/species-

specific differential turnover rates and response to ecological niche space. Therefore, I propose using species-specific turnover rates instead of bulk-specific primary productivity measurements as they could provide more refined insights into intrinsic biological processes and spatial scales of differential microbial biodiversity change and community assembly processes. This could also help to track the distribution of indicator and/or invasive species such as harmful algae bloom development and advection. For example, observations along species-specific length scales could further resolve the mechanisms and role of advection on toxic *Alexandrium catenella* bloom initiation and development, which pose serious threats to marine life and local communities in the Arctic (Anderson et al. 2021).

Moreover, the spatial scales of microbial beta diversity change appear to be much smaller (< 10 km; **Chapter 2**, Fig. C2.2) than tested in this thesis (> 50 km; **Chapter 2**, section 4.1). Therefore, I could not directly link microbial beta diversity change to the spatial extents of the calculated productivity-specific length scales. Thus, observations at finer spatial scales are needed to validate the obtained beta diversity length scales. Moreover, the ship-based sampling of the respective campaign did not follow water mass direction. Following a Lagrangian Ansatz for observations could provide a means to separate between current and (a)biotically induced beta diversity patchiness. In addition to ship-based observations, this could, for example, also be achieved through the implementation of technological tools (e.g. ecogenomic sensors; Scholin 2013) of marine observatories (Crise et al. 2018). This could allow us to test and refine productivity-specific length scales and, by that, to observe microbial biodiversity and processes at scales that are relevant to internal dynamics and ecosystem function.

Refining measurement methodologies

New advancements in molecular microbial diversity observations offer more refined investigations of microbial diversity and function at the molecular level. For example, the amplicon sequencing technique used in this thesis is biased by the use of primers that do not cover the whole variety of taxonomic groups within one sample. I am currently working on a project where I use shotgun metagenomic samples and map the sequences of the ribosomal subunit protein L16 as a primer-

independent approach to identify microbial phylogenetic diversity of bacteria and archaea, which was previously not possible through amplicon sequence analysis due to primer bias of a few archaeal lineages (reviewed in Tahon et al. 2021). Preliminary results revealed that archaea showed more pronounced hydrographic-dependent beta diversity patterns than beta diversity patterns of bacteria, an observation that Louca (2021) also made. Identifying the diversity of key phylogenetic groups, such as archaea, in relation to ocean dispersal and hydrographic boundaries can help to refine the beta diversity patterns described in this thesis and to find indicator species to delineate ocean ecosystems.

A key missing link between phylogeny and biogeochemical rates is the inclusion of functional genes using metagenomic and metatranscriptomic approaches. Closing this gap could also provide insights into functional redundancy across phylogenetically diverse microbial communities. Metagenomic-based analyses can provide insights into ecosystem function and its relation to microbial mediated processes through culture-independent, whole-genome analyses that - for example - recently revealed the existence of non-diazotrophic *Trichodesmium* species (Delmont 2021). Metatranscriptomics could further provide insights into the organism-level responses to current environmental conditions than functional potential *per se*. Therefore, metatranscriptomic approaches could be used to elucidate key functional processes that are relevant to the survival of horizontal dispersal (e.g., a required “critical mass” of individual cells that is relevant to a species’ fitness; Angeles-Martinez and Hatzimanikatis 2021) and local environmental filtering (e.g., temperature selection; Logares et al. 2018).

Towards the whole picture of functional microbial biogeography

Functional microbial biogeography includes both spatial and temporal scales of patterns in microbial diversity and activity. At a global scale, this can only be achieved through integrated analyses of multiple datasets that are interlinked. Interlinking multiple datasets requires the provision of relevant metadata, and consistent and complete method documentation. However, these information are often limited in their access and shared only institute-wide or within a particular research community. Additionally, the multitude of methodological approaches to assess

marine microbiomes (e.g. HPLC, meta-omics) is currently only weakly interlinked and cross-compared. Such cross-comparisons can provide, for example, quantitative information (through pigment concentrations via HPLC) to the more qualitative data of genomic sequences, as shown in **Chapter 1** of this thesis. Ideally, any individual datasets should be nested in a global microbial diversity database, including the currently missing information on functional traits (also reviewed in (Martini et al. 2021), and rate measurements of C and N (sketched in Supplementary 4 Fig. S1). If approaches and collected data on microbial functional traits are harmonized and shared globally, then this could aid capacity exchange and co-development of innovative approaches within the entire research community. I believe that such collaborative efforts are urgently needed facing the current global biodiversity crisis, and I am looking forward to further contribute to and address some of these issues and outstanding questions in my future research.

5. Conclusions

In this thesis, I show the mosaic-like structure of pelagic prokaryotic and eukaryotic beta diversity patterns with heterogeneous alpha diversity patterns, and identified spatial scales of functional microbial biogeography in the Atlantic and the Indian Ocean as well as Arctic and sub-Arctic fjords. Generally, I did not observe consistent alpha diversity patterns with environmental variables (**Chapter 1, 2**), likely due to different responses of functional groups and/or correlations between environmental parameters that can amplify or suppress correlation structures with an environmental variable of interest. Specifically, primary productivity correlated distinctly with individual trophic functional groups, supporting refined trait-based observations when researching microbial alpha diversity in relation to primary productivity. Trophic functional groups also exhibited distinct basin-wide beta diversity patterns in the Atlantic Ocean (**Chapter 2**). This shows that specific lifestyles (e.g. autotrophy vs. mixotrophy) could be more beneficial for species diversification under high or low chlorophyll *a* conditions. On a broader scale, I could confirm that beta diversity patterns of all trophic functional groups follow Longhurst provinces. I could also show that microbial diversity patterns could be used to refine Longhurst' province delineations, if the persistence of these patterns are confirmed across temporal scales.

The integrated analysis of trophic-level diversity analyses bio-physically coupled to primary productivity measurement, and horizontal advection via surface currents (productivity-specific length scales; **Chapter 2**) supports a new framing of pelagic microbial ecosystems. Using this strategy, I was able to identify first-order estimated scales of oceanographic patchiness relevant to primary production that is nested within wider ecosystem boundaries. In **Chapter 3**, I found that prokaryotic and eukaryotic picoplankton beta diversity was significantly correlated with hydrodynamic distance across Arctic and sub-Arctic fjords, supporting the hypothesis that advection is a key determinant for microbial beta diversity structure. My thesis encompasses

multiple scales from a microbe's perspective to basin-wide observation. This breadth of spatial observation, as anticipated in **Chapter 4**, can be further improved and extrapolated through consistent documentation and observational cross-comparisons. Furthermore, my thesis contributes to an advanced understanding of ecosystem function and environmental heterogeneity by identifying spatial boundaries of microbial biodiversity and microbial activity (i.e., carbon and nitrogen uptake) through an oceanographic lens.

Acknowledgement

First, I would like to thank my PhD thesis Committee: Prof. Dr. Anya Waite, Prof. Dr. Matthias Ullrich, Prof. Dr. Frank Oliver Glöckner, Dr. Uwe John, Dr. Pier Luigi Buttigieg, Dr. Eric Raes, and Prof. Dr. Dieter Wolf-Gladrow. Thank you for giving me the opportunity to conducting my PhD under your supervision, thank your for your advice insights and insiring discussions. In particular, I would like to thank Prof. Dr. Anya Waite for making this project possible and for her guidance trust, energy, and focus throughout my PhD despite the geographic distance. Your insights and passion have been truly inspiring!

Additionally, I would like to thank Dr. Uwe John for adopting me in his group at AWI. His openness and his continuous support of my science, and his incredible excitement and positivity.

I would also like to thank Dr. Eric Raes for all our collaborations, which were always insightful and joyful, including to shake hundreds of N₂ fixation bottles together on the *Investigator 2am* in the morning ☺

Pier Luigi Buttigieg, I would like to thank you for your guidance, trust and support. Thank you for all your time, for your patience, understanding, and especially also your help to find a way to express myself. I'm honored that I had the chance to learn from your insights and perspectives.

Thanks to Jay Pearlman for his trust and mentorship, and Jordan van Stavel for her time and energy in all our collaborations for the OBPS. Moreover, although our collaborations are beyond my PhD work, I would also like to express my gratitude to my colleagues in the DEI task team of the UN Ocean Deace ECOP programme who are all great and inspiring women in science!

I want to thank my lab, office and working group mates Sabrina, Steffi, Nancy, Claudia, Kostas, Antonia and Sylke for all their help, comments and the fun lunch and coffee breaks. Thanks to Andreas and Susanne. Thanks to Pauline Thome for her tiredless work during the PS113, in the lab and during her Bachelor project. Of course, I would also like to thank everyone during the PS113 and IN_2019_V03 expeditions and particularly Astrid, Sonja, Simon and Jan for their support. Thanks to Jennifer Verduin for kindly hosting me during my time in Perth. Thanks to Martin

Ostrowski for the opportunity to join him and his team for a research stay at University of Technology Sydney.

I would also like to thank Niels Fuchs and Matthew Schechter for their helpful critique and insightful comments on my work.

Thanks to the graduate schools POLMAR and MARMIC, to my marmic2021 master class and specifically thanks to Matthew Schechter, David Benito Merino and Kai Blumberg for coffee and great discussions.

I would like to thank my former and current roommate(s) Niels, Damien, Nico, Pedro and Josefa for all the great time, discussions and micro-adventures. I'm particularly thankful to Niels Fuchs. Thank you for your support, insightful comments, encouraging discussions and your refreshing interdisciplinary perspectives on my work. Thanks to my close friends Jana and Tara, who always inspire me personally and scientifically. Special thanks also to Marco for your friendship and incredible excitement for any living organism. Thanks to Sanna and Aurelia for their friendship throughout so many years by now. Of course, many thanks to my family, especially my parents and my sister, for their constant support.

Last but not least, I want to express my deepest gratitude to nature. The wonder and inspirations we are able to experience every day. I am aware that knowledge gain never comes without a price. Expeditions, travel and computing add another burden in times of high pressure. I hope that in return my work will help to a more harmonized co-existence of living on this planet.

6. References

- Acevedo-Trejos, E., E. Maran, and A. Merico. 2018. Phytoplankton size diversity and ecosystem function relationships across oceanic regions. *Proc. R. Soc. B Biol. Sci.* **285**. doi:10.1098/rspb.2018.0621
- Achterberg, P., W. de Koster, and J. van der Waal. 2015. A science confidence gap: Education, trust in scientific methods, and trust in scientific institutions in the United States. *Public Underst. Sci.* **26**: 704–720. doi:10.1177/0963662515617367
- Aerts, J. 2018. From machine-readable CDISC Standard Specifications to the e-Protocol. *PhUSE EU Connect 2018*. 1–8.
- Aiken, J., Y. Pradhan, R. Barlow, S. Lavender, A. Poulton, P. Holligan, and N. Hardman-Mountford. 2009. Phytoplankton pigments and functional types in the Atlantic Ocean: A decadal assessment, 1995-2005. *Deep. Res. Part II Top. Stud. Oceanogr.* **56**: 899–917. doi:10.1016/j.dsr2.2008.09.017
- Albuquerque, R., A. Bode, J. I. González-Gordillo, C. M. Duarte, and H. Queiroga. 2021. Trophic Structure of Neuston Across Tropical and Subtropical Oceanic Provinces Assessed With Stable Isotopes. *Front. Mar. Sci.* **7**. doi:10.3389/fmars.2020.606088
- Alderkamp, A. C., V. Garcon, H. J. W. de Baar, and K. R. Arrigo. 2011. Short-term photoacclimation effects on photoinhibition of phytoplankton in the Drake Passage (Southern Ocean). *Deep. Res. Part I Oceanogr. Res. Pap.* **58**: 943–955. doi:10.1016/j.dsr.2011.07.001
- Alwazae, M. M. S., E. Perjons, and H. Kjellin. 2014. Quality measures for documentation of best practices. *Proc. Annu. Hawaii Int. Conf. Syst. Sci.* 3410–3419. doi:10.1109/HICSS.2014.423
- Anderson, D. M., E. Fachon, R. S. Pickart, and others. 2021. Evidence for massive and recurrent toxic blooms of *Alexandrium catenella* in the Alaskan Arctic. *Proc. Natl. Acad. Sci. U. S. A.* **118**. doi:10.1073/pnas.2107387118
- Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **26**: 32–46. doi:10.1046/j.1442-9993.2001.01070.x
- Anderson, T. R., and H. W. Ducklow. 2001. Microbial loop carbon cycling in ocean environments studied using a simple steady-state model. *Aquat. Microb. Ecol.* **26**: 37–49. doi:10.3354/ame026037
- Angeles-Martinez, L., and V. Hatzimanikatis. 2021. Spatio-temporal modeling of the crowding conditions and metabolic variability in microbial communities. *PLoS Comput. Biol.* **17**: 1–20. doi:10.1371/journal.pcbi.1009140
- von Appen, W.-J., V. H. Strass, A. Bracher, H. Xi, C. Hörstmann, M. H. Iversen, and A. M. Waite. 2020.

- High-resolution physical–biogeochemical structure of a filament and an eddy of upwelled water off northwest Africa. *Ocean Sci.* **16**: 253–270. doi:10.5194/os-16-253-2020
- Armstrong, F. A. J. 1951. The determination of silicate in sea water. *J. Mar. Biol. Assoc. United Kingdom* **30**: 149–160. doi:10.1017/S0025315400012649
- Azam, F., T. Fenchel, J. G. Field, J. S. Gray, L. A. Meyer - Reil, and F. Thingstad. 1983. The Ecological Role of Water - Column Microbes in the Sea. *Mar. Ecol. Prog. Ser.* **10**: 257–263. doi:10.3354/meps010257
- Azam, F., and F. Malfatti. 2007. Microbial structuring of marine ecosystems. *Nature* **5**: 782–791. doi:10.1038/nrmicro1747
- Azzaro, M., S. Aliani, G. Maimone, and others. 2021. Short-term dynamics of nutrients, planktonic abundances, and microbial respiratory activity in the Arctic Kongsfjorden (Svalbard, Norway). *Polar Biol.* **44**: 361–378. doi:10.1007/s00300-020-02798-w
- Baas Becking, L. G. M. 1934. *Geobiologie of inleiding tot de milieukunde*, The Hague, the Netherlands: W.P. Van Stockum & Zoon (in Dutch).
- Baltar, F., and J. Arístegui. 2017. Fronts at the Surface Ocean Can Shape Distinct Regions of Microbial Activity and Community Assemblages Down to the Bathypelagic Zone: The Azores Front as a Case Study. *Front. Mar. Sci.* **4**: 1–13. doi:10.3389/fmars.2017.00252
- Baltar, F., K. Currie, E. Stuck, S. Roosa, and S. E. Morales. 2016. Oceanic fronts: Transition zones for bacterioplankton community composition. *Environ. Microbiol. Rep.* **8**: 132–138. doi:10.1111/1758-2229.12362
- Bar-on, Y. M., R. Phillips, and R. Milo. 2018. The biomass distribution on Earth. *PNAS* **115**: 6506–6511. doi:10.1073/pnas.1711842115
- Barbier, M., A. Reitz, K. Pabortsava, A. C. Wolfl, T. Hahn, and F. Whoriskey. 2018. Ethical recommendations for ocean observation. *Adv. Geosci.* **45**: 343–361. doi:10.5194/adgeo-45-343-2018
- Barlow, R., V. Stuart, V. Lutz, and others. 2007. Seasonal pigment patterns of surface phytoplankton in the subtropical southern hemisphere. *Deep. Res. Part I Oceanogr. Res. Pap.* **54**: 1687–1703. doi:10.1016/j.dsr.2007.06.010
- Barnosky, A. D., E. A. Hadly, J. Bascompte, and others. 2012. Approaching a state shift in Earth's biosphere. *Nature* **486**: 52–58. doi:10.1038/nature11018
- Behrenfeld, M. J., K. H. Halsey, and A. J. Milligan. 2008. Evolved physiological responses of phytoplankton to their integrated growth environment. *Philos. Trans. R. Soc. B Biol. Sci.* **363**: 2687–2703. doi:10.1098/rstb.2008.0019
- Behrenfeld, M. J., R. T. O'Malley, D. A. Siegel, and others. 2006. Climate-driven trends in contemporary ocean productivity. *Nature* **444**: 752–755. doi:10.1038/nature05317

- Belkin, I. M., and J. E. O'Reilly. 2009. An algorithm for oceanic front detection in chlorophyll and SST satellite imagery. *J. Mar. Syst.* **78**: 319–326. doi:10.1016/j.jmarsys.2008.11.018
- Bell, J. 2014. *Doing Your Research Project: A guide for first-time researchers*, 6th ed. M.-H.E. (UK) [ed.]. (UK), McGraw-Hill Education.
- Biers, E. J., S. Sun, and E. C. Howard. 2009. Prokaryotic genomes and diversity in surface ocean waters: Interrogating the global ocean sampling metagenome. *Appl. Environ. Microbiol.* **75**: 2221–2229. doi:10.1128/AEM.02118-08
- Blain, S., B. Quéguiner, L. Armand, and others. 2007. Effect of natural iron fertilization on carbon sequestration in the Southern Ocean. *Nature* **446**: 1070–1074. doi:10.1038/nature05700
- Blain, S., G. Sarthou, and P. Laan. 2008. Distribution of dissolved iron during the natural iron-fertilization experiment KEOPS (Kerguelen Plateau, Southern Ocean). *Deep. Res. Part II Top. Stud. Oceanogr.* **55**: 594–605. doi:10.1016/j.dsr2.2007.12.028
- Boatman, T. G., P. A. Davey, T. Lawson, and R. J. Geider. 2018. The physiological cost of diazotrophy for *Trichodesmium erythraeum* IMS101. *PLoS One* **13**: e0195638. doi:10.1371/journal.pone.0195638
- Bokulich, N. A., B. D. Kaehler, J. R. Rideout, M. Dillon, E. Bolyen, R. Knight, G. A. Huttley, and J. Gregory Caporaso. 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**: 1–17. doi:10.1186/s40168-018-0470-z
- Bolaños, L. M., L. Karp-Boss, C. J. Choi, and others. 2020. Small phytoplankton dominate western North Atlantic biomass. *ISME J.* 1–12. doi:10.1038/s41396-020-0636-0
- Boss, E., A. M. Waite, J. Uitz, and others. 2020. Recommendations for plankton measurements on the GO-SHIP program with relevance to other sea-going expeditions. SCOR Working Group 154, GO-SHIP Report.
- Boyd, P. W., R. Strzepek, F. Fu, and D. A. Hutchins. 2010. Environmental control of open-ocean phytoplankton groups: Now and in the future. *Limnol. Oceanogr.* **55**: 1353–1376. doi:10.4319/lo.2010.55.3.1353
- Boyd, P. W., A. J. Watson, C. S. Law, and others. 2000. A mesoscale phytoplankton bloom in the polar Southern Ocean stimulated by iron fertilization. *Nature* **407**: 695–702. doi:10.1038/35037500
- Bracher, A., H. Xi, T. Dinter, A. Mangin, V. Strass, W. J. von Appen, and S. Wiegmann. 2020. High Resolution Water Column Phytoplankton Composition Across the Atlantic Ocean From Ship-Towed Vertical Undulating Radiometry. *Front. Mar. Sci.* **7**: 1–22. doi:10.3389/fmars.2020.00235
- Breitbarth, E., A. Oschlies, J. Laroche, E. Breitbarth, A. Oschlies, and J. L. Physiological. 2007. Physiological constraints on the global distribution of *Trichodesmium*? effect of temperature on diazotrophy. *Biogeosciences, Eur. Geosci. Union* **4**: 53–61. doi:10.5194/bg-4-53-2007
- Brett, A., J. Leape, M. Abbott, and others. 2020. COMMENT: Ocean data need a sea change to help navigate

- the warming world. *Nature* **582**: 181–183. doi:10.1038/d41586-020-01668-z
- Brown, S. L., and M. R. Landry. 2001. Mesoscale variability in biological community structure and biomass in the Antarctic Polar Front region at 170°W during austral spring 1997. *J. Geophys. Res. Ocean.* **106**: 13917–13930. doi:10.1029/1999JC000188
- Bruhn, C. S., S. Wohlrab, B. Krock, N. Lundholm, and U. John. 2021. Seasonal plankton succession is in accordance with phycotoxin occurrence in Disko Bay, West Greenland. *Harmful Algae* **103**: 101978. doi:10.1016/j.hal.2021.101978
- Buchan, A., G. R. Leclerc, C. A. Gulvik, and J. M. González. 2014. Master recyclers : features and functions of bacteria associated with phytoplankton blooms. *Nat. Publ. Gr.* doi:10.1038/nrmicro3326
- Buongiorno, J., L. C. Herbert, L. M. Wehrmann, and others. 2019. Complex Microbial Communities Drive Iron and Sulfur Cycling in Arctic Fjord Sediments. *Appl. Environ. Microbiol.* **85**: 1–16. doi:10.1128/AEM.00949-19
- Buttigieg, P. L., E. Fadeev, C. Bienhold, L. Hehemann, P. Offre, and A. Boetius. 2018. ScienceDirect Marine microbes in 4D — using time series observation to assess the dynamics of the ocean microbiome and its links to ocean health. *Curr. Opin. Microbiol.* **43**: 169–185. doi:10.1016/j.mib.2018.01.015
- Buttigieg, P. L., P. Simpson, S. Caltagirone, and J. S. Pearlman. 2019. The Ocean Best Practices System - Supporting a Transparent and Accessible Ocean. *IEEE. OCEANS 2019 MTS/IEEE SEATTLE*.
- Cadotte, M. W. 2006. Dispersal and species diversity: A meta-analysis. *Am. Nat.* **167**: 913–924. doi:10.1086/504850
- Caldeira, K., and M. Wickett. 2005. Ocean model predictions of chemistry changes from carbon dioxide emissions to the atmosphere and ocean. *J. Geophys. Res. C Ocean.* **110**: 1–12. doi:10.1029/2004JC002671
- Callahan, B. J., P. J. Mcmurdie, M. J. Rosen, A. W. Han, and A. J. A. 2016. DADA2: High resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**: 581–583. doi:10.1038/nmeth.3869.DADA2
- Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden. 2009. BLAST+ : architecture and applications. *BMC Bioinformatics* **9**: 1–9. doi:10.1186/1471-2105-10-421
- Camarena-Gómez, M. T., T. Lipsewiers, J. Piiparinen, E. Eronen-Rasimus, D. Perez-Quemaliños, L. Hoikkala, C. Sobrino, and K. Spilling. 2018. Shifts in phytoplankton community structure modify bacterial production, abundance and community composition. *Aquat. Microb. Ecol.* **81**: 149–170. doi:10.3354/ame01868
- Cantoni, C., M. J. Hopwood, J. S. Clarke, J. Chiggiato, E. P. Achterberg, and S. Cozzi. 2020. Glacial Drivers of Marine Biogeochemistry Indicate a Future Shift to More Corrosive Conditions in an Arctic Fjord.

- J. Geophys. Res. Biogeosciences **125**: 1–24. doi:10.1029/2020JG005633
- Caporaso, J. G., C. L. Lauber, W. A. Walters, D. Berg-Lyons, C. A. Lozupone, P. J. Turnbaugh, N. Fierer, and R. Knight. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 4516–4522. doi:10.1073/pnas.1000080107
- Caporaso, J. G., C. L. Laubner, E. K. Costello, and others. 2016. Moving pictures of the human microbiome. *Genome Biol.* **36**: 50–80. doi:10.1016/j.soilbio.2014.09.028
- Cardinale, B. J., D. M. Bennett, C. E. Nelson, and K. Gross. 2009. Does productivity drive diversity or vice versa? A test of the multivariate productivity-diversity hypothesis in streams. *Ecology* **90**: 1227–1241. doi:10.1890/08-1038.1
- Cardinale, B. J., K. L. Matulich, D. U. Hooper, and others. 2011. The functional role of producer diversity in ecosystems. *Am. J. Bot.* **98**: 572–592. doi:10.3732/ajb.1000364
- Cavicchioli, R., W. J. Ripple, K. N. Timmis, and others. 2019. Scientists’ warning to humanity: microorganisms and climate change. *Nat. Rev. Microbiol.* **17**: 569–586. doi:10.1038/s41579-019-0222-5
- Chafee, M., A. Fernández-Guerra, P. L. Buttigieg, G. Gerdt, A. M. Eren, H. Teeling, and R. I. Amann. 2018. Recurrent patterns of microdiversity in a temperate coastal marine environment. *ISME J.* **12**: 237–252. doi:10.1038/ismej.2017.165
- Chalup, M. S., and E. A. Laws. 1990. A test of the assumptions and predictions of recent microalgal growth models with the marine phytoplankter *Pavlova lutheri*. *Limnol. Oceanogr.* **35**: 583–596. doi:10.4319/lo.1990.35.3.0583
- Chao, A., N. J. Gotelli, T. C. Hsieh, E. L. Sander, R. K. Colwell, and A. M. Ellison. 2014. Rarefaction and Extrapolation with Hill Numbers: A Framework for Sampling and Estimation in Species Diversity Studies. *Ecol. Monogr.* **84**: 45–67. doi:10.1890/13-0133.1
- Chapman, C. C., M. A. Lea, A. Meyer, J. B. Sallée, and M. Hindell. 2020. Defining Southern Ocean fronts and their influence on biological and physical processes in a changing climate. *Nat. Clim. Chang.* **10**: 209–219. doi:10.1038/s41558-020-0705-4
- Chase, J. M., and M. A. Leibold. 2002. Spatial scale dictates the productivity-biodiversity relationship. *Nature* **416**: 427–430. doi:10.1038/416427a
- Cheng, W. H., H. P. Lu, C. C. Chen, S. Jan, and C. H. Hsieh. 2020. Vertical Beta-Diversity of Bacterial Communities Depending on Water Stratification. *Front. Microbiol.* **11**: 1–10. doi:10.3389/fmicb.2020.00449
- Clayton, S., S. Dutkiewicz, O. Jahn, and M. J. Follows. 2013. Dispersal, eddies, and the diversity of marine phytoplankton. *Limnol. Oceanogr. Fluids Environ.* **3**: 182–197. doi:10.1215/21573689-2373515
- Connell, J. H., and E. Orias. 1964. The Ecological Regulation of Species Diversity. *Am. Nat.* **98**: 399–414.

doi:10.1086/282335

- Constable, A. J., J. Melbourne-Thomas, S. P. Corney, and others. 2014. Climate change and Southern Ocean ecosystems I: How changes in physical habitats directly affect marine biota. *Glob. Chang. Biol.* **20**: 3004–3025. doi:10.1111/gcb.12623
- Conway, J. R., A. Lex, and N. Gehlenborg. 2017. UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**: 2938–2940. doi:10.1093/bioinformatics/btx364
- Cottier, F. R., F. Nilsen, R. Skogseth, V. Tverberg, J. Skardhamar, and H. Svendsen. 2010. Arctic fjords: A review of the oceanographic environment and dominant physical processes. *Geol. Soc. Spec. Publ.* **344**: 35–50. doi:10.1144/SP344.4
- Crawford, D. W., S. N. Wyatt, I. A. Wrohan, A. O. Cefarelli, K. E. Giesbrecht, B. Kelly, and D. E. Varela. 2015. Low particulate carbon to nitrogen ratios in marine surface waters of the Arctic. *Global Biogeochem. Cycles* **29**: 2021–2033. doi:10.1002/2015GB005200
- Crise, A., M. R. d’Alcalà, P. Mariani, G. Petihakis, J. Robidart, D. Iudicone, R. Bachmayer, and F. Malfatti. 2018. A conceptual framework for developing the next generation of Marine OBservatories (MOBs) for science and society. *Front. Mar. Sci.* **5**: 1–8. doi:10.3389/fmars.2018.00318
- Crummett, L. T. 2020. Acidification decreases microbial community diversity in the Salish Sea, a region with naturally high pCO₂. *PLoS One* **15**: 1–17. doi:10.1371/journal.pone.0241183
- Crump, B. C., C. S. Hopkinson, M. L. Sogin, and J. E. Hobbie. 2004. Microbial Biogeography along an Estuarine Salinity Gradient: Combined Influences of Bacterial Growth and Residence Time. *Appl. Environ. Microbiol.* **70**: 1494–1505. doi:10.1128/AEM.70.3.1494-1505.2004
- Curtin, C. G., and J. P. Parker. 2014. Foundations of resilience thinking. *Conserv. Biol.* **28**: 912–923. doi:10.1111/cobi.12321
- D’Ovidio, F., S. De Monte, S. Alvain, Y. Dandonneau, and M. Lévy. 2010. Fluid dynamical niches of phytoplankton types. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 18366–18370. doi:10.1073/pnas.1004620107
- Dance, A. 2020. The search for microbial dark matter. *Nature* **582**: 301–303. doi:10.1038/d41586-020-01684-z
- Davies, N., and D. Field. 2012. Sequencing data: A genomic network to monitor Earth. *Nature* **481**: 145. doi:10.1038/481145a
- Davies, N., D. Field, L. Amaral-Zettler, and others. 2014. The founding charter of the Genomic Observatories Network. *Gigascience* **3**: 1–5. doi:10.1186/2047-217X-3-2
- Debroas, D., I. Domaizon, J. F. Humbert, L. Jardillier, C. Lepère, A. Oudart, and N. Taib. 2017. Overview of freshwater microbial eukaryotes diversity: A first analysis of publicly available metabarcoding data. *FEMS Microbiol. Ecol.* **93**: 1–14. doi:10.1093/femsec/fix023

- Dee, D. P., S. M. Uppala, A. J. Simmons, and others. 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**: 553–597. doi:10.1002/qj.828
- Delmont, T. O. 2021. Discovery of nondiazotrophic *Trichodesmium* species abundant and widespread in the open ocean. *PNAS* **118**: e2112355118. doi:10.1073/pnas.2112355118
- Delong, E. F., C. M. Preston, T. Mincer, and others. 2006. Community Genomics Among Stratified Microbial Assemblages in the Ocean's Interior. *Science*. **311**: 496–504. doi:10.1126/science.1120250
- Devred, E., S. Sathyendranath, and T. Platt. 2007. Delineation of ecological provinces using ocean colour radiometry. *Mar. Ecol. Prog. Ser.* **346**: 1–13. doi:10.3354/meps07149
- Dickey, T. D. 2003. Emerging ocean observations for interdisciplinary data assimilation systems. *J. Mar. Syst.* **40–41**: 5–48. doi:10.1016/S0924-7963(03)00011-3
- Diepenbroek, M., and F. O. Glöckner. 2014. Towards an Integrated Biodiversity and Ecological Research Data Management and Archiving Platform: The German Federation for the Curation of Biological Data (GFBio), p. 1711–1721. *In* Plödereder, E., Grunske, L., Schneider, E. & Ull, D. (Hrsg.), Informatik 2014. Gesellschaft für Informatik e.V.
- Döös, K. 1995. Inter-ocean exchange of water masses. *J. Geophys. Res.* **100**: 499–514. doi:10.1029/95jc00337
- Duffy, J. E., and J. J. Stachowicz. 2006. Why biodiversity is important to oceanography: Potential roles of genetic, species, and trophic diversity in pelagic ecosystem processes. *Mar. Ecol. Prog. Ser.* **311**: 179–189. doi:10.3354/meps311179
- Egbert, G. D., and S. Y. Erofeeva. 2002. Efficient inverse modeling of barotropic ocean tides. *J. Atmos. Ocean. Technol.* **19**: 183–204. doi:10.1175/1520-0426(2002)019<0183:EIMOB>2.0.CO;2
- Eichner, M., S. A. Kranz, and B. Rost. 2014. Combined effects of different CO₂ levels and N sources on the diazotrophic cyanobacterium *Trichodesmium*. *Physiol. Plant.* 316–330. doi:10.1111/ppl.12172
- Elena, S. F., and R. E. Lenski. 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature* **4**: 457–469. doi:10.1038/nrg1088
- Elferink, S., U. John, S. Neuhaus, and S. Wohlrab. 2020. Functional genomics differentiate inherent and environmentally influenced traits in dinoflagellate and diatom communities. *Microorganisms* **8**: 1–22. doi:10.3390/microorganisms8040567
- Elferink, S., S. Neuhaus, S. Wohlrab, and others. 2017. Molecular diversity patterns among various phytoplankton size-fractions in West Greenland in late summer. *Deep. Res. Part I Oceanogr. Res. Pap.* **121**: 54–69. doi:10.1016/j.dsr.2016.11.002
- Evans, C., P. G. Thomson, A. T. Davidson, A. R. Bowie, R. van den Enden, H. Witte, and C. P. D. Brussaard. 2011. Potential climate change impacts on microbial distribution and carbon cycling in the Australian Southern Ocean. *Deep. Res. Part II Top. Stud. Oceanogr.* **58**: 2150–2161.

- doi:10.1016/j.dsr2.2011.05.019
- Fadeev, E., I. Salter, V. Schourup-Kristensen, and others. 2018. Microbial communities in the east and west fram strait during sea ice melting season. *Front. Mar. Sci.* **5**: 1–21. doi:10.3389/fmars.2018.00429
- Fadeev, E., M. Wietz, W. von Appen, and others. 2021. Submesoscale physicochemical dynamics directly shape bacterioplankton community structure in space and time. *Limnol. Oceanogr.* 1–13. doi:10.1002/lno.11799
- Falcini, F., R. Corrado, M. Torri, and others. 2020. Seascape connectivity of European anchovy in the Central Mediterranean Sea revealed by weighted Lagrangian backtracking and bio-energetic modelling. *Sci. Rep.* **10**: 1–13. doi:10.1038/s41598-020-75680-8
- Falkowski, P. G., T. Fenchel, and E. F. Delong. 2008. The Microbial Engines That Drive Earth ' s Biogeochemical Cycles. *Microb. Ecol.* **320**: 1034–1040. doi:10.1126/science.1153213
- Falkowski, P. G., M. E. Katz, A. H. Knoll, A. Quigg, J. A. Raven, O. Schofield, and F. J. R. Taylor. 2004. The evolution of modern eukaryotic phytoplankton. *Science.* **305**: 354–360. doi:10.1126/science.1095964
- Faure, E., S. D. Ayata, and L. Bittner. 2021. Towards omics-based predictions of planktonic functional composition from environmental data. *Nat. Commun.* **12**: 1–15. doi:10.1038/s41467-021-24547-1
- Fedoseev, A. 1970. Geostrophic circulation of surface waters on the shelf of north-west Africa. *Rapp PV Reun Cons Int Explor Mer* **159**: 32–37.
- Fernández-Méndez, M., K. A. Turk-Kubo, P. L. Buttigieg, J. Z. Rapp, T. Krumpen, J. P. Zehr, and A. Boetius. 2016. Diazotroph diversity in the sea ice, melt ponds, and surface waters of the eurasian basin of the Central Arctic Ocean. *Front. Microbiol.* **7**: 1–18. doi:10.3389/fmicb.2016.01884
- Fernandez, C., L. Farías, and O. Ulloa. 2011. Nitrogen fixation in denitrified marine waters. *PLoS One* **6**. doi:10.1371/journal.pone.0020539
- Le Fèvre, J. 1987. Aspects of the Biology of Frontal Systems, p. 163–299. *In* J.H.S. Blaxter and S. A.J. [eds.], *Advances in Marine Biology*. Academic Press INC. (London) LTD.
- Field, C. B., M. J. Behrenfeld, J. T. Randerson, P. Falkowski, C. B. Field, M. J. Behrenfeld, and J. T. Randerson. 1998. Primary Production of the Biosphere : Integrating Terrestrial and Oceanic Components. *Science.* **281**: 237–240. doi:10.1126/science.281.5374.237
- Fitzpatrick, M. C. 2020. Generalized Dissimilarity Modeling. doi:10.1111/j.1472-4642.2007.00341.x
- Frey, K. E., J. C. Comiso, L. W. Cooper, J. M. Grebmeier, and L. V Stock. 2018. Arctic Ocean Primary Productivity: The Response of Marine Algae to Climate Warming and Sea Ice Decline. *Arct. Rep. Card* 2018 2–7. doi:10.25923/vtdn-2198
- Friedlingstein, P., M. W. Jones, M. O'Sullivan, and others. 2019. Global Carbon Budget 2019. *Earth Syst. Sci. Data* **11**: 1783–1838. doi:10.5194/essd-11-1783-2019

- Friedrichs, A., K. Schwalfenberg, B. P. Koch, and O. Zielinski. 2017. Physical oceanography during MARIA S. MERIAN cruise MSM56 (MECAF). Institute for Chemistry and Biology of the Marine Environment, Carl-von-Ossietzky University of Oldenburg, Germany, PANGAEA, doi:10.1594/PANGAEA.871015.
- Frost, C. M., W. J. Allen, F. Courchamp, J. M. Jeschke, W. C. Saul, and D. A. Wardle. 2019. Using Network Theory to Understand and Predict Biological Invasions. *Trends Ecol. Evol.* **34**: 831–843. doi:10.1016/j.tree.2019.04.012
- Fuhrman, J. A., J. A. Cram, and D. M. Needham. 2015. Marine microbial community interpretation. *Nat. Publ. Gr.* **13**. doi:10.1038/nrmicro3417
- Fuhrman, J. A., J. A. Steele, I. Hewson, M. S. Schwalbach, M. V. Brown, J. L. Green, and J. H. Brown. 2008. A latitudinal diversity gradient in planktonic marine bacteria. *Proc. Natl. Acad. Sci.* **105**: 7774–8. doi:10.1073/pnas.0803070105
- Galand, P. E., O. Pereira, C. Hochart, J. C. Auguet, and D. Debroas. 2018. A strong link between marine microbial community composition and function challenges the idea of functional redundancy. *ISME J.* **12**: 2470–2478. doi:10.1038/s41396-018-0158-1
- Gallet, A., P. Koubbi, N. Léger, M. Scheifler, M. Ruiz-Rodriguez, M. T. Suzuki, Y. Desdevises, and S. Duperron. 2019. Low-diversity bacterial microbiota in Southern Ocean representatives of lanternfish genera *Electrona*, *Protomyctophum* and *Gymnoscopelus* (family *Myctophidae*). *PLoS One* **14**: 1–17. doi:10.1371/journal.pone.0226159
- Gamfeldt, L., and H. Hillebrand. 2008. Biodiversity effects on aquatic ecosystem functioning - Maturation of a new paradigm. *Int. Rev. Hydrobiol.* **93**: 550–564. doi:10.1002/iroh.200711022
- Garcia, N. S., F. Fu, C. L. Breene, P. W. Bernhardt, M. R. Mulholland, J. A. Sohm, and D. A. Hutchins. 2011. Interactive Effects Of Irradiance and CO₂ on CO₂ Fixation and N₂ Fixation in the Diazotroph *Trichodesmium Erythraeum* (Cyanobacteria). *J. Phycol.* **47**: 1292–1303. doi:10.1111/j.1529-8817.2011.01078.x
- Gasol, J. M., and X. Moran. 2015. Flow Cytometric Determination of Microbial Abundances and Its Use to Obtain Indices of Community Structure and Relative Activity. doi:10.1007/8623_2015_139
- Geider, R. J. 1987. Light and Temperature Dependence of the Carbon to Chlorophyll a Ratio in Microalgae and Cyanobacteria : Implications for Physiology and Growth of Phytoplankton. *New Phytol.* **106**: 1–34. doi:https://doi.org/10.1111/j.1469-8137.1987.tb04788.x
- Georgieva, M. N., S. Taboada, A. Riesgo, and others. 2020. Evidence of Vent-Adaptation in Sponges Living at the Periphery of Hydrothermal Vent Environments: Ecological and Evolutionary Implications. *Front. Microbiol.* **11**. doi:10.3389/fmicb.2020.01636
- Gibbons, S. M., J. G. Caporaso, M. Pirrung, D. Field, R. Knight, and J. A. Gilbert. 2013. Evidence for a

- persistent microbial seed bank throughout the global ocean. *Proc. Natl. Acad. Sci.* **110**: 4651–4655. doi:10.1073/pnas.1217767110
- Gibney, E., and R. Van Noorden. 2013. Scientists losing data at a rapid rate. *Nature*. doi:10.1038/nature.2013.14416
- Giebel, H. A., T. Brinkhoff, W. Zwisler, N. Selje, and M. Simon. 2009. Distribution of *Roseobacter* RCA and SAR11 lineages and distinct bacterial communities from the subtropics to the Southern Ocean. *Environ. Microbiol.* **11**: 2164–2178. doi:10.1111/j.1462-2920.2009.01942.x
- Gillespie, R. G. 2013. The International Biogeography Society: enabling a dynamic discipline. *Front. Biogeogr.* **5**: 0–5. doi:10.21425/f5fbg17001
- Gloor, G. B., J. M. Macklaim, V. Pawlowsky-Glahn, and J. J. Egozcue. 2017. Microbiome datasets are compositional: And this is not optional. *Front. Microbiol.* **8**: 1–6. doi:10.3389/fmicb.2017.02224
- González, M. L., V. Molina, L. Oriol, and A. J. Cavagna. 2014. Nitrogen fixation in the Southern Ocean: a case of study of the Fe-fertilized Kerguelen region (KEOPS II cruise). *Biogeosciences Discuss.* **11**: 17151–17185. doi:10.5194/bgd-11-17151-2014
- Goodwin, K. D., L. R. Thompson, B. Duarte, T. Kahlke, A. R. Thompson, J. C. Marques, and I. Caçador. 2017. DNA Sequencing as a Tool to Monitor Marine Ecological Status. *Front. Mar. Sci.* **4**. doi:10.3389/fmars.2017.00107
- Gradoville, M. R., D. Bombar, B. C. Crump, R. M. Letelier, J. P. Zehr, and A. E. White. 2017. Diversity and activity of nitrogen-fixing communities across ocean basins. *Limnol. Oceanogr.* **62**: 1895–1909. doi:10.1002/lno.10542
- Graham, E. B., J. E. Knelman, A. Schindlbacher, and others. 2016. Microbes as Engines of Ecosystem Function : When Does Community Structure Enhance Predictions of Ecosystem Processes ? *Front. Microbiol.* **7**: 1–10. doi:10.3389/fmicb.2016.00214
- Gregoracci, G. B., A. C. Dos Santos Soares, M. D. Miranda, R. Coutinho, and F. L. Thompson. 2015. Insights into the microbial and viral dynamics of a coastal downwelling-upwelling transition. *PLoS One* **10**: 1–14. doi:10.1371/journal.pone.0137090
- Griffith, G. P., H. Hop, M. Vihtakari, A. Wold, K. Kalhagen, and G. W. Gabrielsen. 2019. Ecological resilience of Arctic marine food webs to climate change. *Nat. Clim. Chang.* **9**: 868–872. doi:10.1038/s41558-019-0601-y
- Guidi, L., S. Chaffron, L. Bittner, D. Eveillard, M. Marin, and S. B. De Roscoff. 2016. Plankton networks driving carbon export in the oligotrophic ocean. *Nature* **532**: 465–470. doi:10.1038/nature16942.Plankton
- Guillou, L., D. Bachar, S. Audic, and others. 2013. The Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic*

- Acids Res. **41**: 597–604. doi:10.1093/nar/gks1160
- Hager, S. W., D. D. Harmon, and A. E. Alpine. 1984. Chemical Determination of Particulate Nitrogen in San Francisco Bay . Nitrogen Chlorophyll a Ratios in Plankton. *Estuar. Coast. Shelf Sci.* **19**: 193–204. doi:https://doi.org/10.1016/0272-7714(84)90064-7
- Halm, H., P. Lam, T. G. Ferdelman, G. Lavik, T. Dittmar, J. Laroche, S. D'Hondt, and M. M. M. Kuypers. 2012. Heterotrophic organisms dominate nitrogen fixation in the south pacific gyre. *ISME J.* **6**: 1238–1249. doi:10.1038/ismej.2011.182
- Hamilton-Taylor, J., and N. B. Price. 1983. The geochemistry of iron and manganese in the waters and sediments of Bolstadfjord, S.W. Norway. *Estuar. Coast. Shelf Sci.* **17**: 1–19. doi:10.1016/0272-7714(83)90041-0
- Han, D., H. Y. Kang, C. K. Kang, T. Unno, and H. G. Hur. 2020. Seasonal Mixing-Driven System in Estuarine–Coastal Zone Triggers an Ecological Shift in Bacterial Assemblages Involved in Phytoplankton-Derived DMSP Degradation. *Microb. Ecol.* **79**: 12–20. doi:10.1007/s00248-019-01392-w
- Hansen, M. O., T. G. Nielsen, C. A. Stedmon, and P. Munk. 2012. Oceanographic regime shift during 1997 in Disko Bay, Western Greenland. *Limnol. Oceanogr.* **57**: 634–644. doi:10.4319/lo.2012.57.2.0634
- Hanson, C. A., J. A. Fuhrman, M. C. Horner-Devine, and J. B. H. Martiny. 2012. Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nat. Rev. Microbiol.* **10**: 497–506. doi:10.1038/nrmicro2795
- Hardman-Mountford, N. J., T. Hirata, K. A. Richardson, and J. Aiken. 2008. An objective methodology for the classification of ecological pattern into biomes and provinces for the pelagic ocean. *Remote Sens. Environ.* **112**: 3341–3352. doi:10.1016/j.rse.2008.02.016
- Hartmann, M., C. Grob, G. A. Tarran, A. P. Martin, P. H. Burkill, D. J. Scanlan, and M. V. Zubkov. 2012. Mixotrophic basis of Atlantic oligotrophic ecosystems. *Proc. Natl. Acad. Sci. U. S. A.* **109**: 5756–5760. doi:10.1073/pnas.1118179109
- Hattermann, T., P. E. Isachsen, W. J. Von Appen, J. Albretsen, and A. Sundfjord. 2016. Eddy-driven recirculation of Atlantic Water in Fram Strait. *Geophys. Res. Lett.* **43**: 3406–3414. doi:10.1002/2016GL068323
- Hegseth, E. N., and V. Tverberg. 2013. Effect of Atlantic water inflow on timing of the phytoplankton spring bloom in a high Arctic fjord (Kongsfjorden, Svalbard). *J. Mar. Syst.* **113–114**: 94–105. doi:10.1016/j.jmarsys.2013.01.003
- Hellweger, F. L., E. Van Sebille, and N. D. Fredrick. 2014. Biogeographic patterns in ocean microbes emerge in a neutral agent-based model. *Science.* **345**: 1346–1349. doi:10.1126/science.1254421
- Herbert, L. C., Q. Zhu, A. B. Michaud, and others. 2021. Benthic iron flux influenced by climate-sensitive

- interplay between organic carbon availability and sedimentation rate in Arctic fjords. *Limnol. Oceanogr.* **66**: 3374–3392. doi:10.1002/lno.11885
- Hernandez, D. J., A. S. David, E. S. Menges, C. A. Searcy, and M. E. Afkhami. 2021. Environmental stress destabilizes microbial networks. *ISME J.* **15**: 1722–1734. doi:10.1038/s41396-020-00882-x
- Hernando-Morales, V., J. Ameneiro, and E. Teira. 2017. Water mass mixing shapes bacterial biogeography in a highly hydrodynamic region of the Southern Ocean. *Environ. Microbiol.* **19**: 1017–1029. doi:10.1111/1462-2920.13538
- Herndl, G. J., H. Agogue, F. Baltar, T. Reinthaler, E. Sintes, and M. M. Varela. 2008. Regulation of aquatic microbial processes: The “microbial loop” of the sunlit surface waters and the dark ocean dissected. *Aquat. Microb. Ecol.* **53**: 59–68. doi:10.3354/ame01225
- Hill, M. O. 1973. Diversity and Evenness : A Unifying Notation and Its Consequences. *Ecology* **54**: 427–432. doi:10.2307/1934352
- Hillebrand, H. 2004. On the generality of the latitudinal diversity gradient. *Am. Nat.* **163**: 192–211. doi:10.1086/381004
- Hirata, T., R. J. W. Brewin, J. Aiken, R. Barlow, K. Suzuki, and T. Isada. 2011. Synoptic relationships between surface Chlorophyll- a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences* **8**: 311–327. doi:10.5194/bg-8-311-2011
- Holm, S. 1979. A Simple Sequentially Rejective Multiple Test Procedure. *Scand. J. Stat.* **6**: 65–70.
- Holmes, F. A., N. Kirchner, J. Kuttenueler, J. Krützfeldt, and R. Noormets. 2019. Relating ocean temperatures to frontal ablation rates at Svalbard tidewater glaciers: Insights from glacier proximal datasets. *Sci. Rep.* **9**: 1–11. doi:10.1038/s41598-019-45077-3
- Hood, R. R., V. J. Coles, and D. G. Capone. 2004. Modeling the distribution of Trichodesmium and nitrogen fixation in the Atlantic Ocean. *J. Geophys. Res. C Ocean.* **109**: 1–25. doi:10.1029/2002JC001753
- Hopwood, M. J., D. Carroll, T. Dunse, and others. 2020. Review article: How does glacier discharge affect marine biogeochemistry and primary production in the Arctic? *Cryosphere* **14**: 1347–1383. doi:10.5194/tc-14-1347-2020
- Horner-Devine, M. C., M. Lage, J. B. Hughes, and B. J. M. Bohannan. 2004. A taxa-area relationship for bacteria. *Nature* **432**: 750–753. doi:10.1038/nature03073
- Hörstmann, C., P. L. Buttigieg, U. John, E. J. Raes, D. Wolf-Gladrow, A. Bracher, and A. M. Waite. 2021. Microbial diversity through an oceanographic lens: refining the concept of ocean provinces through trophic-level analysis and productivity-specific length scales. *Environ. Microbiol.* **00**. doi:10.1111/1462-2920.15832
- Hörstmann, C., P. L. Buttigieg, P. Simpson, J. Pearlman, and A. M. Waite. 2020. Towards a Best Practice for Developing Best Practices in Ocean Observation (BP4BP): Supporting Methodological Evolution

- through Actionable Documentation. Intergov. Oceanogr. Comm. Manuals Guid. **84**. doi:10.25607/OBP-781
- Hutchins, D. A., N. G. Walworth, E. A. Webb, M. A. Saito, D. Moran, M. R. McIlvin, J. Gale, and F. X. Fu. 2015. Irreversibly increased nitrogen fixation in *Trichodesmium* experimentally adapted to elevated carbon dioxide. *Nat. Commun.* **6**: 1–7. doi:10.1038/ncomms9155
- Hutchinson, G. E. 1957. Population studies-animal ecology and demography-concluding remarks. *Cold Spring Harbor symposia on quantitative biology*. COLD SPRING HARBOR LAB PRESS, PUBLICATIONS DEPT. 415–427.
- Ibarbalz, F. M., N. Henry, M. C. Brandão, and others. 2019. Global Trends in Marine Plankton Diversity across Kingdoms of Life. *Cell* **179**: 1084–1097.e21. doi:10.1016/j.cell.2019.10.008
- IPCC. 2021. 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, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, Cambridge Univ. Press 3949.
- Irwin, A. J., Z. V Finkel, O. M. E. Schofield, and P. G. Falkowski. 2006. Scaling-up from nutrient physiology to the size-structure of phytoplankton communities. *J. pl* **28**. doi:10.1093/plankt/fbi148
- Isachsen, P. E. 2015. Baroclinic instability and the mesoscale eddy field around the Lofoten Basin. *J. Geophys. Res. Ocean.* **120**: 28884–2903. doi:10.1038/175238c0
- John, U., and G. Rohardt. 2016. Physical oceanography during HEINCKE cruise HE431. Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA, doi:10.1594/PANGAEA.863438.
- John, U., and A. Wisotzki. 2017. Physical oceanography during HEINCKE cruise HE492. Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA, doi:10.1594/PANGAEA.881306.
- John, U., and A. Wisotzki. 2019. Physical oceanography during HEINCKE cruise HE533 Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA, doi:10.1594/PANGAEA.903511.
- Jönsson, B. F., and J. R. Watson. 2016. The timescales of global surface-ocean connectivity. *Nat. Commun.* **7**: 1–6. doi:10.1038/ncomms11239
- Kadoya, T., G. Gellner, and K. S. McCann. 2018. Potential oscillators and keystone modules in food webs. *Ecol. Lett.* **21**: 1330–1340. doi:10.1111/ele.13099
- Kang, S., J. L. M. Rodrigues, J. P. Ng, and T. J. Gentry. 2016. Hill number as a bacterial diversity measure framework with high-throughput sequence data. *Sci. Rep.* **6**: 1–4. doi:10.1038/srep38263
- Karl, D., A. Michaels, B. Bergman, and others. 2002. Dinitrogen fixation in the world's oceans. *Biogeochemistry* **57–58**: 47–98. doi:10.1023/A:1015798105851

- Kellogg, C. T. E., J. W. McClelland, K. H. Dunton, and B. C. Crump. 2019. Strong Seasonality in Arctic Estuarine Microbial Food Webs. *Front. Microbiol.* **10**. doi:10.3389/fmicb.2019.02628
- Kemp, W. M., and W. J. Mitsch. 1979. Turbulence and phytoplankton diversity: A general model of the “paradox of plankton.” *Ecol. Modell.* **7**: 201–222. doi:10.1016/0304-3800(79)90070-X
- Kilias, E., C. Wolf, E. M. Nöthig, I. Peeken, and K. Metfies. 2013. Protist distribution in the Western Fram Strait in summer 2010 based on 454-pyrosequencing of 18S rDNA. *J. Phycol.* **49**: 996–1010. doi:10.1111/jpy.12109
- Kim, H. H., J. Bowman, Y.-W. Luo, H. Ducklow, O. Schofield, D. Steinberg, and S. Doney. 2020a. Microbial diversity-informed modelling of polar marine ecosystem functions. *Biogeosciences Discuss.* **30**: 1–32. doi:10.5194/bg-2020-302
- Kim, H., S. Y. Kwon, K. Lee, and others. 2020b. Input of terrestrial organic matter linked to deglaciation increased mercury transport to the Svalbard fjords. *Sci. Rep.* **10**: 1–11. doi:10.1038/s41598-020-60261-6
- Klawonn, I., G. Lavik, P. Böning, H. K. Marchant, J. Dekaezemacker, W. Mohr, and H. Ploug. 2015. Simple approach for the preparation of 15-15N₂-enriched water for nitrogen fixation assessments: Evaluation, application and recommendations. *Front. Microbiol.* **6**: 1–11. doi:10.3389/fmicb.2015.00769
- Knap, A., A. Michaels, A. Close, H. Ducklow, and A. Dickson. 1996. Protocols for the Joint Global Ocean Flux Study (JGOFS) Core Measurements. JGOFS Reoprt Nr. 19 198. doi:10013/epic.27912
- Knapp, A. N. 2012. The sensitivity of marine N₂ fixation to dissolved inorganic nitrogen. *Front. Microbiol.* **3**: 1–14. doi:10.3389/fmicb.2012.00374
- Kollmann, J., S. T. Meyer, R. Bateman, and others. 2016. Integrating ecosystem functions into restoration ecology—recent advances and future directions. *Restor. Ecol.* **24**: 722–730. doi:10.1111/rec.12422
- Kopf, A., M. Bicak, R. Kottmann, J. Schnetzer, and et al. 2015. The ocean sampling day consortium. *Gigascience* **4**. doi:10.1186/s13742-015-0065-6
- Kopprio, G. A., L. H. Cuong, N. D. Luyen, T. M. Duc, T. H. Ha, L. M. Huong, and A. Gärdes. 2021. Carrageenophyte-attached and planktonic bacterial communities in two distinct bays of Vietnam: Eutrophication indicators and insights on ice-ice disease. *Ecol. Indic.* **121**. doi:10.1016/j.ecolind.2020.107067
- Kraft, N. J. B., P. B. Adler, O. Godoy, E. C. James, S. Fuller, and J. M. Levine. 2015. Community assembly, coexistence and the environmental filtering metaphor. *Funct. Ecol.* **29**: 592–599. doi:10.1111/1365-2435.12345
- Krüger, K., M. Chafee, T. Ben Francis, T. Glavina del Rio, D. Becher, T. Schweder, R. I. Amann, and H. Teeling. 2019. In marine Bacteroidetes the bulk of glycan degradation during algae blooms is mediated by few clades using a restricted set of genes. *ISME J.* **13**: 2800–2816. doi:10.1038/s41396-019-0476-

y

- Laubscher, R. K., R. Perissinotto, and C. D. McQuaid. 1993. Phytoplankton Production and Biomass at Frontal Zones in the Atlantic Sector of the Southern-Ocean. *Polar Biol.* **13**: 471–481. doi:<https://doi.org/10.1007/BF00233138>
- Laufer-Meiser, K., A. B. Michaud, M. Maisch, J. M. Byrne, A. Kappler, M. O. Patterson, H. Røy, and B. B. Jørgensen. 2021. Potentially bioavailable iron produced through benthic cycling in glaciated Arctic fjords of Svalbard. *Nat. Commun.* **12**: 1–13. doi:[10.1038/s41467-021-21558-w](https://doi.org/10.1038/s41467-021-21558-w)
- Lawson, D. 2019. Increased prevalence of human activity in the Arctic as a result of climate change, and the impacts on the Arctic ecosystem from resulting increases of introduced species by: Delaney Lawson. University of Washington Libraries.
- Lea, D. J., I. Mirouze, M. J. Martin, R. R. King, A. Hines, D. Walters, and M. Thurlow. 2015. Assessing a new coupled data assimilation system based on the met office coupled atmosphere-land-ocean-sea ice model. *Mon. Weather Rev.* **143**: 4678–4694. doi:[10.1175/MWR-D-15-0174.1](https://doi.org/10.1175/MWR-D-15-0174.1)
- Legendre, L. 1981. Hydrodynamic Control of Marine Phytoplankton Production: The Paradox of Stability. Elsevier Oceanogr. Ser. **32**: 191–207. doi:[10.1016/S0422-9894\(08\)70410-0](https://doi.org/10.1016/S0422-9894(08)70410-0)
- Lemonnier, C., M. Perennou, and L. Maignien. 2020. Linking Spatial and Temporal Dynamic of Bacterioplankton Communities With Ecological Strategies Across a Coastal Frontal Area. *Front. Mar. Sci.* **7**: 1–13. doi:[10.3389/fmars.2020.00376](https://doi.org/10.3389/fmars.2020.00376)
- Li, W. K. W., E. C. Carmack, F. A. McLaughlin, R. J. Nelson, and W. J. Williams. 2013. Space-for-time substitution in predicting the state of picoplankton and nanoplankton in a changing Arctic Ocean. *J. Geophys. Res. Ocean.* **118**: 5750–5759. doi:[10.1002/jgrc.20417](https://doi.org/10.1002/jgrc.20417)
- Litchman, E., K. F. Edwards, and P. W. Boyd. 2021. Toward trait-based food webs: Universal traits and trait matching in planktonic predator–prey and host–parasite relationships. *Limnol. Oceanogr.* 1–16. doi:[10.1002/lno.11924](https://doi.org/10.1002/lno.11924)
- Litchman, E., P. de Tezanos Pinto, K. F. Edwards, C. A. Klausmeier, C. T. Kremer, and M. K. Thomas. 2015. Global biogeochemical impacts of phytoplankton: A trait-based perspective. *J. Ecol.* **103**: 1384–1396. doi:[10.1111/1365-2745.12438](https://doi.org/10.1111/1365-2745.12438)
- Lithgow, G. J., M. Driscoll, and P. Phillips. 2017. A long journey to reproducible results. *Nature* **548**: 387–388.
- Locey, K. J., and J. T. Lennon. 2016. Scaling laws predict global microbial diversity. *Proc. Natl. Acad. Sci. U. S. A.* **113**: 5970–5975. doi:[10.1073/pnas.1521291113](https://doi.org/10.1073/pnas.1521291113)
- Logares, R., I. M. Deutschmann, C. R. Giner, and others. 2018. Different processes shape prokaryotic and picoeukaryotic assemblages in the sunlit ocean microbiome. *bioRxiv* 37–49. doi:[10.1101/374298](https://doi.org/10.1101/374298)
- Longhurst, A. 2007. *Ecological Geography of the sea*, 2nd ed. Academic Press.

- van der Loos, L. M., and R. Nijland. 2020. Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. *Mol. Ecol.* **30**: 3270–3288. doi:10.1111/mec.15592
- Louca, S. 2021. The rates of global bacterial and archaeal dispersal. *ISME J.* 24–31. doi:10.1038/s41396-021-01069-8
- Louca, S., F. Mazel, M. Doebeli, and L. W. Parfrey. 2019. A census-based estimate of earth’s bacterial and archaeal diversity. *PLoS Biol.* **17**: 1–30. doi:10.1371/journal.pbio.3000106
- Louca, S., M. F. Polz, F. Mazel, and others. 2018. Function and functional redundancy in microbial systems. *Nat. Ecol. Evol.* **2**. doi:10.1038/s41559-018-0519-1
- Louca, S., L. Wegener Parfrey, and M. Doebeli. 2016. Decoupling function and taxonomy in the global ocean microbiome. *Science* (80-.). **353**: 1272–1277.
- Lowndes, J. S. S., B. D. Best, C. Scarborough, J. C. Afflerbach, M. R. Frazier, C. C. O’Hara, N. Jiang, and B. S. Halpern. 2017. Our path to better science in less time using open data science tools. *Nat. Ecol. Evol.* **1**. doi:10.1038/s41559-017-0160
- Luo, Y.-W., S. C. Doney, L. A. Anderson, and others. 2012. Database of diazotrophs in global ocean: abundance, biomass and nitrogen fixation rates. *Earth Syst. Sci. Data* **4**: 47–73. doi:10.5194/essd-4-47-2012
- MacArthur, R. H., and E. O. Wilson. 1967. *The Theory of Island Biogeography*, Princeton University Press.
- Mackas, D. L., K. L. Denman, and M. R. Abbott. 1985. Plankton patchiness: biology in the physical vernacular. *Bull. Mar. Sci.* **37**: 653–674.
- Maguire, B. J. 1973. Niche Response Structure and the Analytical Potentials of Its Relationship to the Habitat. *Am. Nat.* **107**: 213–246. doi:10.1086/282827
- Malviya, S., E. Scalco, S. Audic, and others. 2016. Insights into global diatom distribution and diversity in the world’s ocean. *Proc. Natl. Acad. Sci.* **113**: E1516–E1525. doi:10.1073/pnas.1509523113
- Mansfeldt, C., S. Achermann, Y. Men, J. C. Walser, K. Villez, A. Joss, D. R. Johnson, and K. Fenner. 2019. Microbial residence time is a controlling parameter of the taxonomic composition and functional profile of microbial communities. *ISME J.* **13**: 1589–1601. doi:10.1038/s41396-019-0371-6
- Maranger, R., D. Vaqué, D. Nguyen, M. P. Hébert, and E. Lara. 2015. Pan-Arctic patterns of planktonic heterotrophic microbial abundance and processes: Controlling factors and potential impacts of warming. *Prog. Oceanogr.* **139**: 221–232. doi:10.1016/j.pocean.2015.07.006
- Marquardt, M., A. Vader, E. I. Stübner, M. Reigstad, and T. M. Gabrielsen. 2016. Strong seasonality of marine microbial eukaryotes in a high-Arctic fjord (Isfjorden, in West Spitsbergen, Norway). *Appl. Environ. Microbiol.* **82**: 1868–1880. doi:10.1128/AEM.03208-15
- Martin, K., K. Schmidt, A. Toseland, and others. 2021. The biogeographic differentiation of algal microbiomes in the upper ocean from pole to pole. *Nat. Commun.* **12**: 1–15. doi:10.1038/s41467-021-

25646-9

- Martini, S., F. Larras, A. Boyé, and others. 2021. Functional trait-based approaches as a common framework for aquatic ecologists. *Limnol. Oceanogr.* **66**: 965–994. doi:10.1002/lno.11655
- Martiny, J. B. H., B. J. M. Bohannan, J. H. Brown, and others. 2006. Microbial biogeography : putting microorganisms on the map. *Nature* **4**: 102–112. doi:10.1038/nrmicro1341
- Massana, R., and R. Logares. 2013. Eukaryotic versus prokaryotic marine picoplankton ecology. *Environ. Microbiol.* **15**: 1254–1261. doi:10.1111/1462-2920.12043
- Matthews, T. J., and R. J. Whittaker. 2014. Neutral theory and the species abundance distribution: Recent developments and prospects for unifying niche and neutral perspectives. *Ecol. Evol.* **4**: 2263–2277. doi:10.1002/ece3.1092
- Maturana-Martínez, C., C. Fernández, H. E. González, and P. E. Galand. 2021. Different Active Microbial Communities in Two Contrasted Subantarctic Fjords. *Front. Microbiol.* **12**. doi:10.3389/fmicb.2021.620220
- Mayol, E., M. A. Jiménez, G. J. Herndl, C. M. Duarte, and J. M. Arrieta. 2014. Resolving the abundance and air- sea fluxes of airborne microorganisms in the North Atlantic Ocean. *Front. Microbiol.* **5**: 1–9. doi:10.3389/fmicb.2014.00557
- Mcclain, C. R. 2009. A Decade of Satellite Ocean Color Observations. *Ann. Rev. Mar. Sci.* **1**: 19–42. doi:10.1146/annurev.marine.010908.163650
- McGovern, M., A. E. Poste, E. Oug, P. E. Renaud, and H. C. Trannum. 2020. Riverine impacts on benthic biodiversity and functional traits: A comparison of two sub-Arctic fjords. *Estuar. Coast. Shelf Sci.* **240**: 106774. doi:10.1016/j.ecss.2020.106774
- McKenzie, C. H., S. S. Bates, J. L. Martin, and others. 2021. Three decades of Canadian marine harmful algal events: Phytoplankton and phycotoxins of concern to human and ecosystem health. *Harmful Algae* **102**. doi:10.1016/j.hal.2020.101852
- Meire, L., J. Mortensen, P. Meire, and others. 2017. Marine-terminating glaciers sustain high productivity in Greenland fjords. *Glob. Chang. Biol.* **23**: 5344–5357. doi:10.1111/gcb.13801
- Mestre, M., and J. Höfer. 2021. The Microbial Conveyor Belt: Connecting the Globe through Dispersion and Dormancy. *Trends Microbiol.* **29**: 482–492. doi:10.1016/j.tim.2020.10.007
- Milici, M., J. Tomasch, M. L. Wos-Oxley, and others. 2016. Bacterioplankton biogeography of the Atlantic ocean: A case study of the distance-decay relationship. *Front. Microbiol.* **7**: 1–15. doi:10.3389/fmicb.2016.00590
- Mittelbach, G. G., and D. W. Schemske. 2015. Ecological and evolutionary perspectives on community assembly. *Trends Ecol. Evol.* **30**: 241–247. doi:10.1016/j.tree.2015.02.008
- Mohr, W., T. Großkopf, D. W. R. Wallace, and J. LaRoche. 2010. Methodological underestimation of

- oceanic nitrogen fixation rates. *PLoS One* **5**: 1–7. doi:10.1371/journal.pone.0012583
- Le Moigne, F. A. C., A. J. Poulton, S. A. Henson, and others. 2015. Carbon export efficiency and phytoplankton community composition in the Atlantic sector of the Arctic Ocean. *J. Geophys. Res. Ocean.* **120**: 3896–3912. doi:10.1038/175238c0
- Lo Monaco, C., M. Álvarez, R. M. Key, and others. 2010. Assessing the internal consistency of the CARINA database in the Indian sector of the Southern Ocean. *Earth Syst. Sci. Data* **2**: 51–70. doi:10.5194/essd-2-51-2010
- Mongin, M., E. Molina, and T. W. Trull. 2008. Seasonality and scale of the Kerguelen plateau phytoplankton bloom: A remote sensing and modeling analysis of the influence of natural iron fertilization in the Southern Ocean. *Deep. Res. Part II Top. Stud. Oceanogr.* **55**: 880–892. doi:10.1016/j.dsr2.2007.12.039
- Monier, A., J. Comte, M. Babin, A. Forest, A. Matsuoka, and C. Lovejoy. 2015. Oceanographic structure drives the assembly processes of microbial eukaryotic communities. *ISME J.* **9**: 990–1002. doi:10.1038/ismej.2014.197
- Montoya, J. P., M. Voss, P. Kahler, and D. G. Capone. 1996. A Simple , High-Precision , High-Sensitivity Tracer Assay for N₂ Fixation. *Appl. Environ. Microbiol.* **62**: 986–993. doi:10.1128/AEM.62.3.986-993.1996
- Mousing, E. A., K. Richardson, J. Bendtsen, I. Cetinić, and M. J. Perry. 2016. Evidence of small-scale spatial structuring of phytoplankton alpha- and beta-diversity in the open ocean. *J. Ecol.* **104**: 1682–1695. doi:10.1111/1365-2745.12634
- Müller, O., L. Seuthe, G. Bratbak, and M. L. Paulsen. 2018. Bacterial response to permafrost derived organic matter input in an Arctic Fjord. *Front. Mar. Sci.* **5**. doi:10.3389/fmars.2018.00263
- Murphy, J., and J. Riley. 1962. A modified single solution method for the determination of phosphate in natural waters. *Anal. Chem. ACTA* **27**: 31–36. doi:10.1016/S0003-2670(00)88444-5
- Neukermans, G., L. Oziel, and M. Babin. 2018. Increased intrusion of warming Atlantic water leads to rapid expansion of temperate phytoplankton in the Arctic. *Glob. Chang. Biol.* **24**: 2545–2553. doi:10.1111/gcb.14075
- Nogales, B., M. P. Lanfranconi, J. M. Piña-Villalonga, and R. Bosch. 2011. Anthropogenic perturbations in marine microbial communities. *FEMS Microbiol. Rev.* **35**: 275–298. doi:10.1111/j.1574-6976.2010.00248.x
- Nöthig, E. M., A. Bracher, A. Engel, and others. 2015. Summertime plankton ecology in Fram strait—a compilation of long- and short-term observations. *Polar Res.* **34**. doi:10.3402/polar.v34.23349
- Nti, I. K., A. F. Adekoya, M. Opoku, and P. Nimbe. 2020. Synchronising social media into teaching and learning settings at tertiary education. *Int. J. Soc. Media Interact. Learn. Environ.* **6**: 230. doi:10.1504/ijsmile.2020.109228

- Nurhas, I., T. De Fries, S. Geisler, and J. Pawlowski. 2018. Positive Computing as Paradigm to Overcome Barriers to Global Co-authoring of Open Educational Resources. *IEEE Xplore*. 283–292.
- O'Connor, B., S. Bojinski, C. Rösli, and M. E. Schaepman. 2020. Monitoring global changes in biodiversity and climate essential as ecological crisis intensifies. *Ecol. Inform.* **55**: 101033. doi:10.1016/j.ecoinf.2019.101033
- Oldham, C. E., D. E. Farrow, and S. Peiffer. 2013. A generalized Damköhler number for classifying material processing in hydrological systems. *Hydrol. Earth Syst. Sci.* **17**: 1133–1148. doi:10.5194/hess-17-1133-2013
- Oliver, M. J., and A. J. Irwin. 2008. Objective global ocean biogeographic provinces. *Geophys. Res. Lett.* **35**: 1–6. doi:10.1029/2008GL034238
- Olson, D. M., E. Dinerstein, E. D. Wikramanayake, and others. 2001. Terrestrial ecoregions of the world: A new map of life on Earth. *Bioscience* **51**: 933–938. doi:10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2
- Oziel, L., A. Baudena, M. Ardyna, and others. 2020. Faster Atlantic currents drive poleward expansion of temperate phytoplankton in the Arctic Ocean. *Nat. Commun.* **11**: 1–8. doi:10.1038/s41467-020-15485-5
- Padgham, M., M. D. Sumner, and C. F. F. Karney. 2020. geodist R package version 0.0.4.
- Parada, A. E., D. M. Needham, and J. A. Fuhrman. 2016. 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. doi:10.1111/1462-2920.13023
- Partensky, F., W. R. Hess, and D. Vaultot. 1999. *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol. Mol. Biol. Rev.* **63**: 106–127. doi:10.1128/MMBR.63.1.106-127.1999
- Pearlman, J., M. Bushnell, L. Coppola, and others. 2019. Evolving and sustaining ocean best practices and standards for the next decade. *Front. Mar. Sci.* **6**: 1–19. doi:10.3389/fmars.2019.00277
- Pechar, E., T. Bernauer, and F. Mayer. 2018. Beyond Political Ideology: The Impact of Attitudes Towards Government and Corporations on Trust in Science. *SAGE journals Sci. Commun.* **40**: 291–313. doi:10.1177/1075547018763970
- Peschel, S., C. L. Müller, E. von Mutius, A.-L. Boulesteix, and M. Depner. 2020. NetCoMi: network construction and comparison for microbiome data in R. *Brief. Bioinform.* **00**: 1–18. doi:10.1093/bib/bbaa290
- Peter, H., S. Beier, S. Bertilsson, E. S. Lindstro, S. Langenheder, and L. J. Tranvik. 2011. Function-specific response to depletion of microbial diversity. *ISME J.* **5**: 351–361. doi:10.1038/ismej.2010.119
- Pielou, E. C. 1975. *Ecological diversity*, John Wiley & Sons.
- Pinhassi, J., M. M. Sala, H. Havskum, F. Peters, Ò. Guadayol, A. Malits, and C. Marrasé. 2004. Changes in

- bacterioplankton composition under different phytoplankton regimens. *Appl. Environ. Microbiol.* **70**: 6753–6766. doi:10.1128/AEM.70.11.6753-6766.2004
- Pironon, S., J. Villellas, W. Thuiller, V. M. Eckhart, M. A. Geber, D. A. Moeller, and M. B. García. 2018. The ‘Hutchinsonian niche’ as an assemblage of demographic niches: implications for species geographic ranges. *Ecography (Cop.)*. **41**: 1103–1113. doi:10.1111/ecog.03414
- Polyakov, I. V., A. V. Pnyushkov, and E. C. Carmack. 2018. Stability of the arctic halocline: A new indicator of arctic climate change. *Environ. Res. Lett.* **13**. doi:10.1088/1748-9326/aaec1e
- Powell, T. M., P. J. Richerson, T. M. Dillon, B. A. Agee, B. J. Dozier, D. A. Godden, and L. O. Myrup. 1975. Spatial scales of current speed and phytoplankton biomass fluctuations in Lake Tahoe. *Science*. **189**: 1088–1090. doi:10.1126/science.189.4208.1088
- Preston, F. W. 1960. Time and Space and the Variation of Species. *Ecology* **41**: 612–627. doi:10.2307/1931793
- Pruesse, E., J. Peplies, F. O. Glöckner, A. Editor, and J. Wren. 2012. SINA : Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**: 1823–1829. doi:10.1093/bioinformatics/bts252
- Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, F. O. Glo, and P. Yarza. 2013. The SILVA ribosomal RNA gene database project : improved data processing and web-based tools. *Nucleic Acids Res.* **41**: D590–D596. doi:10.1093/nar/gks1219
- Raes, E. J., L. Bodrossy, J. Van De Kamp, A. Bissett, M. Ostrowski, and M. V Brown. 2018. Oceanographic boundaries constrain microbial diversity gradients in the South Pacific Ocean. *PNAS* **115**: E8266–E8275. doi:10.1073/pnas.1719335115
- Raes, E. J., J. Van De Kamp, L. Bodrossy, and others. 2020. N₂ Fixation and New Insights Into Nitrification From the Ice-Edge to the Equator in the South Pacific Ocean. **7**: 1–20. doi:10.3389/fmars.2020.00389
- Raes, E. J., P. A. Thompson, A. S. McInnes, H. M. Nguyen, N. Hardman-mountford, and A. M. Waite. 2015. Sources of new nitrogen in the Indian Ocean. *Global Biogeochem. Cycles* **935:8**: 1283–1297. doi:10.1002/2015GB005194
- Raes, E. J., A. M. Waite, A. S. McInnes, H. Olsen, H. M. Nguyen, N. Hardman-Mountford, and P. A. Thompson. 2014. Changes in latitude and dominant diazotrophic community alter N₂ fixation. *Mar. Ecol. Prog. Ser.* **516**: 85–102. doi:10.3354/meps11009
- Raes, J., I. Letunic, T. Yamada, L. J. Jensen, and P. Bork. 2011. Toward molecular trait-based ecology through integration of biogeochemical, geographical and metagenomic data. *Mol. Syst. Biol.* **7**: 1–9. doi:10.1038/msb.2011.6
- Ras, J., H. Claustre, and J. Uitz. 2008. Spatial variability of phytoplankton pigment distributions in the Subtropical South Pacific Ocean: Comparison between in situ and predicted data. *Biogeosciences* **5**:

- 353–369. doi:10.5194/bg-5-353-2008
- Reese, A. T., and R. R. Dunn. 2018. Drivers of Microbiome Biodiversity: A Review of General Rules, Feces, and Ignorance. *ASM Journals* **9**: e01294-18. doi:10.1128/mBio.01294-18
- Ren, L., Y. Liu, T. L. Lauridsen, and others. 2021. Warming exacerbates the impact of nutrient enrichment on microbial functional potentials important to the nutrient cycling in shallow lake mesocosms. *Limnol. Oceanogr.* **66**: 2481–2495. doi:10.1002/lno.11766
- Révelard, A., J. Tintoré, J. Verron, and others. 2021. Ocean Integration: the needs and challenges of effective coordination within the ocean observing system. *Front. Mar. Sci.* [*provisionally accepted*]
- Richter, D., R. Watteaux, T. Vannier, and others. 2019. Genomic evidence for global ocean plankton biogeography shaped by large-scale current systems evidence for global ocean plankton biogeography shaped by large-scale current systems. [preprint] *bioRxiv* 1–38. doi:10.1101/867739
- Rohde, K. 1992. Latitudinal Gradients in Species Diversity: The Search for the Primary Cause. *Nord. Soc. Oikos* **65**: 514–527.
- Rothstein, H. R., A. J. Sutton, and M. Borenstein. 2005. Publication Bias in Meta-Analysis, p. 1–7. *In* Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments.
- Röttgers, L., and K. Faust. 2018. From hairballs to hypotheses—biological insights from microbial networks. *FEMS Microbiol. Rev.* **42**: 761–780. doi:10.1093/femsre/fuy030
- Ryabinin, V., J. Barbière, P. Haugan, and others. 2019. The UN decade of ocean science for sustainable development. *Front. Mar. Sci.* **6**. doi:10.3389/fmars.2019.00470
- Sambrotto, R. N., and B. J. Mace. 2000. Coupling of biological and physical regimes across the Antarctic Polar Front as reflected by nitrogen production and recycling. *Deep. Res. Part II* **47**: 3339–3367. doi:10.1016/S0967-0645(00)00071-0
- Samuel, R. M., R. Meyer, P. L. Buttigieg, and others. 2021. Toward a Global Public Repository of Community Protocols to Encourage Best Practices in Biomolecular Ocean Observing and Research. *Front. Mar. Sci.* **8**: 1–8. doi:10.3389/fmars.2021.758694
- Santi, I., A. Tsiola, P. D. Dimitriou, and others. 2019. Prokaryotic and eukaryotic microbial community responses to N and P nutrient addition in oligotrophic Mediterranean coastal waters: Novel insights from DNA metabarcoding and network analysis. *Mar. Environ. Res.* **150**: 104752. doi:10.1016/j.marenvres.2019.104752
- Sarmiento, H., and J. M. Gasol. 2012. Use of phytoplankton-derived dissolved organic carbon by different types of bacterioplankton. *Environ. Microbiol.* **14**: 2348–2360. doi:10.1111/j.1462-2920.2012.02787.x
- Sarmiento, H., J. M. Montoya, E. Vázquez-Domínguez, D. Vaqué, and J. M. Gasol. 2010. Warming effects on marine microbial food web processes: How far can we go when it comes to predictions? *Philos.*

- Trans. R. Soc. B Biol. Sci. **365**: 2137–2149. doi:10.1098/rstb.2010.0045
- Sarmiento, J. L., R. Slater, R. Barber, and others. 2004. Response of ocean ecosystems to climate warming. *Global Biogeochem. Cycles* **18**. doi:10.1029/2003GB002134
- Scholin, C. A. 2013. Ecogenomic Sensors. *Encycl. Biodivers.* Second Ed. **2**: 690–700. doi:10.1016/B978-0-12-384719-5.00408-1
- Seebens, H., T. M. Blackburn, E. E. Dyer, and others. 2018. Global rise in emerging alien species results from increased accessibility of new source pools. *Proc. Natl. Acad. Sci. U. S. A.* **115**: E2264–E2273. doi:10.1073/pnas.1719429115
- Seibold, S., M. W. Cadotte, J. S. Macivor, S. Thorn, and J. Müller. 2018. The Necessity of Multitrophic Approaches in Community Ecology. *Trends Ecol. Evol.* **33**: 754–764. doi:10.1016/j.tree.2018.07.001
- Seifert, M., M. Hoppema, C. Burau, and others. 2019. Influence of glacial meltwater on summer biogeochemical cycles in Scoresby Sund, East Greenland. *Front. Mar. Sci.* **6**. doi:10.3389/fmars.2019.00412
- Seymour, J. R., S. A. Amin, J. B. Raina, and R. Stocker. 2017. Zooming in on the phycosphere: The ecological interface for phytoplankton-bacteria relationships. *Nat. Microbiol.* **2**. doi:10.1038/nmicrobiol.2017.65
- Shchepetkin, A. F., and J. C. McWilliams. 2005. The regional oceanic modeling system (ROMS): A split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Model.* **9**: 347–404. doi:10.1016/j.ocemod.2004.08.002
- Shiozaki, T., D. Bombar, L. Riemann, and others. 2017. Basin scale variability of active diazotrophs and nitrogen fixation in the North Pacific, from the tropics to the subarctic Bering Sea. *Global Biogeochem. Cycles* **31**: 996–1009. doi:10.1002/2017GB005681
- Shiozaki, T., A. Fujiwara, M. Ijichi, N. Harada, S. Nishino, S. Nishi, T. Nagata, and K. Hamasaki. 2018. Diazotroph community structure and the role of nitrogen fixation in the nitrogen cycle in the Chukchi Sea (western Arctic Ocean). *Limnol. Oceanogr.* doi:10.1002/lno.10933
- Simpson, P., F. Pearlman, and J. Pearlman. 2018. Evolving and Sustaining Ocean Best Practices Workshop, 15 – 17 November 2017. *Intergovernmental Oceanographic Commission, Paris, France: Proceedings. AtlantOS/ODIP/OORCN Ocean Best Practices Working Group.* 74pp.
- Simpson, P., J. Pearlman, and F. Pearlman. 2020. Evolving and Sustaining Ocean Best Practices Workshop III. *IOC- IODE: GOOS and IEEE Oceanic Engineering Society.* 37pp.
- Sipler, R. E., D. Gong, S. E. Baer, M. P. Sanderson, Q. N. Roberts, M. R. Mulholland, and D. A. Bronk. 2017. Preliminary estimates of the contribution of Arctic nitrogen fixation to the global nitrogen budget. *Limnol. Oceanogr. Lett.* 159–166. doi:10.1002/lol2.10046
- Sommeria-Klein, G., R. Watteaux, D. Iudicone, C. Bowler, and H. Morlon. 2021. Global drivers of

- eukaryotic plankton bioregionalization in the upper ocean. *Science*. **374**: 594–599. doi:10.1126/science.abb3717
- De Souza, R. B., and I. S. Robinson. 2004. Lagrangian and satellite observations of the Brazilian Coastal Current. *Cont. Shelf Res.* **24**: 241–262. doi:10.1016/j.csr.2003.10.001
- Stec, K. F., L. Caputi, P. L. Buttigieg, and others. 2017. Modelling plankton ecosystems in the meta-omics era. Are we ready? *Mar. Genomics* **32**: 1–17. doi:10.1016/j.margen.2017.02.006
- Steiner, C. F., and M. A. Leibold. 2004. Cyclic Assembly Trajectories and Scale-dependent Productivity – Diversity Relationships. *Ecology* **85**: 107–113. doi:10.1890/03-3010
- Stoeck, T., D. Bass, M. Nebel, R. Christen, and D. Meredith. 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* **19**: 21–31. doi:10.1111/j.1365-294X.2009.04480.x
- Storkey, D., E. W. Blockley, R. Furner, and others. 2010. Forecasting the ocean state using NEMO:The new FOAM system. *J. Oper. Oceanogr.* **3**: 3–15. doi:10.1080/1755876X.2010.11020109
- Strass, V. 2018. The Expedition PS113 of the Research Vessel POLARSTERN to the Atlantic Ocean in 2018. *Berichte zur Polar- und Meeresforsch.* **714**. doi:10.2312/BzPM_0724_2018
- Strass, V. H., A. C. Naveira Garabato, R. T. Pollard, and others. 2002. Mesoscale frontal dynamics: Shaping the environment of primary production in the Antarctic Circumpolar Current. *Deep. Res. Part II Top. Stud. Oceanogr.* **49**: 3735–3769. doi:10.1016/S0967-0645(02)00109-1
- Stukel, M. R., M. R. Landry, C. R. Benitez-Nelson, and R. Goericke. 2011. Trophic cycling and carbon export relationships in the California current ecosystem. *Limnol. Oceanogr.* **56**: 1866–1878. doi:10.4319/lo.2011.56.5.1866
- Sunagawa, S., L. P. Coelho, S. Chaffron, and others. 2015. Structure and function of the global ocean microbiome. *Science* **348**: 1261359. doi:10.1126/science.1261359
- Swart, N. C., S. T. Gille, J. C. Fyfe, and N. P. Gillett. 2018. Recent Southern Ocean warming and freshening driven by greenhouse gas emissions and ozone depletion. *Nat. Geosci.* **11**: 836–841. doi:10.1038/s41561-018-0226-1
- Swingland, I. R. 2001. Biodiversity, Definition of. *Encycl. Biodivers.* **1**: 377–391.
- Szeligowska, M., E. Trudnowska, R. Boehnke, A. M. Dąbrowska, K. Dragańska-Deja, K. Deja, M. Darecki, and K. Błachowiak-Samołyk. 2021. The interplay between plankton and particles in the Isfjorden waters influenced by marine- and land-terminating glaciers. *Sci. Total Environ.* **780**. doi:10.1016/j.scitotenv.2021.146491
- Tahon, G., P. Geesink, and T. J. G. Ettema. 2021. Expanding Archaeal Diversity and Phylogeny: Past, Present, and Future. *Annu. Rev. Microbiol.* **75**: 359–381. doi:10.1146/annurev-micro-040921-050212
- Talaber, I., J. Francé, V. Flander-Putrlé, and P. Mozetič. 2018. Primary production and community structure

- of coastal phytoplankton in the Adriatic Sea: Insights on taxon-specific productivity. *Mar. Ecol. Prog. Ser.* **604**: 65–81. doi:10.3354/meps12721
- Tang, W., Z. Li, and N. Cassar. 2019. Machine Learning Estimates of Global Marine Nitrogen Fixation. *JGR Biogeosciences* 717–730. doi:10.1029/2018JG004828
- Tanhua, T., S. Pouliquen, J. Hausman, and others. 2019. Ocean FAIR data services. *Front. Mar. Sci.* **6**. doi:10.3389/fmars.2019.00440
- Taylor, B. B., E. Torrecilla, A. Bernhardt, M. H. Taylor, I. Peeken, R. Röttgers, J. Piera, and A. Bracher. 2011. Bio-optical provinces in the eastern Atlantic Ocean and their biogeographical relevance. *Biogeosciences* **8**: 3609–3629. doi:10.5194/bg-8-3609-2011
- Teeling, H., B. M. Fuchs, D. Becher, and others. 2012. Substrate-Controlled Succession of Marine Bacterioplankton Populations Induced by a Phytoplankton Bloom. *Science* (80-.). **336**: 608–611. doi:10.1126/science.1218344
- Teira, E., R. Logares, A. Gutiérrez-Barral, I. Ferrera, M. M. Varela, X. A. G. Morán, and J. M. Gasol. 2019. Impact of grazing, resource availability and light on prokaryotic growth and diversity in the oligotrophic surface global ocean. *Environ. Microbiol.* **21**: 1482–1496. doi:10.1111/1462-2920.14581
- Testor, P., B. DeYoung, D. L. Rudnick, and others. 2019. OceanGliders: A component of the integrated GOOS. *Front. Mar. Sci.* **6**: 1–32. doi:10.3389/fmars.2019.00422
- Teytelman, L. 2018. No more excuses for non-reproducible methods. *Nature* **560**: 411. doi:10.1038/d41586-018-06008-w
- Thomas, M. K., C. T. Kremer, C. A. Klausmeier, and E. Litchman. 2012. A global pattern of thermal adaptation in marine phytoplankton. *Science* (80-.). **338**: 1085–1088. doi:10.1126/science.1224836
- Thompson, A. W., R. A. Foster, A. Krupke, B. J. Carter, N. Musat, D. Vaultot, M. M. M. Kuypers, and J. P. Zehr. 2012. Unicellular Cyanobacterium Symbiotic with a Single-Celled Eukaryotic Alga. *Science*. **337**: 1546–1550. doi:10.1126/science.1222700
- Thompson, L. R., C. Field, T. Romanuk, D. Kamanda Ngugi, R. Siam, H. El Dorry, and U. Stingl. 2013. Patterns of ecological specialization among microbial populations in the Red Sea and diverse oligotrophic marine environments. *Ecol. Evol.* **3**: 1780–1797. doi:10.1002/ece3.593
- Tilman, D., D. Wedin, and J. Knops. 1996. Evolutionary History of the Marsupials and an Analysis of Osteological Characters. Frederick S. Szalay. *Nature* **379**: 718–720. doi:10.1086/419118
- Tipton, L., C. L. Müller, Z. D. Kurtz, L. Huang, E. Kleerup, A. Morris, R. Bonneau, and E. Ghedin. 2018. Fungi stabilize connectivity in the lung and skin microbial ecosystems. *Microbiome* **6**: 1–14. doi:10.1186/s40168-017-0393-0
- Tonelli, M., C. N. Signori, A. Bendia, J. Neiva, B. Ferrero, V. Pellizari, and I. Wainer. 2021. Climate Projections for the Southern Ocean Reveal Impacts in the Marine Microbial Communities Following

- Increases in Sea Surface Temperature. *Front. Mar. Sci.* **8**: 1–11. doi:10.3389/fmars.2021.636226
- Townsend, D. W., and N. R. Pettigrew. 1997. Nitrogen limitation of secondary production on Georges Bank. *J. Plankton Res.* **19**: 221–235. doi:10.1093/plankt/19.2.221
- Troussellier, M., A. Escalas, T. Bouvier, and D. Mouillot. 2017. Sustaining rare marine microorganisms: Macroorganisms as repositories and dispersal agents of microbial diversity. *Front. Microbiol.* **8**: 1–17. doi:10.3389/fmicb.2017.00947
- Turner, J. R. G. 2004. Explaining the global biodiversity gradient: Energy, area, history and natural selection. *Basic Appl. Ecol.* **5**: 435–448. doi:10.1016/j.baae.2004.08.004
- Uitz, J., H. Claustre, A. Morel, and S. B. Hooker. 2006. Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. *J. Geophys. Res. Ocean.* **111**. doi:10.1029/2005JC003207
- Ustick, L. J., A. A. Larkin, C. A. Garcia, and others. 2021. Metagenomic analysis reveals global-scale patterns of ocean nutrient limitation. *Science* (80-.). **372**: 287–291. doi:10.1126/science.abe6301
- Vallina, S. M., M. J. Follows, S. Dutkiewicz, J. M. Montoya, P. Cermeno, and M. Loreau. 2014. Global relationship between phytoplankton diversity and productivity in the ocean. *Nat. Commun.* **5**: 1–10. doi:10.1038/ncomms5299
- Vaulot, D., W. Eikrem, M. Viprey, and H. Moreau. 2008. The diversity of small eukaryotic phytoplankton ($\leq 3 \mu\text{m}$) in marine ecosystems. *FEMS Microbiol. Rev.* **32**: 795–820. doi:10.1111/j.1574-6976.2008.00121.x
- Vellend, M., and A. Agrawal. 2010. Conceptual Synthesis in Community Ecology. *Q. Rev. Biol.* **85**: 183–206. doi:https://doi.org/10.1086/652373
- Vellend, M., D. S. Srivastava, K. M. Anderson, and others. 2014. Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos* **123**: 1420–1430. doi:10.1111/oik.01493
- Verde, C., D. Giordano, C. M. Bellas, G. di Prisco, and A. M. Anesio. 2016. Polar Marine Microorganisms and Climate Change, p. 187–215. *In* *Advances in Microbial Physiology*. Elsevier Ltd.
- Vichi, M., J. I. Allen, S. Masina, and N. J. H. Mountford. 2011. The emergence of ocean biogeochemical provinces : A quantitative assessment and a diagnostic for model evaluation. *Global Biogeochem. Cycles* **25**: 1–17. doi:10.1029/2010GB003867
- Vidussi, F., H. Claustre, B. B. Manca, A. Luchetta, and J.-C. Marty. 2001. Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter. *J. Geophys. Res.* **106**: 939–956. doi:10.1029/1999JC000308
- Villarino, E., J. R. Watson, B. Jönsson, and others. 2018. Large-scale ocean connectivity and planktonic body size. *Nat. Commun.* **9**. doi:10.1038/s41467-017-02535-8
- Vilmi, A., C. Gibert, G. Escarguel, and others. 2021. Dispersal–niche continuum index: a new quantitative

- metric for assessing the relative importance of dispersal versus niche processes in community assembly. *Ecography (Cop.)*. **44**: 370–379. doi:10.1111/ecog.04083
- Vincent, W. F. 2010. Microbial ecosystem responses to rapid climate change in the Arctic. *ISME J.* **4**: 1089–1091. doi:10.1038/ismej.2010.108
- Violle, C., M.-L. Navas, D. Vile, E. Kazakou, C. Fortunel, I. Hummel, and E. Garnier. 2007. Let the concept of trait be functional! *Oikos* **116**: 882–892. doi:10.1111/j.2007.0030-1299.15559.x
- Waite, A. M., B. A. Muhling, C. M. Holl, L. E. Beckley, J. P. Montoya, J. Strzelecki, P. A. Thompson, and S. Pesant. 2007. Food web structure in two counter-rotating eddies based on $\delta^{15}\text{N}$ and $\delta^{13}\text{C}$ isotopic analyses. *Deep. Res. Part II Top. Stud. Oceanogr.* **54**: 1055–1075. doi:10.1016/j.dsr2.2006.12.010
- Wang, J., F. Pan, J. Soininen, J. Heino, and J. Shen. 2016. Nutrient enrichment modifies temperature-biodiversity relationships in large-scale field experiments. *Nat. Commun.* **7**: 1–9. doi:10.1038/ncomms13960
- Wassmann, P., K. N. Kosobokova, D. Slagstad, and others. 2015. The contiguous domains of Arctic Ocean advection: trails of life and death. *Prog. Oceanogr.* **139**: 42–65. doi:10.1016/j.pocean.2015.06.011
- Wehrmann, L. M., N. J. Knab, H. Pirlet, V. Unnithan, C. Wild, and T. G. Ferdelman. 2009. Carbon mineralization and carbonate preservation in modern cold-water coral reef sediments on the Norwegian shelf. *Biogeosciences* **6**: 663–680. doi:10.5194/bg-6-663-2009
- Weichselgartner, J., and R. Kaspersen. 2010. Barriers in the science-policy-practice interface: Toward a knowledge-action-system in global environmental change research. *Glob. Environ. Chang.* **20**: 266–277. doi:10.1016/j.gloenvcha.2009.11.006
- Weissgerber, T. L., V. D. Garovic, S. J. Winham, N. M. Milic, and E. M. Prager. 2016. Transparent reporting for reproducible science. *J. Neurosci. Res.* **94**: 859–864. doi:10.1002/jnr.23785
- Whitt, C., J. Pearlman, B. Polagye, and others. 2020. Future Vision for Autonomous Ocean Observations. *Front. Mar. Sci.* **7**. doi:10.3389/fmars.2020.00697
- Wiedmann, I., E. Ershova, B. A. Bluhm, E. M. Nöthig, R. R. Gradinger, K. Kosobokova, and A. Boetius. 2020. What Feeds the Benthos in the Arctic Basins? Assembling a Carbon Budget for the Deep Arctic Ocean. *Front. Mar. Sci.* **7**. doi:10.3389/fmars.2020.00224
- Wiedmann, I., M. Reigstad, M. Marquardt, A. Vader, and T. M. Gabrielsen. 2016. Seasonality of vertical flux and sinking particle characteristics in an ice-free high arctic fjord-Different from subarctic fjords? *J. Mar. Syst.* **154**: 192–205. doi:10.1016/j.jmarsys.2015.10.003
- Wilkins, D., F. M. Lauro, T. J. Williams, and others. 2013. Biogeographic partitioning of Southern Ocean microorganisms revealed by metagenomics. *Environ. Microbiol.* **15**: 1318–1333. doi:10.1111/1462-2920.12035
- Wilkinson, M. D., M. Dumontier, Ij. J. Aalbersberg, and others. 2016. Comment: The FAIR Guiding

- Principles for scientific data management and stewardship. *Sci. Data* **3**: 1–9. doi:10.1038/sdata.2016.18
- Wolf, A. A., J. L. Funk, P. C. Selman, C. N. Morozumi, D. L. Hernández, J. R. Pasari, and E. S. Zavaleta. 2021. Trait-based filtering mediates the effects of realistic biodiversity losses on ecosystem functioning. *Proc. Natl. Acad. Sci. U. S. A.* **118**. doi:10.1073/pnas.2022757118
- Wood, E. D., F. A. J. Armstrong, and F. A. Richards. 1967. Determination of nitrate in sea water by cadmium-copper reduction to nitrite. *J. Mar. Biol. Assoc. United Kingdom* **47**: 23–31. doi:10.1017/S002531540003352X
- Woodall, L. C., S. Talma, O. Steeds, P. Stefanoudis, M. M. Jeremie-Muzungaile, and A. De Comarmond. 2021. Co-development, co-production and co-dissemination of scientific research: A case study to demonstrate mutual benefits. *Biol. Lett.* **17**. doi:10.1098/rsbl.2020.0699
- Woodward, F. I. 1987. *Climate and plant distribution.*, Cambridge University Press.
- Woodward, F. I., M. R. Lomas, and C. K. Kelly. 2004. Global climate and the distribution of plant biomes. *Philos. Trans. R. Soc. B Biol. Sci.* **359**: 1465–1476. doi:10.1098/rstb.2004.1525
- Yang, H., G. Lohmann, U. Krebs-Kanzow, and others. 2020a. Poleward Shift of the Major Ocean Gyres Detected in a Warming Climate. *Geophys. Res. Lett.* **47**. doi:10.1029/2019GL085868
- Yang, J., H. Pei, J. Lv, Q. Liu, F. Nan, and X. Liu. 2020b. Seasonal co-occurrence patterns of bacterial and eukaryotic phytoplankton communities and ecological effects of keystone taxa in an urban river. *Authorea Prepr.* 1–20. doi:https://doi.org/10.22541/au.159602488.89942318
- Yeh, Y.-C., D. M. Needham, E. T. Sieradzki, and J. A. Fuhrman. 2018. Taxon Disappearance from Microbiome Analysis Reinforces the Value of Mock Communities as a Standard in Every Sequencing Run. *MSystems* **3**: e00023-18.
- Yilmaz, P., R. Kottmann, D. Field, and others. 2011. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIXS) specifications. *Nature* **29**: 415–420. doi:10.1038/nbt.1823
- Zeebe, R., and D. Wolf-Gladrow. 2001. *CO2 in seawater: equilibrium, kinetics, isotopes.* No. 65, Gulf Professional Publishing.
- Zhou, J., Y. Deng, L. Shen, and others. 2016. Temperature mediates continental-scale diversity of microbes in forest soils. *Nat. Commun.* **7**. doi:10.1038/ncomms12083
- Zhou, J., Y. Deng, P. Zhang, and others. 2014. Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. *Proc. Natl. Acad. Sci. U. S. A.* **111**. doi:10.1073/pnas.1324044111
- Zhou, J., S. Kang, C. W. Schadt, and C. T. Garten. 2008. Spatial scaling of functional gene diversity across various microbial taxa. *Proc. Natl. Acad. Sci. U. S. A.* **105**: 7768–7773. doi:10.1073/pnas.0709016105
- Zielinski, O., D. Voß, D. Meier, R. Henkel, L. Holinde, S. P. Garaba, and A. Cembella. 2013. Physical

- oceanography during Maria S. Merian cruise MSM21/3 (ARCHEMHAB). Institute for Chemistry and Biology of the Marine Environment, Carl-von-Ossietzky University of Oldenburg, Germany, PANGAEA, doi:10.1594/PANGAEA.819731.
- Zinger, L., L. A. Amaral-Zettler, J. A. Fuhrman, and others. 2011. Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. *PLoS One* **6**: 1–11. doi:10.1371/journal.pone.0024570
- Zorz, J., C. Willis, A. M. Comeau, M. G. I. Langille, C. L. Johnson, W. K. W. Li, and J. LaRoche. 2019. Drivers of Regional Bacterial Community Structure and Diversity in the Northwest Atlantic Ocean. *Front. Microbiol.* **10**: 1–24. doi:10.3389/fmicb.2019.00281
- Zwirgmaier, K., L. Jardillier, M. Ostrowski, and others. 2008. Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. *Environ. Microbiol.* **10**: 147–161. doi:10.1111/j.1462-2920.2007.01440.x

Supplementary 1

Supplementary to Chapter 1

Supplementary 1 A Diagnostic pigments

For pigment analysis we used a high performance liquid chromatography (HPLC) analyser and calculated pigment concentration according to Kiliyas et al. (2013). The relative proportion of phytoplankton functional types (PFTs) to the sum of diagnostic pigments was calculated after Hirata et al. (2011) (Table S1, Fig. S6). We checked for ultraoligotrophic conditions ($\text{chl } a < 0.08 \text{ mg m}^{-3}$; Brewin et al. 2010), however, none of our stations was ultraoligotrophic according to these terms. Relative proportion of each phytoplankton size class (PSC) was calculated using a linear regression model proposed by Uitz et al. (2006). ΣDP is the sum of all diagnostic pigment concentrations and is expressed as the sum of 1.41 [Fuco] (Fucoxanthin) + 1.41 [Perid] (Perdinin) + 1.27 [Hex-fuco] (19'-hexanoyloxyfucoxanthin) + 0.6 [Allo] (Alloxanthin) + 0.35 [But-fuco] (19'-but-fucoxanthin) + 1.01 [TChl-b] (Chlorophyll b) + 0.86 [Zea] (Zeaxanthin). The calculated size classes were microplankton (20–200 μm), nanoplankton (2–20 μm) and picoplankton (0.2–2 μm) (Sieburth et al. 1978). For our data, the difference between total chl *a* and accessory pigments varied in the given range between 7% and 27%, with one outlier in station 37 (51%) because of the high quantification of diatoxanthin (and beta carotenoids). The regression between total chl.a and accessory pigments had a slope of 0.96 and $R^2 > 0.9$. We note that diagnostic pigment analysis has its inaccuracies and may fail on a case-by-case basis (e.g. the size of diatoms can range between the micro and nanoplankton size fraction); for more detail, see Aiken et al. (2009) and Uitz et al. (2006). Nevertheless, this method allows a comprehensive description of the distribution of PFTs and PSCs on global scales.

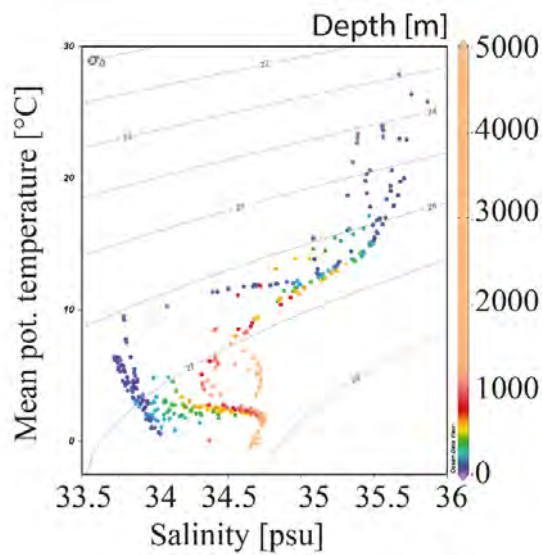


Figure S1. Temperature- Salinity- density plot of all CTD depth profiles. Potential temperature was calculated in Ocean Data View based on measured Temperature and Salinity and reference density of 0.

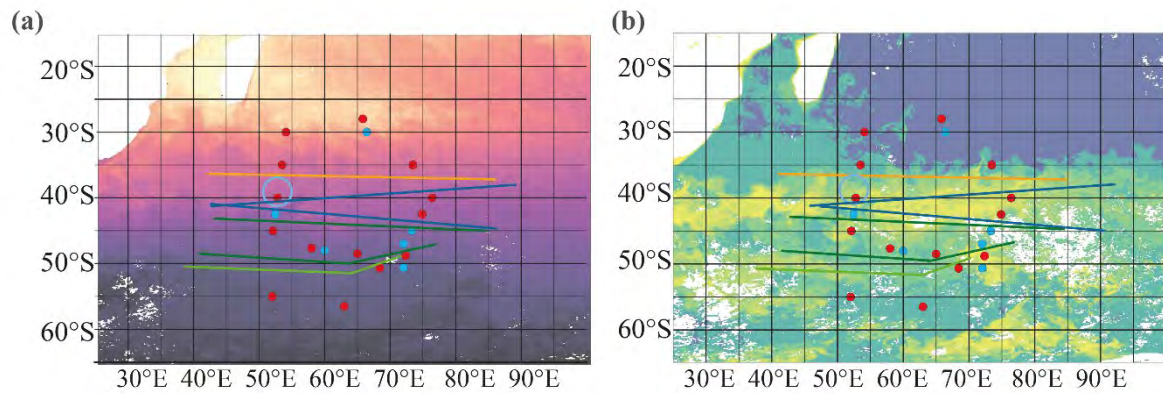


Figure S2. Sea surface temperature (a) and Surface chlorophyll a concentration (b) from MODIS satellite data (<https://neo.sci.gsfc.nasa.gov/>) of the study region in January 2017. Red stations indicate sampling for C- and N- rate measurements and samples for microbial community composition. Blue stations indicate stations with CTD sensor data. Colored lines are drawn according to separation of water masses in this study. The yellow line highlights the southern boundary of the Indian Ocean gyre (ISSG), light blue circle the Subtropical Front (STF), blue line the Subantarctic Front (SAF), green line the Polar Front Zone (PFZ) and the light green line indicates the boundary of the northern Antarctic Zone (AZ).

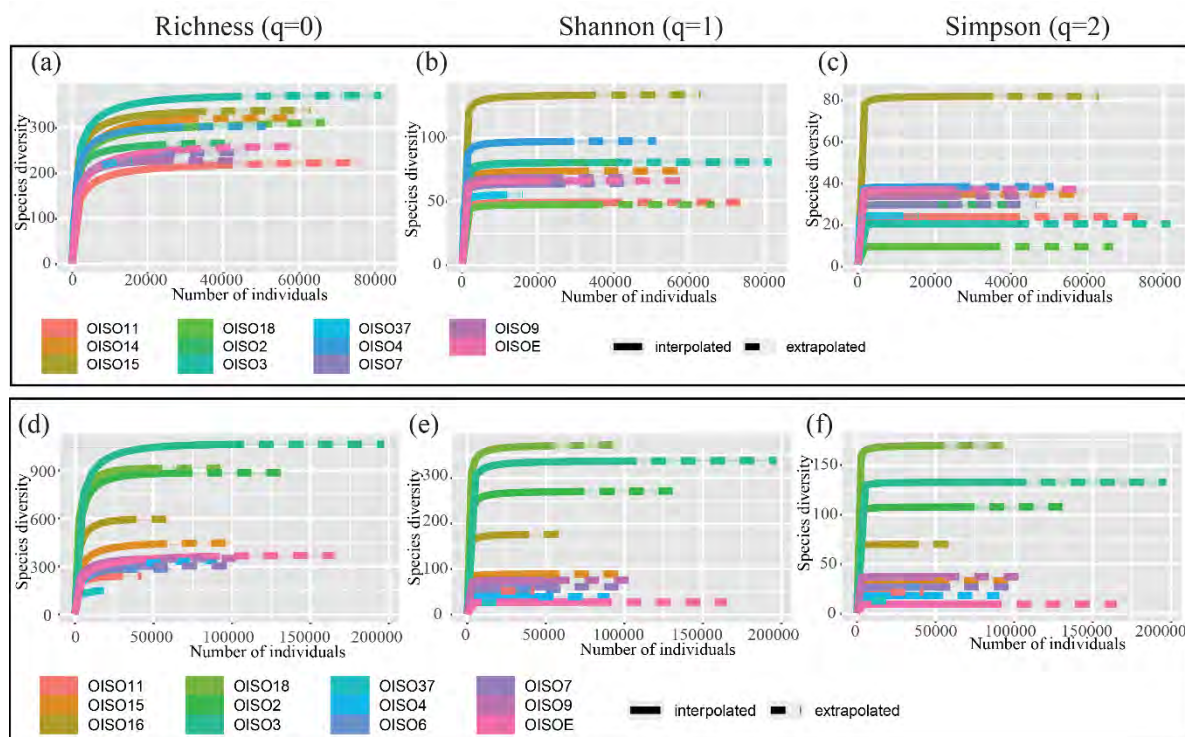


Figure S3. microbial alpha diversity of prokaryotes (a-c) and eukaryotes (d-f). (a) Richness ($q=0$) and extrapolates richness estimate for prokaryotes. (b) Shannon entropy ($q=1$) and extrapolates Shannon estimate for prokaryotes. (c) Inverse Simpson Diversity ($q=2$) and extrapolates Inverse Simpson Diversity estimate for prokaryotes. (d) Richness ($q=0$) and extrapolates richness estimate for eukaryotes. (e) Shannon entropy ($q=1$) and extrapolates Shannon estimate for eukaryotes. (f) Inverse Simpson Diversity ($q=2$) and extrapolates Inverse Simpson Diversity estimate for eukaryotes.

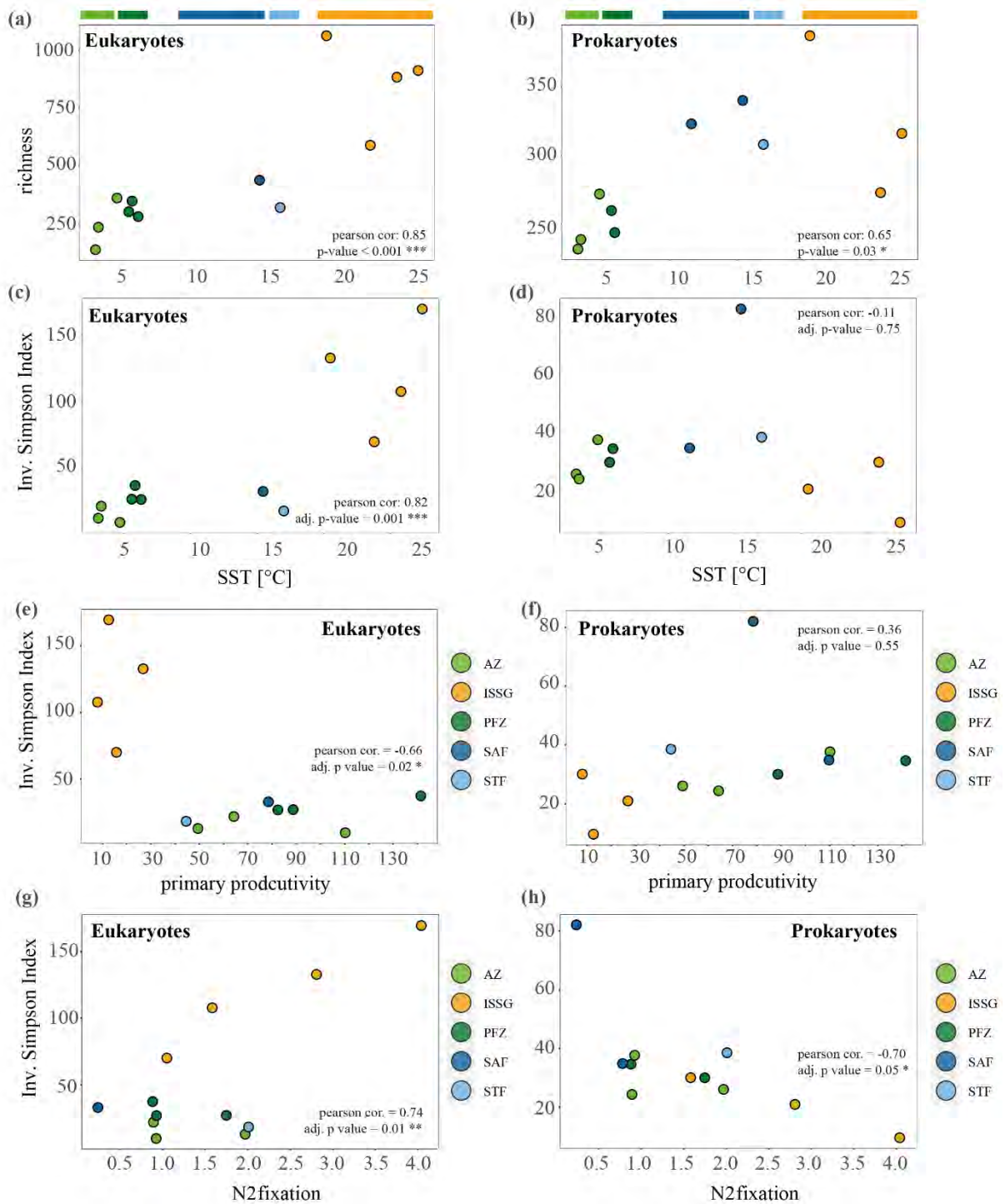


Figure S4. Microbial alpha diversity (a) Eukaryotic ASV richness of 18S amplicon sequences changing over sea surface temperature (SST). (b) Prokaryotic ASV richness of 16S amplicon sequences changing over SST. (c) alpha diversity of eukaryotic ASVs calculated as Inverse Simpson Index against SST. (d) alpha diversity of prokaryotic ASVs calculated as Inverse Simpson Index against SST. (e) Eukaryotic Inverse Simpson Index against primary productivity. (f) Prokaryotic Inverse Simpson Index against primary productivity. (g) Eukaryotic Inverse Simpson Index against nitrogen fixation. (h) Prokaryotic Inverse Simpson Index against nitrogen fixation; Pearson correlation (r) and p-values are indicated in the figures; dark green color indicates the Antarctic Zone (AZ), light green the Polar Front Zone (PFZ), dark blue color Subantarctic Front (SAF), light blue indicates the Subtropical Front (STF) and yellow the Indian South Subtropical Gyre (ISSG).

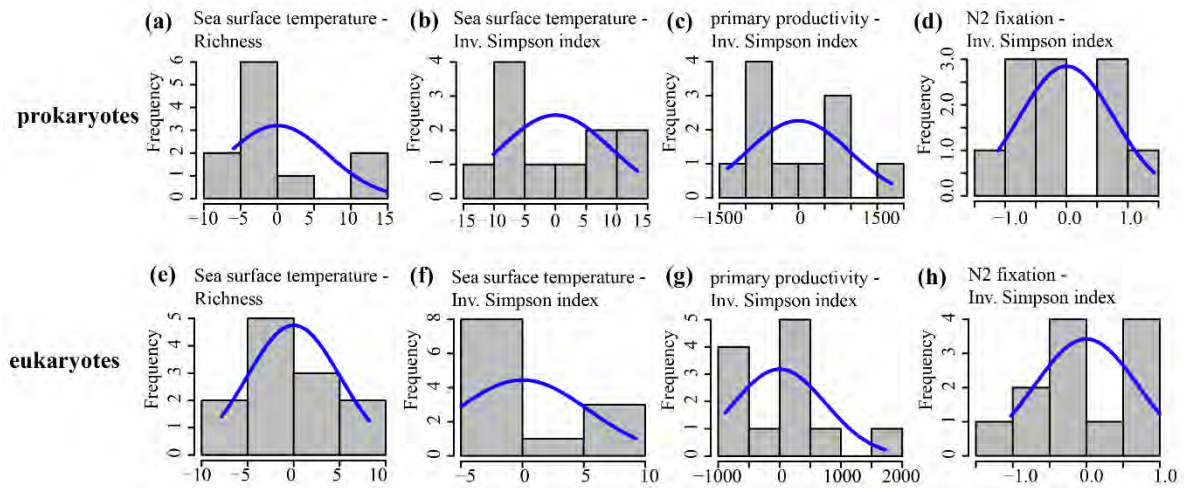


Figure S5. Residual Histograms of Pearson correlation plots from Fig. S4 of prokaryotes (a – d) and eukaryotes (e – h).

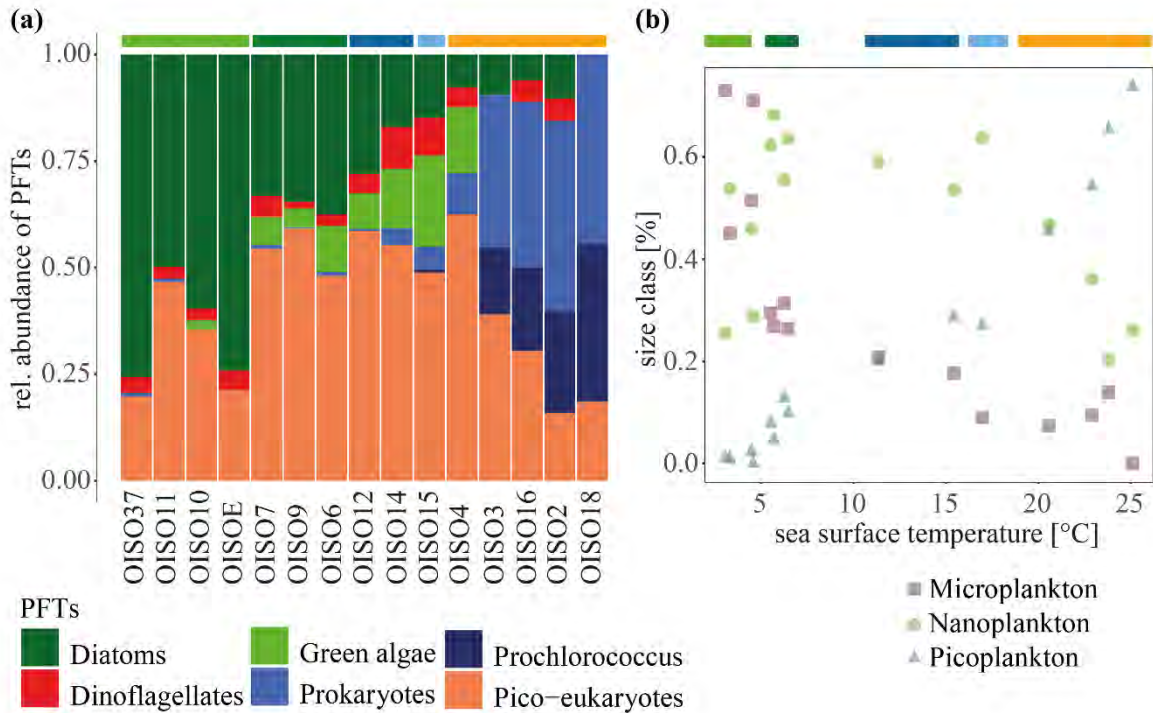


Figure S6. (a) relative abundance of different phytoplankton functional types (PFTs) calculated after Hirata et al. (2011), stations are sorted after increasing temperature. (b) relative abundance of phytoplankton size classes calculated after Uitz et al. 2006 against sea surface temperature. Colored bars indicate water masses according to their sea surface temperature; dark green color indicates the Antarctic Zone (AZ), light green the Polar Front Zone (PFZ), dark blue color Subantarctic Front (SAF), light blue indicates the Subtropical Front (STF) and yellow the Indian South Subtropical Gyre (ISSG).

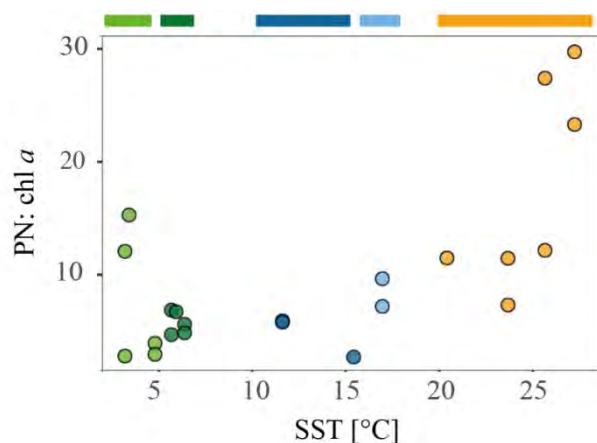
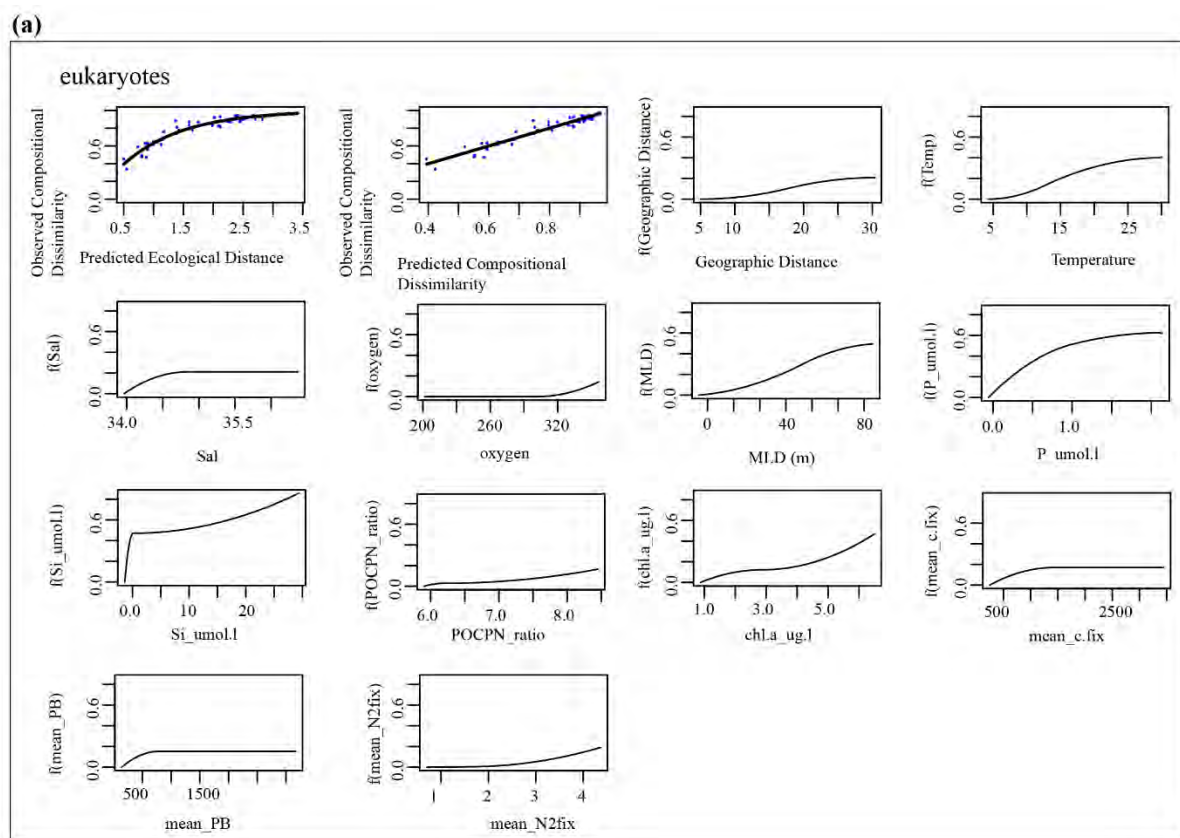


Figure S7. Ratio of particulate organic nitrogen (PN) concentration and chlorophyll *a* (chl *a*) concentration against sea surface temperature. Colored bars indicate water masses according to their sea surface temperature; dark green color indicates the Antarctic Zone (AZ), light green the Polar Front Zone (PFZ), dark blue color Subantarctic Front (SAF), light blue indicates the Subtropical Front (STF) and yellow the Indian South Subtropical Gyre (ISSG).



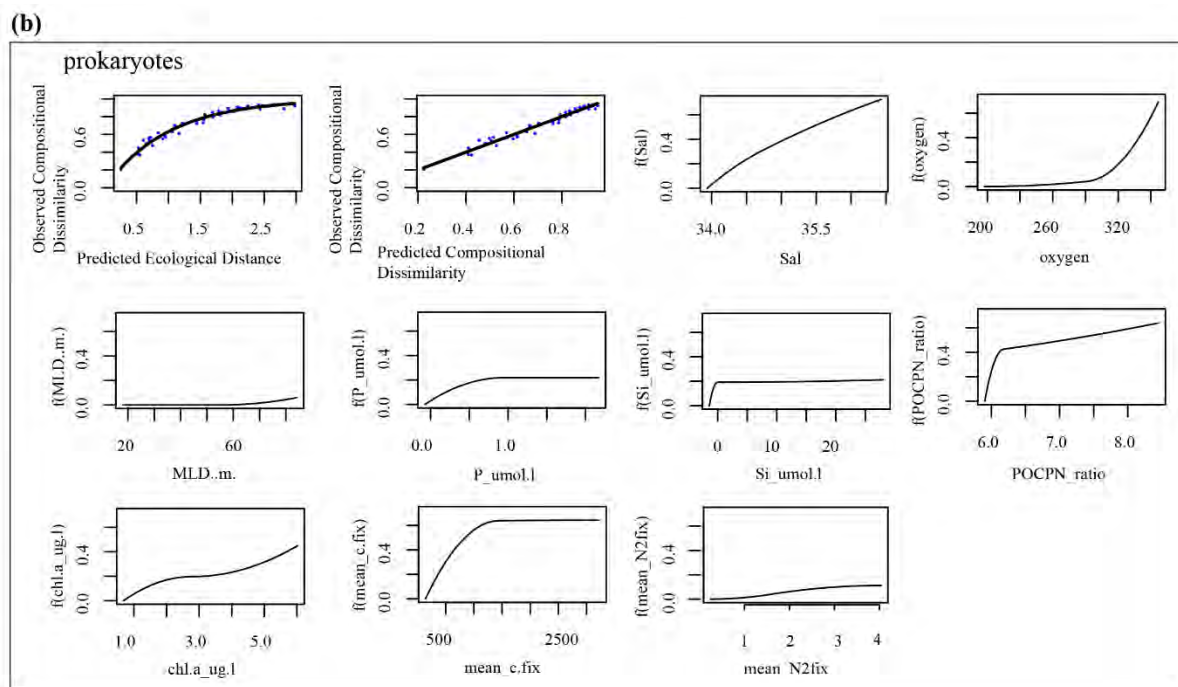


Figure S8. General dissimilarity model (GDM) of (a) eukaryotes and (b) prokaryotes. GDM is used to assess non-linear relationship of microbial community dissimilarity with z-scored environmental parameters (i.e. geographical distance, salinity, mixed layer depth (MLD), temperature, oxygen, nitrate (NO_3), ammonium (NH_4), phosphate (PO_4), silicate (Si), chlorophyll *a*, POC : PN ratio, primary productivity (PP), specific primary productivity (P^B), and N_2 fixation). Displayed is only those plots where environmental parameters contributed to detectable community change. Contribution of each environmental parameter is shown by the magnitude of each spline along the concentration of each environmental variable.

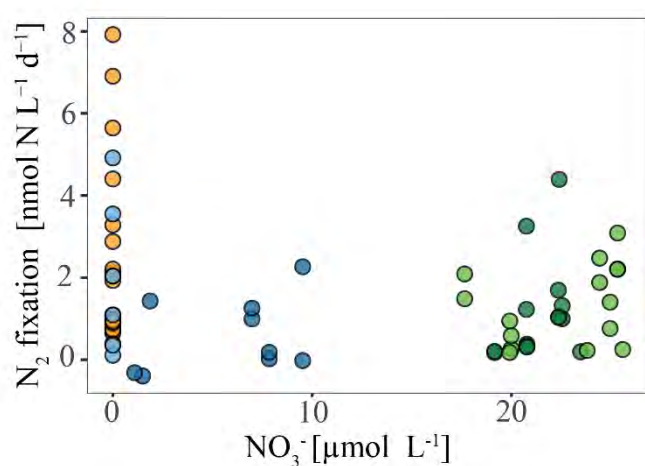


Figure S9. N_2 fixation in $\text{nmol N L}^{-1} \text{d}^{-1}$ against nitrate concentration ($\mu\text{M NO}_3$). Despite highest N_2 fixation at low- NO_3 sites, no clear association between N_2 fixation and NO_3 concentration was found.

Table S1. OISO Formulas used for calculation of phytoplankton functional types and size class fractions based on diagnostic pigments after Hirata et al. (2011) and Uitz et al. (2006).

Reference	Phytoplankton Size Class / Functional Type	Diagnostic Pigments	Estimation Formula
Hirata et al 2011	Diatoms	Fuco	$1.41[\text{Fuco}] / \sum \text{DP}_w$
	Dinoflagellates	Perid	$1.41[\text{Perid}] / \sum \text{DP}_w$
	Green algae	TChl- <i>b</i>	$1.01[\text{Chl-}b] / \sum \text{DP}_w$
	Prokaryotes	Zea	$0.86[\text{Zea}] / \sum \text{DP}_w$
	<i>Prochlorococcus</i>	Divinyl Chlorophyll- <i>a</i> (DV Chl- <i>a</i>)	$0.74[\text{DV Chl-}a] / \sum \text{DP}_w$
	Prymnesiophytes (Haptophytes)	Hex-fuco, But-fuco	Approximated with Nanoplankton – Green algae
	Pico-Eukaryotes	Hex-fuco, Chl- <i>b</i>	Approximated with Picoplankton – Prokaryotes
Uitz et al 2006	Microplankton	Fuco, Perid	$1.41[\text{Fuco} + \text{Perid}] / \sum \text{DP}_w$
	Nanoplankton	Hex-fuco, But-fuco, Allo	$(1.27[\text{Hex-fuco}] + 0.35[\text{But-fuco}] + 0.6[\text{Allo}]) / \sum \text{DP}_w$
	Picoplankton	Zea, Hex-fuco, Chl- <i>b</i>	$(0.86[\text{Zea}] + 1.01[\text{TChl-}b]) / \sum \text{DP}_w$

Table S2. DADA2 pipeline sequencing statistics for 16S and 18S rRNA gene sequences. Numbers show number of reads left in each sample after different steps in the pipeline. Input data is raw input data from Illumina sequencer. Input reads are filtered with filterandtrim() function with minimum quality score of 2 and minimum sequence length of 50bp; denoised (denoisedF are forward reads and denoisedR for reverse reads) with DADA2 default parameters, merged sequences with minimum overlap of 20bp and final step of denovo chimeras removal (method= consensus).

	Station	Input	filtered	denoisedF	denoisedR	merged	nonchim
16S rRNA gene	OISO11_T0	66108	63846	62838	63254	48086	47037
	OISO3_T0	73642	69113	67515	68144	47726	46581
	OISOE_T0	50057	47584	46944	47108	35810	35511
	OISO7_T0	40330	38681	38120	38407	28130	27644
	OISO4_T0	36868	34934	34059	34409	27378	27115
	OISO37_T0	35416	33686	33222	33348	10635	10587
	OISO15_T0	47689	45126	43907	44502	34182	33642
	OISO2_T0	33694	32425	31429	31974	23589	23025
	OISO14_T0	45556	43802	43000	43351	34250	32550
	OISO9_T0	40535	38706	38023	38281	28173	27912
	OISO10_T0	46965	45299	44522	44889	30995	30468
	OISO18_T0	47929	45807	44381	44870	35860	35147
18S rRNA gene	OISO11_T0	36696	25637	25445	25363	24923	24916
	OISO3_T0	180744	127416	125888	125477	117793	116330
	OISOE_T0	179771	128995	128454	128150	107141	107036
	OISO7_T0	103111	69257	68795	68678	60290	60093

OISO6_T0	65038	44344	44031	43900	39237	39116
OISO37_T0	60683	39490	39330	39281	38744	38532
OISO15_T0	78683	56333	55512	55279	52216	50846
OISO2_T0	116014	81494	79958	79546	72577	71621
OISO4_T0	65095	47768	47333	47176	45761	45608

Table S3. SIMPER analysis of the eukaryotic (18S) and prokaryotic (16S) sequence count as relative contribution to the dissimilarity observed between provinces.

18S				
Provinces	Rank	#ASV	Contribution (%)	Taxonomy
SO-ISSG	1	ASV_2	2.10	Eukaryota; Alveolata; Dinoflagellata; Dinophyceae; Dinophyceae_X; Dinophyceae_XX; Dinophyceae_XXX; Dinophyceae_XXX_sp.
	2	ASV_10	2.08	Eukaryota; Hacrobia; Haptophyta; Prymnesiophyceae; Phaeocystales; Phaeocystaceae; Phaeocystis; Phaeocystis_rex
	3	ASV_25	2.04	Eukaryota; Stramenopiles; Ochrophyta; Bacillariophyta; Bacillariophyta_X; Radial-centric-basal-Coscinodiscophyceae; Actinocyclus; Actinocyclus_curvatulus
	4	ASV_31	1.93	Eukaryota; Stramenopiles; Ochrophyta; Bacillariophyta; Bacillariophyta_X; Raphid-pennate; Fragilariopsis
	5	ASV_33	1.86	Eukaryota; Stramenopiles; Ochrophyta; Bacillariophyta; Bacillariophyta_X; Polar-centric-Mediophyceae; Chaetoceros
	6	ASV_134	1.85	Eukaryota; Alveolata; Dinoflagellata; Syndiniales ; Dino-Group-I; Dino-Group-I-Clade-1; Dino-Group-I-Clade-1_X; Dino-Group-I-Clade-1_X_sp.
	7	ASV_18	1.85	Eukaryota; Stramenopiles; Ochrophyta; Bacillariophyta; Bacillariophyta_X; Raphid-pennate.
	8	ASV_67	1.83	Eukaryota; Stramenopiles; Ochrophyta; Bacillariophyta; Bacillariophyta_X; Raphid-pennate; Pseudo-nitzschia; Pseudo-nitzschia_cf_cuspidata
	9	ASV_78	1.81	Unknown Eukaryotes
	10	ASV_26	1.78	Eukaryota; Hacrobia; Haptophyta; Prymnesiophyceae; Phaeocystales; Phaeocystaceae; Phaeocystis; Phaeocystis_rex
SO-SSTC	1	ASV_53	2.56	Eukaryota; Alveolata; Dinoflagellata; Syndiniales; Dino-Group-I; Dino-Group-I-Clade-1; Dino-Group-I-Clade-1_X; Dino-Group-I-Clade-1_X_sp.
	2	ASV_48	2.52	Eukaryota; Alveolata; Dinoflagellata; Dinophyceae; Gonyaulacales; Gonyaulacaceae; Gonyaulax
	3	ASV_31	2.49	Eukaryota; Stramenopiles; Ochrophyta; Bacillariophyta; Bacillariophyta_X; Raphid-pennate; Fragilariopsis
	4	ASV_11	2.47	Eukaryota; Alveolata; Dinoflagellata; Syndiniales; Dino-Group-I; Dino-Group-I-Clade-1; Dino-Group-I-Clade-1_X; Dino-Group-I-Clade-1_X_sp
	5	ASV_22	2.38	Eukaryota; Alveolata; Dinoflagellata; Dinophyceae; Dinophyceae_X; Dinophyceae_XX; Dinophyceae_XXX; Dinophyceae_XXX_sp.
	6	ASV_50	2.35	Eukaryota; Alveolata; Dinoflagellata; Dinophyceae; Dinophyceae_X; Dinophyceae_XX; Dinophyceae_XXX; Dinophyceae_XXX_sp.
	7	ASV_51	2.31	Eukaryota; Alveolata; Dinoflagellata; Dinophyceae; Dinophyceae_X; Dinophyceae_XX; Dinophyceae_XXX; Dinophyceae_XXX_sp.
	8	ASV_56	2.31	Eukaryota; Alveolata; Dinoflagellata; Dinophyceae; Dinophyceae_X; Dinophyceae_XX; Dinophyceae_XXX; Dinophyceae_XXX_sp.
	9	ASV_25	2.30	Eukaryota; Stramenopiles; Ochrophyta; Bacillariophyta; Bacillariophyta_X; Radial-centric-basal-Coscinodiscophyceae; Actinocyclus; Actinocyclus_curvatulus
	10	ASV_94	2.22	Eukaryota; Archaeplastida; Chlorophyta; Chloropicophyceae; Chloropicales; Chloropicaceae; Chloroparvula; Chloroparvula_pacifica
SSTC-ISSG	1	ASV_94	1.63	Eukaryota; Archaeplastida; Chlorophyta; Chloropicophyceae; Chloropicales; Chloropicaceae; Chloroparvula; Chloroparvula_pacifica
	2	ASV_170	1.62	Eukaryota; Alveolata; Dinoflagellata; Dinophyceae; Dinophyceae_X; Dinophyceae_XX; Dinophyceae_XXX; Dinophyceae_XXX_sp.

	3	ASV_176	1.48	Unknown Eukaryotes
	4	ASV_375	1.48	Eukaryota; Alveolata; Dinoflagellata; Dinophyceae; Gonyaulacales; Pyrophacaceae; Fragilidium
	5	ASV_162	1.47	Eukaryota; Alveolata; Dinoflagellata; Dinophyceae; Dinophyceae_X; Dinophyceae_XX; Dinophyceae_XXX; Dinophyceae_XXX_sp.
	6	ASV_23	1.44	Eukaryota; Archaeplastida; Chlorophyta; Chloropicophyceae; Chloropicales; Chloropicaceae; Chloroparvula; Chloroparvula_pacifica
	7	ASV_505	1.39	Eukaryota; Alveolata; Dinoflagellata; Dinophyceae; Dinophyceae_X; Dinophyceae_XX; Dinophyceae_XXX; Dinophyceae_XXX_sp.
	8	ASV_5	1.33	Eukaryota; Archaeplastida; Chlorophyta; Palmophyllophyceae; Prasinococcales; Prasinococcales-Clade-B; Prasinoderma; Prasinoderma_singularis
	9	ASV_258	1.33	Eukaryota; Alveolata; Dinoflagellata; Syndiniales; Dino-Group-II; Dino-Group-II-Clade-10-and-11; Dino-Group-II-Clade-10-and-11_X Dino-Group-II-Clade-10-and-11_X_sp.
	10	ASV_228	1.30	Eukaryota; Alveolata; Dinoflagellata; Dinophyceae
16S				
ISSG-SO	1	ASV_9	2.06	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; Aquibacter
	2	ASV_15	2.02	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; NS5 marine group
	3	ASV_11	1.97	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Planktomarina
	4	ASV_26	1.89	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Planktomarina
	5	ASV_1	1.87	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; Ulvibacter
	6	ASV_23	1.87	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; NS4 marine group
	7	ASV_35	1.85	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; NS4 marine group
	8	ASV_7	1.85	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; NS2b marine group
	9	ASV_18	1.83	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Planktomarina
	10	ASV_16	1.73	Bacteria; Proteobacteria; Gammaproteobacteria; Cellvibrionales; Porticoccaceae; SAR92 clade
SSTC-SO	1	ASV_12	2.56	Bacteria; Proteobacteria; Gammaproteobacteria; Thiomicrospirales; Thioglobaceae; SUP05 cluster
	2	ASV_26	2.55	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Planktomarina
	3	ASV_35	2.51	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; NS4 marine group
	4	ASV_6	2.32	Bacteria; Cyanobacteria; Oxyphotobacteria; Synechococcales; Cyanobiaceae; Synechococcus CC9902
	5	ASV_31	2.23	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; NS9 marine group
	6	ASV_16	2.21	Bacteria; Proteobacteria; Gammaproteobacteria; Cellvibrionales; Porticoccaceae; SAR92 clade
	7	ASV_65	2.21	Bacteria; Proteobacteria; Gammaproteobacteria; Oceanospirillales; Nitrincolaceae
	8	ASV_39	2.17	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Planktomarina
	9	ASV_15	2.16	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; NS5 marine group
	10	ASV_91	2.09	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Cryomorphaceae; NS10 marine group
ISSG-SSTC	1	ASV_24	1.40	Bacteria; Cyanobacteria; Oxyphotobacteria; Synechococcales; Cyanobiaceae; Prochlorococcus MIT9313

	2	ASV_6	1.32	Bacteria; Cyanobacteria; Oxyphotobacteria; Synechococcales; Cyanobiaceae; Synechococcus CC9902
	3	ASV_105	1.31	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae
	4	ASV_122	1.30	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; NS5 marine group
	5	ASV_25	1.27	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; NS4 marine group
	6	ASV_2	1.27	Bacteria; Cyanobacteria; Oxyphotobacteria; Synechococcales; Cyanobiaceae; Prochlorococcus MIT9313
	7	ASV_146	1.25	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; NS5 marine group
	8	ASV_94	1.25	Bacteria; Proteobacteria; Alphaproteobacteria; Puniceispirillales; SAR116 clade
	9	ASV_139	1.21	Bacteria; Proteobacteria; Gammaproteobacteria; SAR86 clade
	10	ASV_7	1.20	Bacteria; Bacteroidetes; Bacteroidia; Flavobacteriales; Flavobacteriaceae; NS2b marine group

Table S4. Upset plots visualize the intersections of multiple sets in a comprehensive and approachable way such that one can see intersection set size of individual sets and intersection size of common samples between sets. The calculations and plots are done based on binary tables (presence/ absence) of samples merged within their provinces. Thus, the analysis is biased by unequal sample number and sequencing depth between samples. For instance, species A is found in sample 1 of the Southern Ocean but not in sample 2 and 3. In the merged sample it would be accounted as “present” for the Southern Ocean, i.e. simplifying the provinces to one homogenous sample. To provide an estimate how well each sample is represented by the merged sample per province we provide in the following table for each sample and its representative with the following categories: 0 = shared absence (not accounted in further analysis), 1 = present in other samples of the same province but absent in sample of interest, 2 = shared between sample of interest and representative sample.

Sample	0	1	2	Percent shared (%)
Eukaryotes (18S)				
OISO37 (and SO)	1923	439	139	24%
OISO11 (and SO)	1923	342	236	41%
OISOE (and SO)	1923	216	362	63%
OISO7 (and SO)	1923	275	303	52%
OSIO9 (and SO)	1923	229	349	60%
OISO4 (and SSTC)	1917	263	321	55%
OISO15 (and SSTC)	1917	145	439	75%
OISO3 (and ISSG)	586	852	1063	56%
OISO2 (and ISSG)	586	1031	884	46%
OISO16 (and ISSG)	586	1325	590	31%
OISO18 (and ISSG)	586	1002	913	48%
Prokaryotes (16S)				
OISO37 (and SO)	551	258	232	47%
OISO11 (and SO)	551	251	239	49%
OISOE (and SO)	551	218	272	56%
OISO7 (and SO)	551	230	260	53%
OISO9 (and SO)	551	246	244	50%
OISO10 (and SO)	551	224	266	54%
OISO4 (and SSTC)	508	225	308	58%
OISO14 (and SSTC)	508	210	323	61%
OISO15 (and SSTC)	508	193	340	64%
OISO3 (and ISSG)	455	199	387	66%
OISO2 (and ISSG)	455	313	273	47%
OISO18 (and ISSG)	455	270	316	54%

References Supplementary 1 A

- Aiken, J., Y. Pradhan, R. Barlow, S. Lavender, A. Poulton, P. Holligan, and N. Hardman-Mountford. 2009. Phytoplankton pigments and functional types in the Atlantic Ocean: A decadal assessment, 1995-2005. *Deep. Res. Part II Top. Stud. Oceanogr.* **56**: 899–917. doi:10.1016/j.dsr2.2008.09.017
- Brewin, R. J. W., S. Sathyendranath, T. Hirata, S. J. Lavender, R. M. Barciela, and N. J. Hardman-mountford. 2010. A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecol. Modell.* **221**: 1472–1483. doi:10.1016/j.ecolmodel.2010.02.014
- Hirata, T., R. J. W. Brewin, J. Aiken, R. Barlow, K. Suzuki, and T. Isada. 2011. Synoptic relationships between surface Chlorophyll- a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences* **8**: 311–327. doi:10.5194/bg-8-311-2011
- Kilias, E., C. Wolf, E. M. Nöthig, I. Peeken, and K. Metfies. 2013. Protist distribution in the Western Fram Strait in summer 2010 based on 454-pyrosequencing of 18S rDNA. *J. Phycol.* **49**: 996–1010. doi:10.1111/jpy.12109
- Sieburth, J. M., V. Smetacek, and J. Lenz. 1978. Pelagic ecosystem structure : Heterotrophic compartments of the plankton and their relationship to plankton size fractions. *Limnol. Oceanogr.* **23**: 1256–1263.
- Uitz, J., H. Claustre, A. Morel, and S. B. Hooker. 2006. Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. *J. Geophys. Res. Ocean.* **111**. doi:10.1029/2005JC003207

Supplementary 2

Supplementary to Chapter 2

S1. Definition of oceanic provinces

Physico-chemical characteristics of ocean provinces were different across a latitudinal gradient and influenced by local ocean dynamics. In the following we briefly describe these differences for each province. The Southwest Atlantic Shelves province (FKLD) is the only province that extends across both continental shelf, and regions of high eddy dynamics at the shelf break (Fig. S3a). The Brazilian Current Coastal (BRAZ) province largely encompasses the Brazilian Current (see Fig. S2). The BRAZ province was associated with active eddy formation, resulting in a large range of temperature and salinity profiles (Fig. S1; Table C2.1), elevated chl *a* concentrations, and primary productivity (Fig. C2.1, Table C2.1). The North Atlantic Subtropical Gyral (NAST-E) province includes the Azores current, which joins the Canary Current south of the province and the flow towards the strait of Gibraltar (Fig. S2a) but was significantly less productive than the adjacent CNRY province (Two sample t-test $t = -3.4$, $p < 0.01$). In comparison to the NAST-E province with temperatures around 19.19 ± 1.9 °C (mean \pm stdev, $n = 12$), we measured colder temperatures ranging from 15.32 °C to 17.07 °C in the North Atlantic Drift (NADR) province. The Western Tropical Atlantic (WTRA) province encompasses the boundaries to the southern and northern subtropical gyres and the equatorial current with sea surface temperatures 27.3–28.7 °C. The North Atlantic Tropical Gyral (NATR) province had relatively high primary productivity rates ($27.1\text{--}86.3$ $\mu\text{mol C L}^{-1} \text{ h}^{-1}$; Table C2.1), despite low biomass concentrations ($0.12\text{--}0.32$ mg C m^{-3}) in previous studies (Longhurst 2007). We consider FKLD, BRAZ, CNRY, NAST-E, and NADR “high chl *a* provinces”; H-CHL (Table C2.1). We note that the higher surface chl-*a* concentrations can vary seasonally and regionally due to current flows and eddy kinetics (1). Thus, we consider the NAST-E province a H-CHL province despite relatively low NO₃, POC, PN, and chl *a* concentrations measured.

On the other hand, the SATL-COLD, SATL-HOT, WTRA and NATR are considered “low chl *a* provinces” with significantly lower PP ($t = -3.85$, $p < 0.001$, $n_1 = 39$, $n_2 = 38$), chl *a* ($t = -3.86$, $p < 0.001$, $n_1 = 39$, $n_2 = 38$), and NO₃ ($t = -3.89$, $p < 0.001$, $n_1 = 39$, $n_2 = 38$; Table C2.1). An exception was found in the North equatorial counter current $\sim 10\text{--}12^\circ\text{N}$, which falls in the NATR province in our dataset. Thus, the samples in the NATR province had NO₃, PN and POC concentrations above the expected oligotrophic conditions (0.08 $\mu\text{mol L}^{-1}$ NO₃).

S2. Dissolved inorganic nutrients and particulate organic matter

Dissolved inorganic nutrients

At each station, we sampled 12 ml of filtered seawater through a 0.22 µm syringe filter (28mm, Sartorius, Germany) for dissolved inorganic nutrient analyses and stored at −80°C until further analysis. Dissolved inorganic nutrients were assayed on an autoanalyzer (Alliance; Evolution III). Detection limits for silicate (H₄SiO₄), phosphate (PO₄), nitrate (NO₃), and nitrite (NO₂) were 0.2 µM, 0.01 µM, 0.01 µM, and 0.015 µM, respectively.

Particulate organic matter

Duplicate samples (4 L) for particulate organic matter (POM) were immediately (T₀) filtered after sampling and stored at −80°C until further analysis. Total carbon and nitrogen elemental analyses were performed on an elemental analyzer (EUROEA3000; EuroVector, Italy).

S3. Experimental design of primary productivity measurements

We took seawater from the ship's underway system in a 20 L canister for primary productivity (PP) experiments. Water in the canister was homogenized by manual shaking (~2 min) and divided between three 1 L polycarbonate bottles. We initiated experiments by spiking each bottle with 200 µmol NaH¹³CO₃ prior to incubation for 24 h. In addition, we incubated seawater in two 1 L bottles at each station without any addition of stable isotopes for 24 h to provide readings of the natural abundance of stable isotopes in samples for PP calculations. Neutral density screens (present on the ship) were used to simulate the light attenuation. Temperature and light attenuation in incubation bins was continuously tracked during the expedition.

S4. Flow Cytometry measurements

Measurements were taken for autofluorescence (FL2-A) and total cell count using SYBR (FL1). Simultaneously, measurements were collected for forward scatter (FSC; roughly equivalent to size), and side scatter (SSC; indicator of granularity). Cell size and detection were calibrated prior to analysis using 2µm ø beads (RCP-20-5 Rainbow Calibration Particles Lot# GKAJ01, Spherotech).

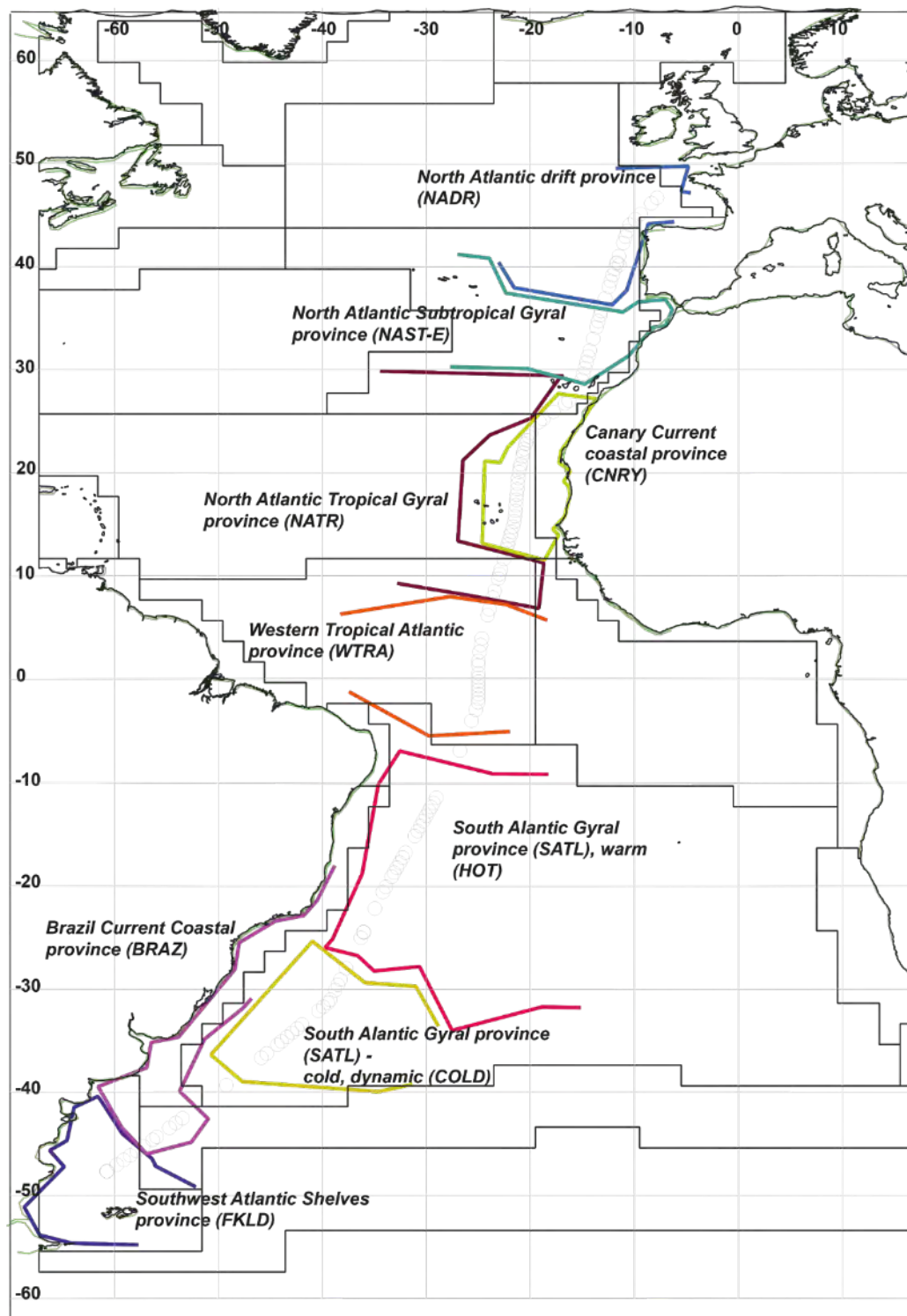


Figure S1. map of PS113 expedition and defined provinces overlayed by Longhurst provinces (Flanders Marine Institute (2009). Longhurst Provinces version 4. Available online at <https://www.marineregions.org/>).

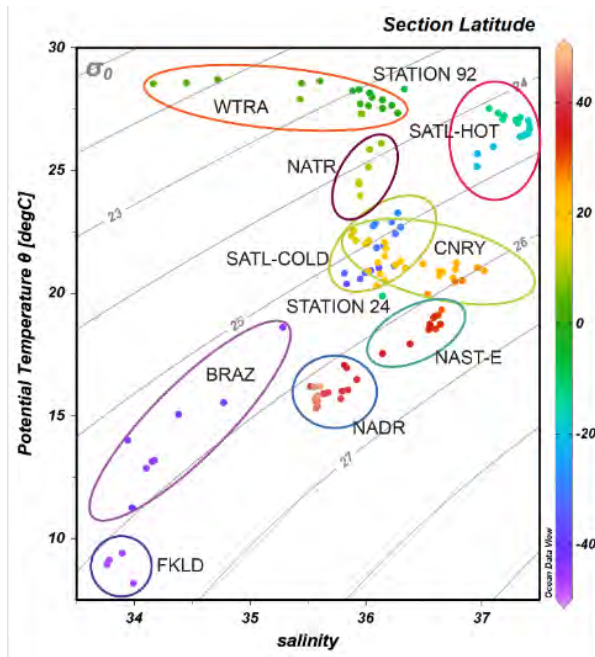


Figure S2. Potential temperature vs salinity from all surface samples. Samples are clustered into oceanic provinces and named according to provinces defined by Longhurst (2007): dark blue circle Falkland province (FKLD); yellow Brazilian province (BRAZ), green South Atlantic tropical province-subdivided into a cold and dynamic part (light green; SATL-COLD) and a homogenous, warm region (dark green; SATL-HOT), orange marks the West Tropical region (WTRA), dark red the North Atlantic tropical region (NATR), yellow-green the Canary region (CNRY), purple the North Atlantic Subtropical eastern province (NAST-E) and blue the North Atlantic drift region (NADR). Potential temperature was calculated in Ocean data view: Potential temperature referenced to reference pressure (0) at 11m depth with sea surface temperature and salinity measured by the ship-board temperature-salinograph.

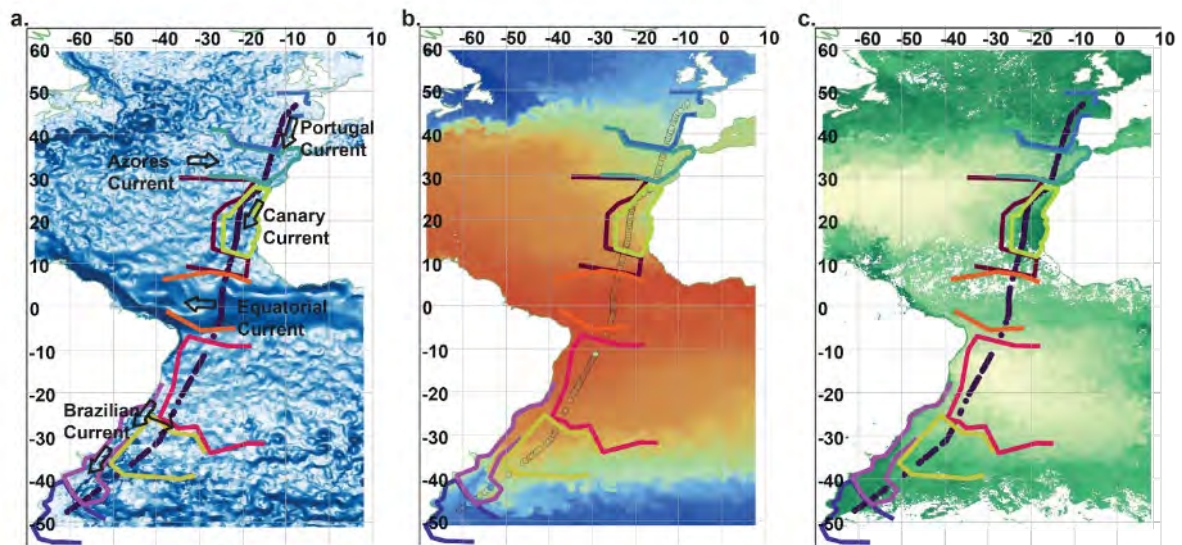


Figure S3. Satellite-derived analysis of ocean provinces taking (a) geostrophic currents, (b) temperature regimes and (c) chl *a* concentration into account. Satellite observations of (a) and (b) were generated using E.U. Copernicus Marine Service Information; monthly averaged from 20180511T00:00Z - 20180608T00:00Z for and NASA MODIS data MY1DMM_CHLORA from May 2018 for chl *a*. The major currents are indicated by arrows and labelled. east and northward component direction calculated with $(180/3.14) * \text{Atn2}([u],[v])$ and velocity with $\text{Sqr}([u]**2 + [v]**2)$ (Butler 2013; <https://www.esri.com/arcgis-blog/products/product/analytics/displaying-speed-and-direction-symbology-from-u-and-v-vectors/>)

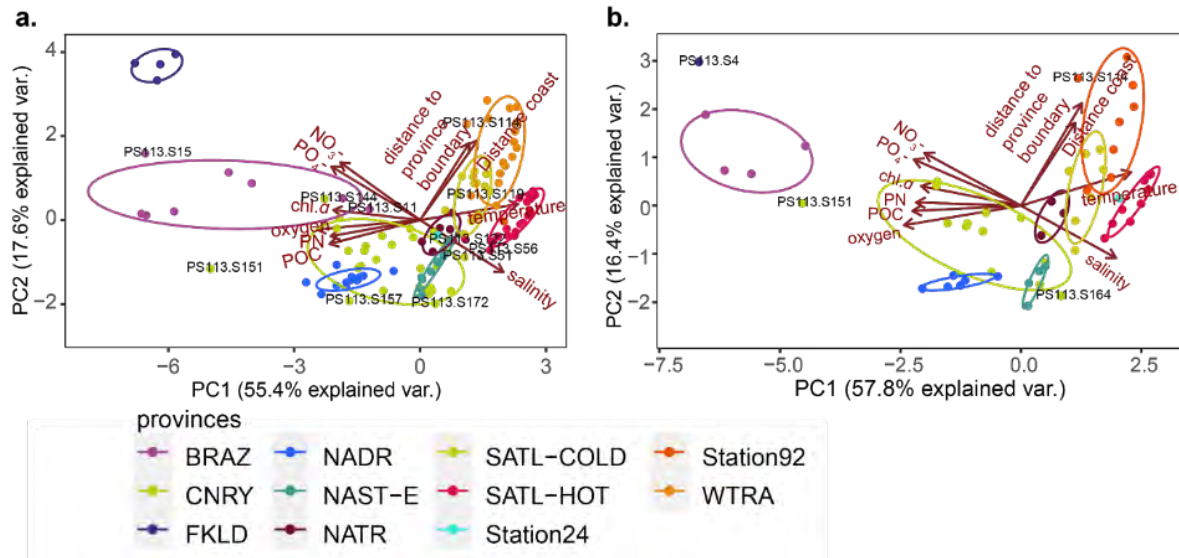


Figure S4. Principal Coordinate Analysis (PCA) of environmental data of (a) the entire transect and (b) the subset of samples where primary productivity (PP) was measured. Sites are colored according to ocean provinces and metadata as explanatory variables.

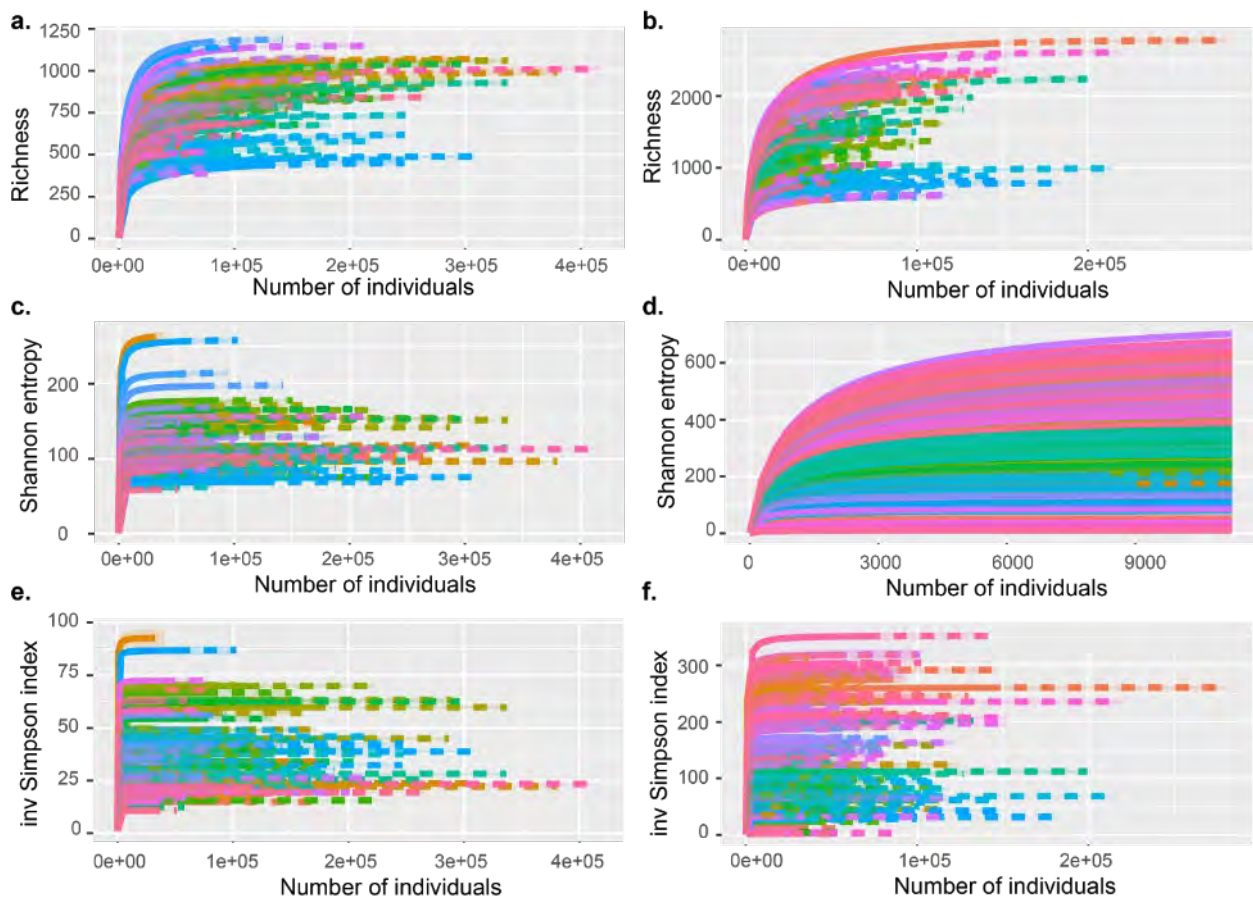


Figure S5. Rarefaction curves of samples of (a) 16S rRNA gene (i.e. prokaryotes) richness ($q = 0$), (b) 18S rRNA gene (i.e. eukaryotes) richness ($q = 0$), (c) 16S rRNA gene sample Shannon entropy ($q = 1$), (d) 18S rRNA gene sample Shannon entropy ($q = 1$) with endpoint set to 11256 (number of ASVs, due to computational power), (e) 16S rRNA gene inverse Simpson index ($q = 2$) (f) 18S rRNA gene sample inverse Simpson index ($q = 2$). All samples were levelling off.

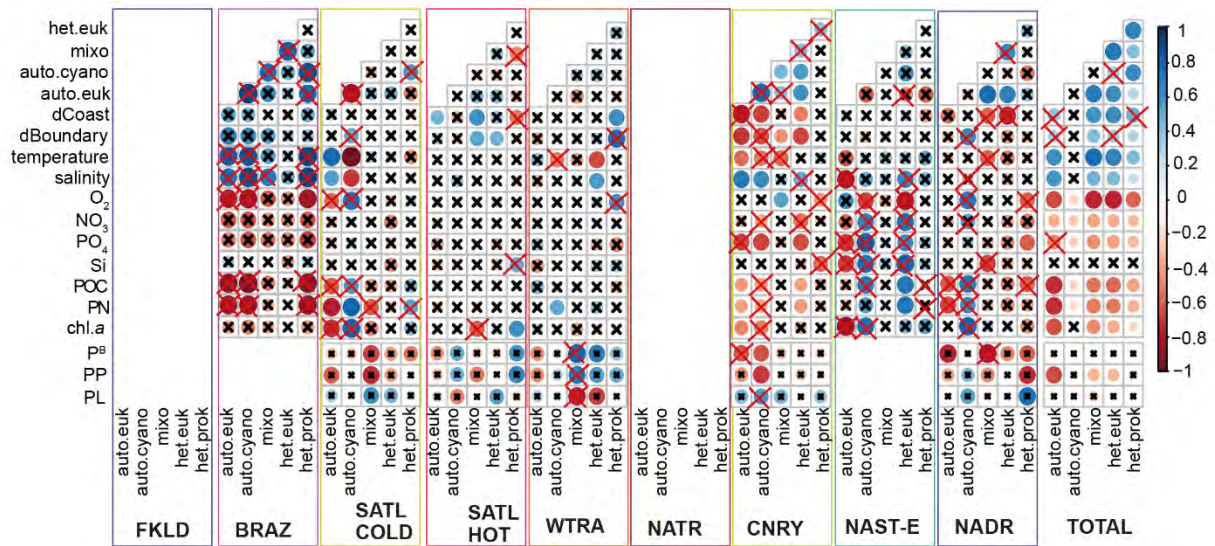


Figure S6. Alpha diversity across ocean provinces. Correlations (Pearson) of Shannon diversity with each other and environmental parameters within ocean provinces. Distance to coast (dCoast) calculated as distance of sample to next shore (in km), dBoundary is the distance to the province boundary of the sample identified to belong to, temperature is sea surface temperature ($^{\circ}\text{C}$), salinity is sea surface salinity, oxygen (μM), nitrate (NO_3) in μM , phosphate (PO_4) in μM , Silicate (Si) in μM , Particulate organic Carbon concentration (POC) in nM , particulate nitrogen (PN) concentration in nM , chlorophyll *a* (chl *a*) in mg m^{-3} , specific primary productivity (P^B) in $\text{nmol C L}^{-1} \text{ chl.}a^{-1}$, primary productivity (PP) in $\text{nmol L}^{-1} \text{ h}^{-1}$ correlation plots with colors indicating gradient from negative (red) to positive (blue) correlation; empty fields indicate “not significant”, i.e. $p\text{-value} > 0.05$, red crosses indicate non-normal distribution of residuals in linear regression model. Within provinces, correlations have been calculated for eukaryotes and prokaryotes, respectively, and each trophic group (autotroph, mixotroph and heterotroph) against different environmental parameters and against each other. Colored boxes indicate correlations within provinces: FKLD, Falkland; BRAZ, Brazilian; SATL, South Atlantic Subtropical Gyre; WTRA, Equator; NATR, North Atlantic Tropical; CNRY, Canary, NAST, North Atlantic Subtropical; NADR, North Atlantic Drift region.

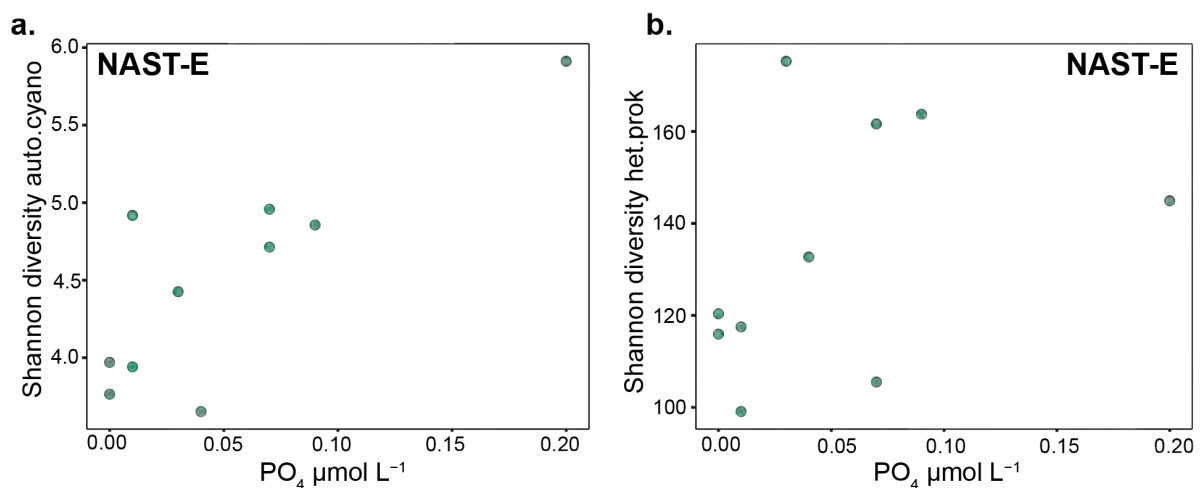


Figure S7. Shannon diversity of cyanobacteria (a) and heterotrophic prokaryotes (b) in the NAST-E province. Correlation between cyanobacteria diversity and phosphate concentration was significantly correlated (Pearson correlation, $r = 0.83$, $p = 0.01$, $n = 10$) while heterotrophic prokaryotic diversity had a positive trend with increasing phosphate concentration, but was not significant (Pearson correlation, $r = 0.39$, $p = 0.27$, $n = 10$).

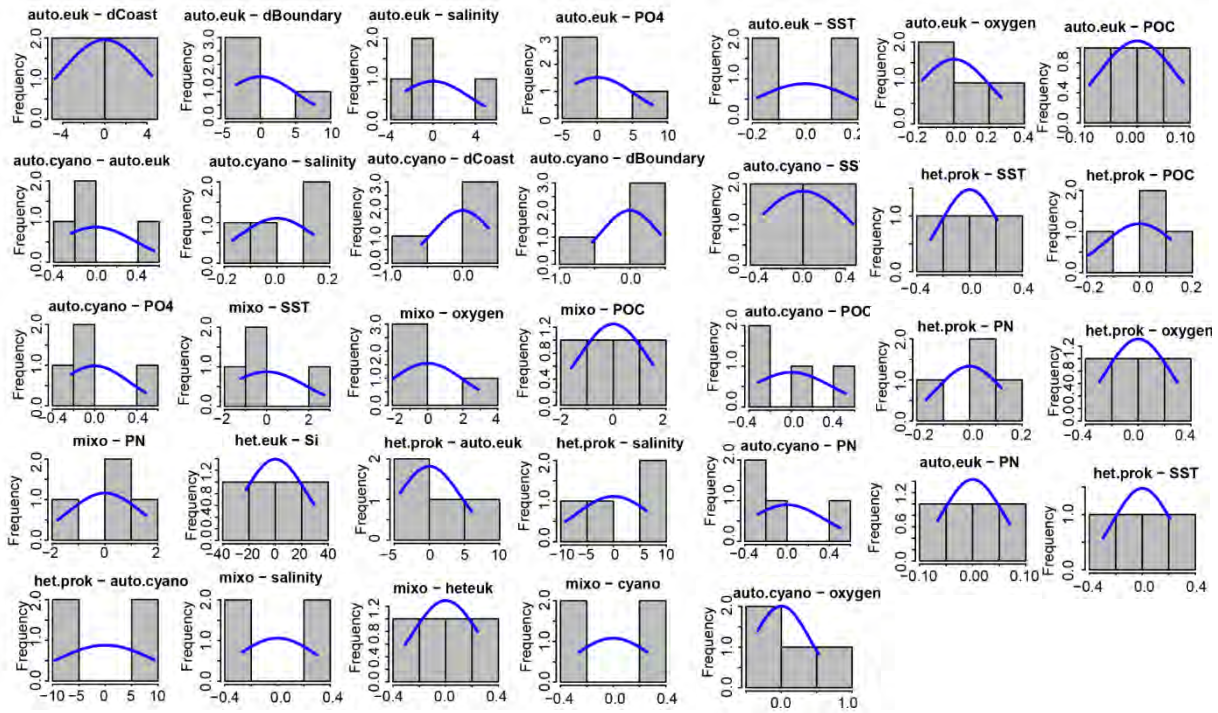


Figure S8. Residual Histograms Brazilian province (n = 8).

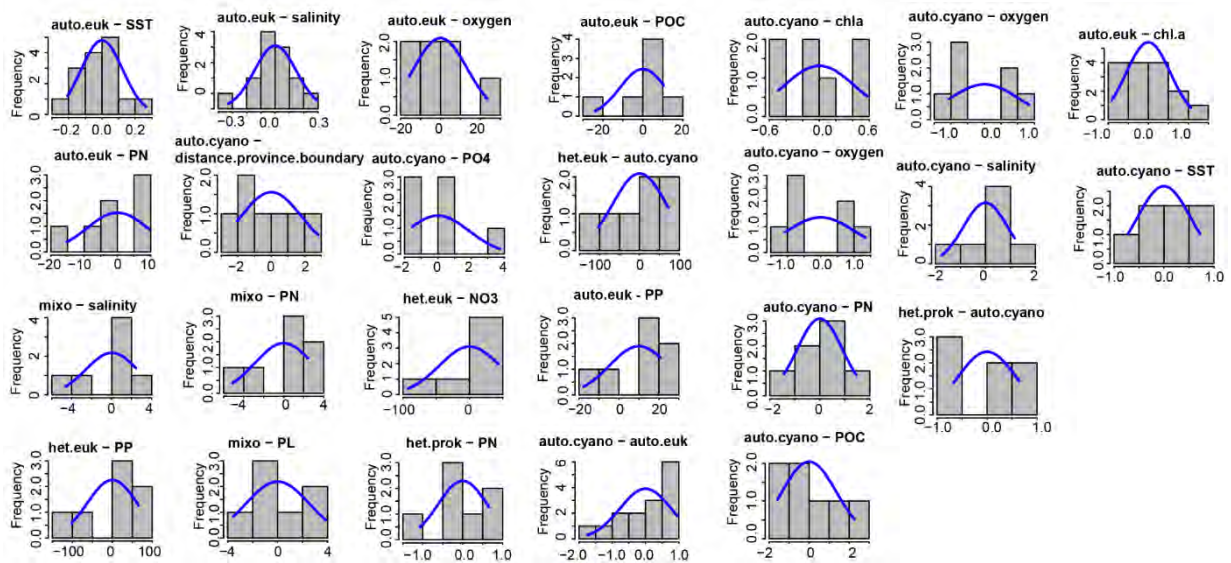


Figure S9. Residual Histograms SATL-COLD province (n = 15).

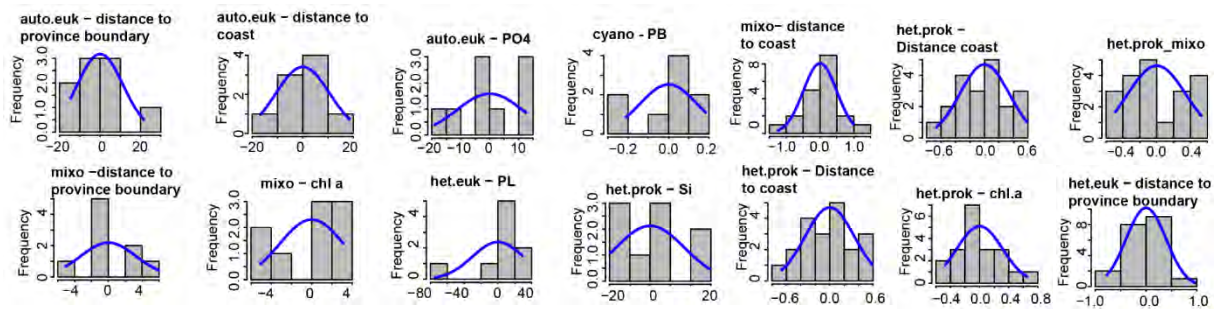


Figure S10. Residual Histograms SATL-HOT province (n = 20).

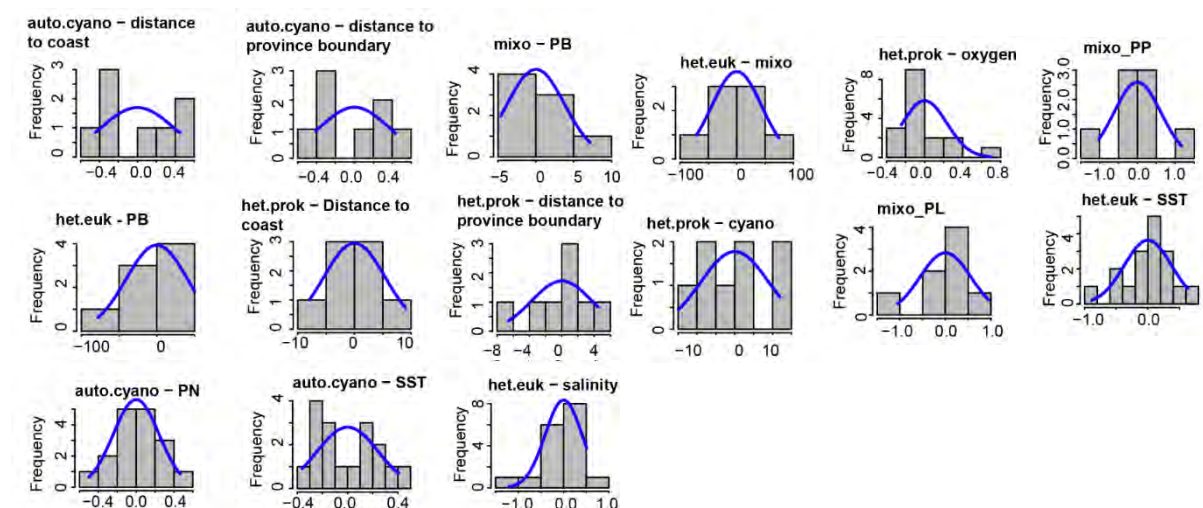


Figure S11. Residual Histogram WTRA province (n = 17).

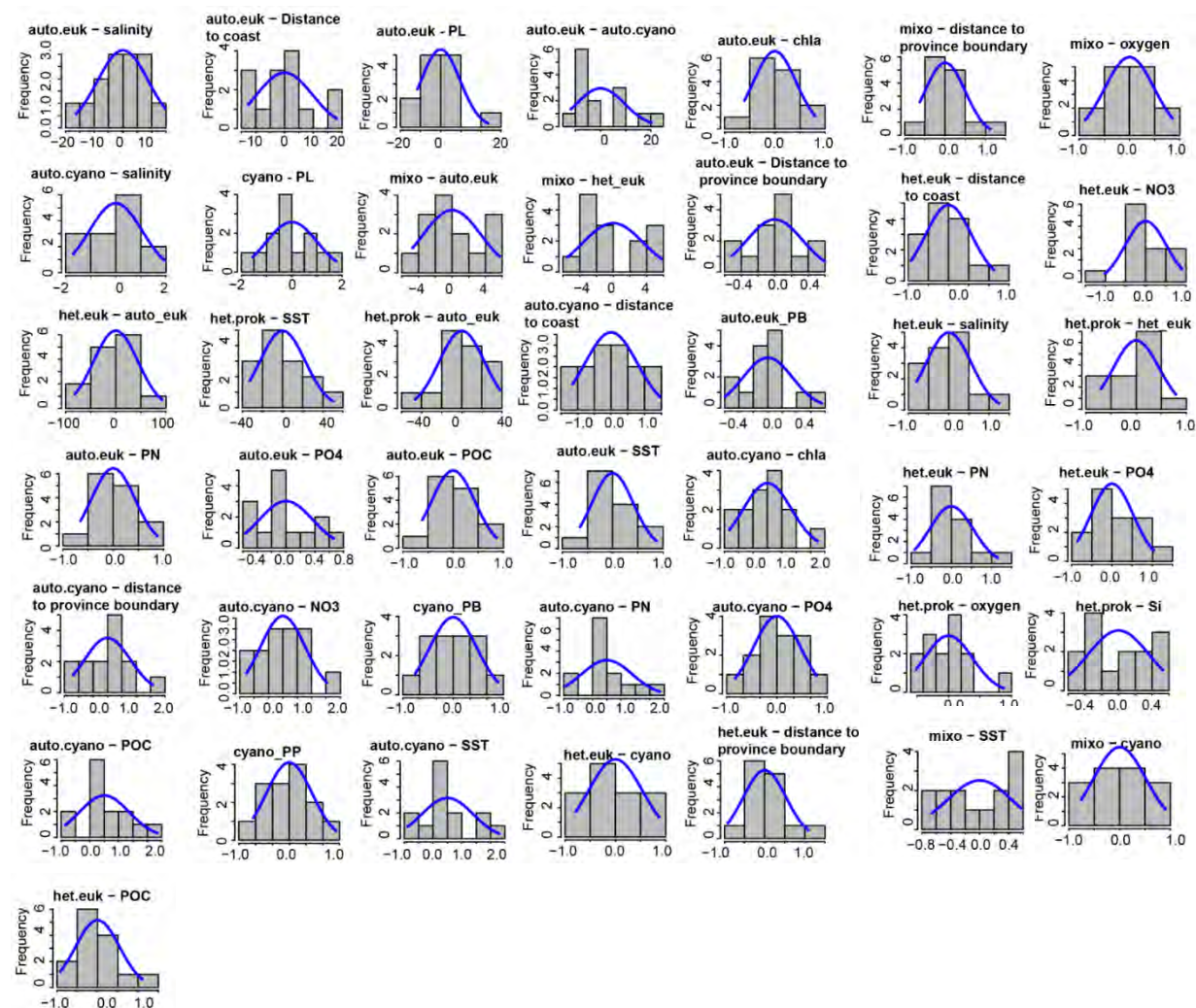


Figure S12. Residual Histogram CNRY province (n = 26).

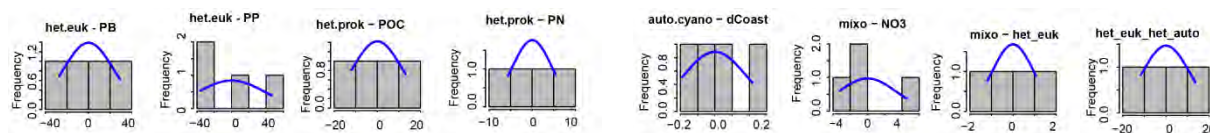


Figure S13. Residual Histograms NAST-E province (n = 10).

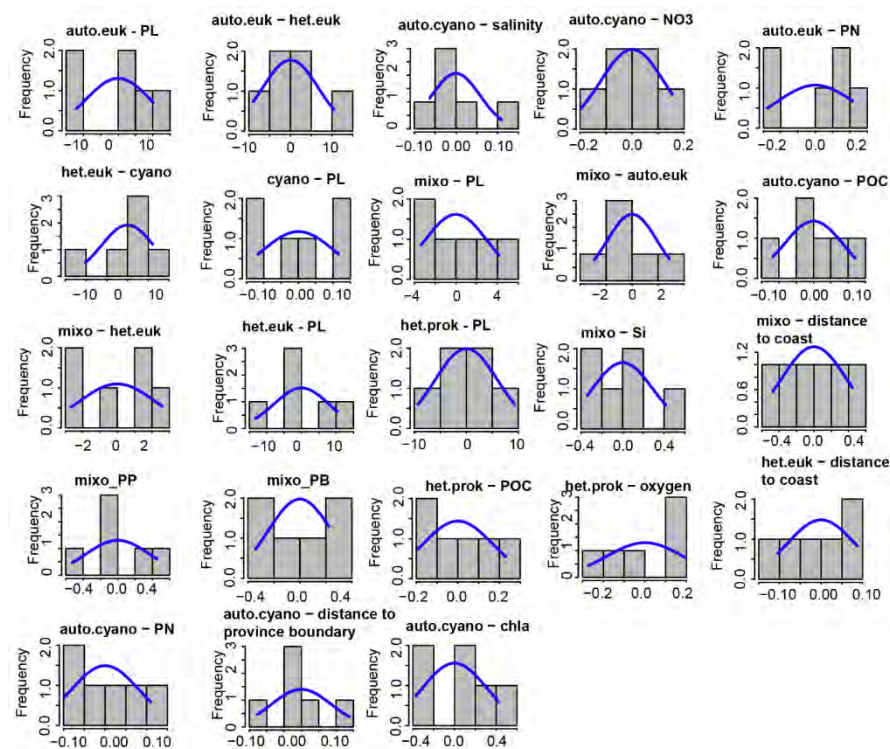


Figure S14. Residual Histograms NADR province (n = 13).

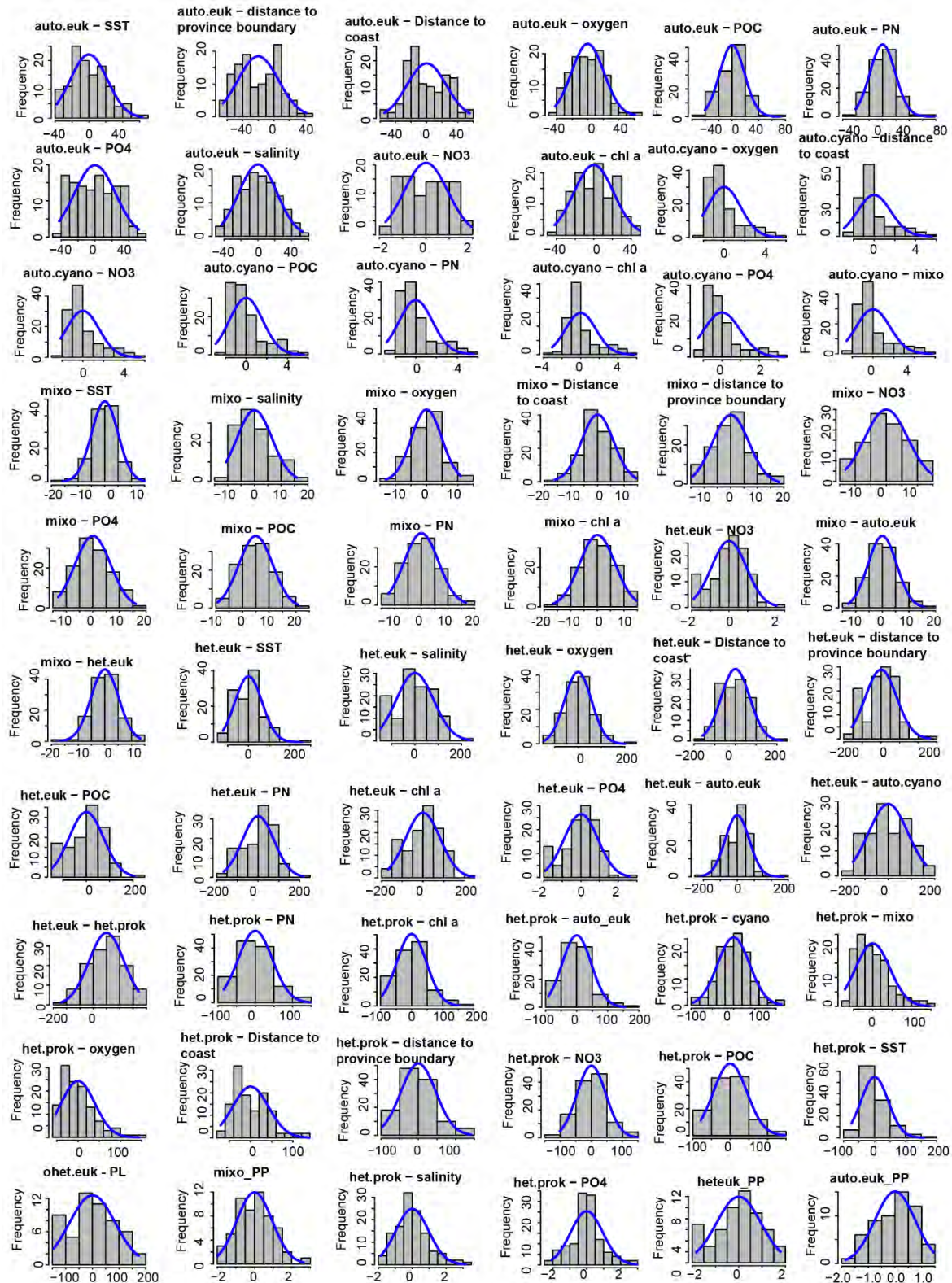


Figure S15. Residual Histograms across all provinces (ALL; n = 121)

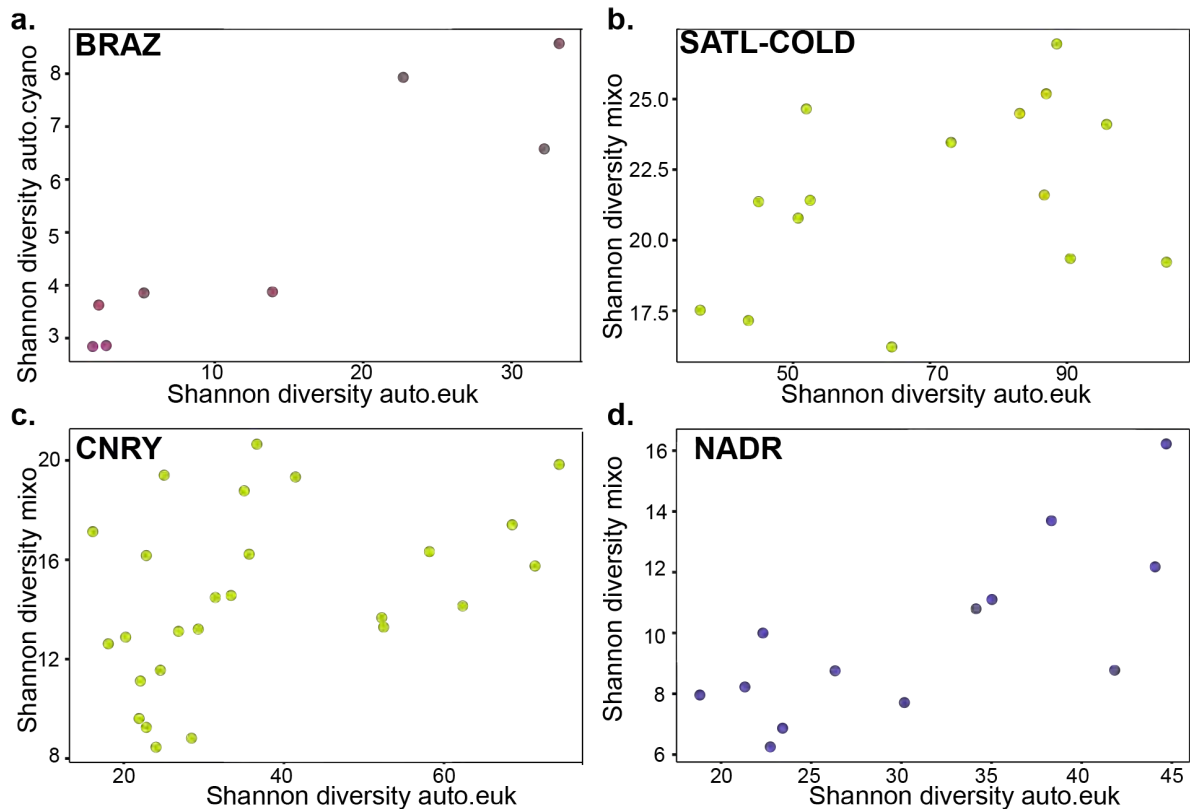


Figure S16. Shannon diversity between functional groups. (a) positive correlation (Pearson, $r = 0.9$, $p < 0.01$, $n = 8$) between Cyanobacteria diversity and autotrophic eukaryotes in the BRAZ province (b) positive correlation (Pearson, $r = 0.6$, $p = 0.01$, $n = 15$) between mixotrophic diversity and autotrophic eukaryotes in the SATL-COLD province, (c) positive correlation (Pearson, $r = 0.7$, $p < 0.001$, $n = 26$) between mixotrophic diversity and autotrophic eukaryotic diversity in the CNRY province, (d) positive correlation (Pearson, $r = 0.7$, $p = 0.01$, $n = 13$) between mixotrophic diversity and autotrophic eukaryotic diversity in the NADR province.

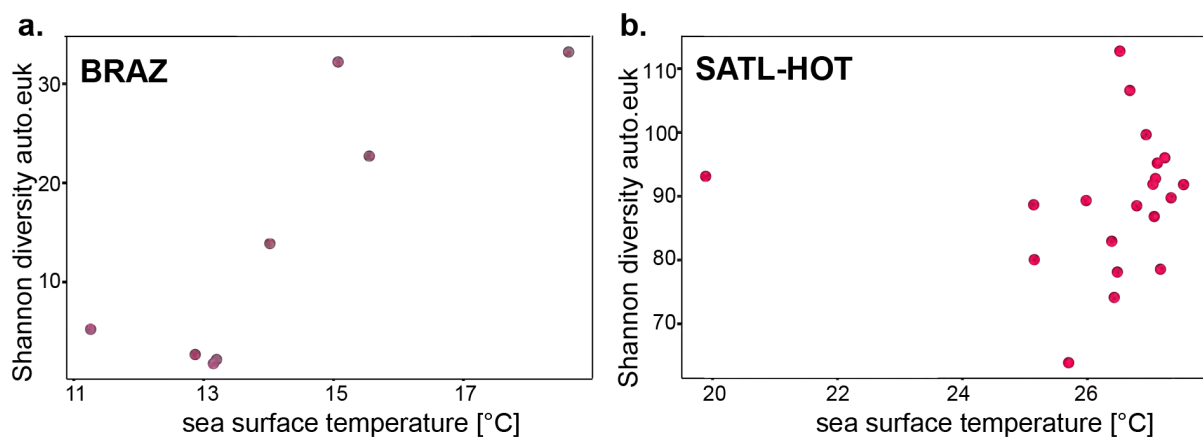


Figure S17. Shannon diversity of autotrophic eukaryotes in the BRAZ (a) and SATL-HOT (b) provinces. The BRAZ province has a large temperature gradient (11.3–18.6°C) and strong correlation with varying alpha diversity (Pearson correlation, $r = 0.85$, $p = 0.007$, $n = 8$). The SATL-HOT province has a small temperature range (25.1–27.6°C; outlier at Station 83 with 19.9°C) and shows no significant correlation with sea surface temperature (Pearson correlation, $r = 0.04$, $p = 0.88$, $n = 20$).

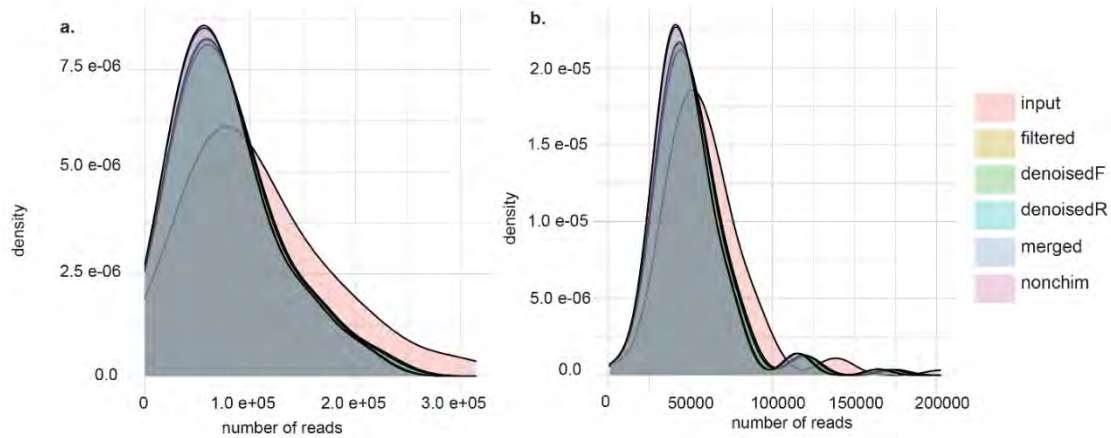


Figure S18. Sequencing statistics density plots of each filtering step for a. 16S rRNA gene sequencing reads and b. 18S rRNA gene sequencing reads.

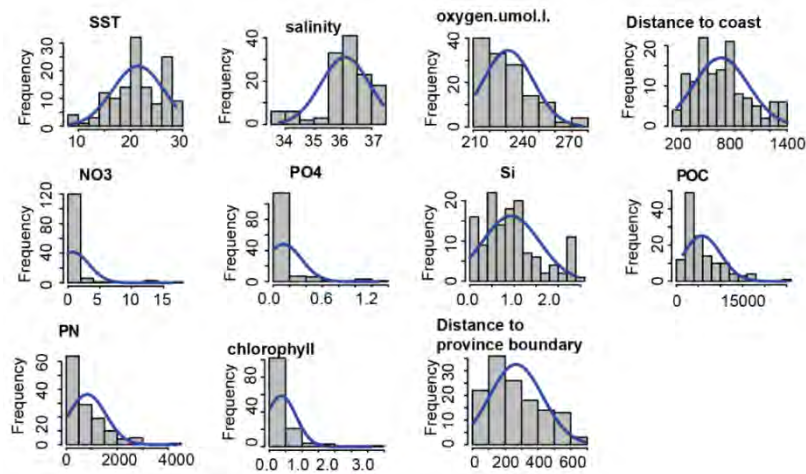


Figure S19. Histograms of metadata across the entire transect (n = 121).

Table S1. Pearson correlation of alpha diversity of trophic functional groups and quantitative cell count against environmental variables across the entire transect. Cells are yellow marked where residuals were normally distributed.

Environmental parameter	auto.euk (n = 124)	auto.cyano (n = 124)	mixo (n = 124)	het.euk (n = 124)	het.prok (n = 124)	Total cell number (n = 115)
temperature	$r = 0.62$, $p < 0.001$	$r = 0.06$, $p = 1$	$r = 0.75$, $p < 0.001$	$r = 0.68$, $p < 0.001$	$r = 0.43$, $p < 0.001$	$r = -0.08$, $p = 1$
salinity	$r = 0.57$, $p < 0.001$	$r = 0.05$, $p = 1$	$r = 0.46$, $p < 0.001$	$r = 0.44$, $p = 0.03$	$r = 0.24$, $p = 0.03$	$r = -0.1$, $p = 1$
O ₂	$r = -0.66$, $p < 0.001$	$r = -0.23$, $p = 0.08$	$r = -0.75$, $p < 0.001$	$r = -0.76$, $p < 0.001$	$r = -0.59$, $p < 0.001$	$r = 0.05$, $p = 1$
NO ₃	$r = -0.34$, $p < 0.001$	$r = -0.23$, $p = 0.08$	$r = -0.37$, $p < 0.001$	$r = -0.33$, $p < 0.001$	$r = -0.29$, $p < 0.001$	$r = -0.02$, $p = 1$
PO ₄	$r = -0.46$, $p < 0.001$	$r = -0.20$, $p = 0.17$	$r = -0.43$, $p < 0.001$	$r = -0.38$, $p < 0.001$	$r = -0.29$, $p < 0.001$	$r = 0.01$, $p = 1$
Si	$r = 0.05$, $p = 0.54$	$r = -0.10$, $p = 1$	$r = 0.12$, $p = 0.16$	$r = -0.03$, $p = 0.75$	$r = -0.02$, $p = 0.8$	$r = 0$, $p = 1$
POC	$r = -0.73$, $p < 0.001$	$r = -0.19$, $p = 0.21$	$r = -0.52$, $p < 0.001$	$r = -0.55$, $p < 0.001$	$r = -0.36$, $p < 0.001$	$r = 0.2$, $p = 0.29$
PN	$r = -0.72$, $p < 0.001$	$r = -0.17$, $p = 0.33$	$r = -0.5$, $p < 0.001$	$r = -0.50$, $p < 0.001$	$r = -0.31$, $p < 0.001$	$r = 0.2$, $p = 0.29$
chl <i>a</i>	$r = -0.62$, $p < 0.001$	$r = -0.03$, $p = 1$	$r = -0.44$, $p < 0.001$	$r = -0.35$, $p = 0.07$	$r = -0.18$, $p = 0.1$	$r = 0.14$, $p = 0.89$
p ^b	$r = -0.18$, $p = 0.15$	$r = -0.06$, $p = 0.65$	$r = 0.14$, $p = 0.28$	$r = 0.08$, $p = 0.51$	$r = 0.02$, $p = 0.87$	$r = 0$, $p = 1$
PP	$r = -0.63$, $p < 0.001$	$r = -0.16$, $p = 0.22$	$r = -0.27$, $p = 0.04$	$r = -0.27$, $p = 0.03$	$r = -0.14$, $p = 0.29$	$r = 0.17$, $p = 0.21$

Table S2. RDA ordination distances between high and low chl *a* provinces and relative ordination distances normalized by distances of all sites to each other within one RDA analyses.

Functional group	provinces	Centroid distances in RDA axes 1+2	high/low chl <i>a</i>	average distance	relative distance (mean distance of chl <i>a</i> provinces/mean distances all provinces)
auto_euk	BRAZ_CNRY	0.09588	high	0.138	0.855
auto_euk	CNRY_FKLD	0.1143694	high		
auto_euk	CNRY_NADR	0.1727393	high		
auto_euk	CNRY_NAST-E	0.1467146	high		
auto_euk	FKLD_NADR	0.1013843	high		
auto_euk	FKLD_NAST-E	0.1543737	high		
auto_euk	NADR_NAST-E	0.1042545	high		
auto_euk	BRAZ_FKLD	0.08769748	high		
auto_euk	BRAZ_NADR	0.1864948	high		
auto_euk	BRAZ_NAST-E	0.2112198	high		
auto_euk	NATR_SATL-COLD	0.1583228	low	0.104	0.649
auto_euk	NATR_SATL-HOT	0.1180996	low		
auto_euk	SATL-COLD_SATL-HOT	0.07251829	low		
auto_euk	WTRA_NATR	0.08980855	low		
auto_euk	WTRA_SATL-COLD	0.1288887	low		
auto_euk	WTRA_SATL-HOT	0.05881104	low		
auto_cyano	BRAZ_CNRY	0.06776943	high	0.101	0.647
auto_cyano	BRAZ_FKLD	0.04163754	high		
auto_cyano	BRAZ_NADR	0.07029267	high		
auto_cyano	BRAZ_NAST-E	0.179421	high		
auto_cyano	CNRY_FKLD	0.04634837	high		
auto_cyano	CNRY_NADR	0.1085454	high		
auto_cyano	CNRY_NAST-E	0.1538886	high		
auto_cyano	FKLD_NADR	0.06223915	high		
auto_cyano	FKLD_NAST-E	0.1396576	high		
auto_cyano	NADR_NAST-E	0.1403244	high		
auto_cyano	NATR_SATL-COLD	0.177128	low	0.13	0.822
auto_cyano	NATR_SATL-HOT	0.1285023	low		
auto_cyano	NATR_WTRA	0.08111078	low		
auto_cyano	SATL-COLD_SATL-HOT	0.05535135	low		
auto_cyano	SATL-COLD_WTRA	0.1977615	low		
auto_cyano	SATL-HOT_WTRA	0.1310024	low		
mixo	BRAZ_CNRY	0.08467257	high	0.116	0.773
mixo	BRAZ_FKLD	0.05857413	high		
mixo	BRAZ_NADR	0.03279495	high		
mixo	BRAZ_NAST-E	0.1983371	high		
mixo	CNRY_FKLD	0.05255136	high		
mixo	CNRY_NADR	0.1427212	high		
mixo	CNRY_NAST-E	0.124576	high		
mixo	FKLD_NADR	0.1129591	high		
mixo	FKLD_NAST-E	0.1666339	high		
mixo	NADR_NAST-E	0.1897172	high		
mixo	NATR_SATL-COLD	0.1224816	low	0.13	0.861
mixo	NATR_SATL-HOT	0.1061296	low		
mixo	NATR_WTRA	0.09472547	low		
mixo	SATL-COLD_SATL-HOT	0.06873986	low		
mixo	SATL-COLD_WTRA	0.2116922	low	0.113	0.711
mixo	SATL-HOT_WTRA	0.1739107	low		
het_euk	BRAZ_CNRY	0.09177388	high		
het_euk	BRAZ_FKLD	0.05746816	high		
het_euk	BRAZ_NADR	0.1071548	high		
het_euk	BRAZ_NAST-E	0.2082722	high		
het_euk	CNRY_FKLD	0.06619234	high		
het_euk	CNRY_NADR	0.1124882	high		
het_euk	CNRY_NAST-E	0.1469013	high		
het_euk	FKLD_NADR	0.05835558	high		

het_euk	FKLD_NAST-E	0.1515959	high	0.138	0.873
het_euk	NADR_NAST-E	0.1282205	high		
het_euk	NATR_SATL-COLD	0.1755226	low		
het_euk	NATR_SATL-HOT	0.155487	low		
het_euk	NATR_WTRA	0.08559509	low		
het_euk	SATL-COLD_SATL-HOT	0.07155067	low		
het_euk	SATL-COLD_WTRA	0.1970334	low		
het_euk	SATL-HOT_WTRA	0.1457272	low	0.126	0.767
het_prok	BRAZ_CNRY	0.1352539	high		
het_prok	BRAZ_FKLD	0.1290217	high		
het_prok	BRAZ_NADR	0.04349993	high		
het_prok	BRAZ_NAST-E	0.08419363	high		
het_prok	CNRY_FKLD	0.2451325	high		
het_prok	CNRY_NADR	0.1786196	high		
het_prok	CNRY_NAST-E	0.1593004	high		
het_prok	FKLD_NADR	0.09890021	high		
het_prok	FKLD_NAST-E	0.09314179	high		
het_prok	NADR_NAST-E	0.09122453	high		
het_prok	NATR_SATL-COLD	0.1900992	low	0.12	0.734
het_prok	NATR_SATL-HOT	0.1067188	low		
het_prok	NATR_WTRA	0.1652539	low		
het_prok	SATL-COLD_SATL-HOT	0.09663984	low		
het_prok	SATL-COLD_WTRA	0.09943583	low		
het_prok	SATL-HOT_WTRA	0.0645592	low		

Table S3. Procrustes analysis of RDA ordinations between different trophic functional groups of pro- and eukaryotes.

	auto.euk	auto.cyano	mixo	het.euk
auto.cyano	284.1			
mixo	498.2	99.9		
het.euk	281.9	291.6	89.9	
het.prok	182.2	209	259.7	559.4

Table S4. Estimated phytoplankton biomass derived from chl *a* concentration and estimated factor (23) based on a quantitative literature value (2). Notably, this is an approximation that varies on regional scales and on the species level. Our estimated phytoplankton biomass falls in the expected range of 1–10% of POC concentration with varying proportions between provinces, and likely underestimation in low, oligotrophic latitudes (<https://doi.org/10.1002/2016GB005458> Artega et al. 2016).

province	Chl <i>a</i> (mg m ⁻³)	total C (mg m ⁻³)	estimated phytoplankton biomass* (mg m ⁻³) * = chl <i>a</i> * 23
BRAZ	0.81–3.1	48.85–203.44	18.63–71.3
Station 24	0.72	36.84	16.56
CNRY	0.13–1.22	38.41–299.06	2.99–28.6
FKLD	0.65	53.51–98.95	14.95
NADR	0.14–0.44	58.85–152.54	3.22–10.12
NAST-E	0.07–0.18	35.38–68.77	1.61–4.14
NATR	0.14–0.43	64.27–109.66	3.22–9.89
SATL-COLD	0.08–0.32	26.26–52.28	1.84–7.36
SATL-HOT	0.04–0.13	18.3–36.5	0.92–2.99
WTRA	0.13–0.35	31.31–53.46	2.99–8.05
Station 90+92	0.09–0.1	32.44	2.07–2.3

Table S5. Literature research and trophic assignment of prokaryotic and eukaryotic taxonomic groups.

Functional group	taxa	Reference	doi/ ISBN	comment
mixotroph	Cryophyceae	Olrik 1998	https://doi.org/10.1007/978-94-017-2668-9_28	
mixotroph	Dinoflagellata	Stoecker 2007	https://doi.org/10.1111/j.1550-7408.1999.tb04619.x	
mixotroph	Chlorarachniophyceae	Ota & Ishida 2008	https://doi.org/10.1002/9780470015902.a0003060.pub2	
mixotroph	Dinoflagellata/ Takayama	Adl et al. 2018	https://doi.org/10.1111/jeu.12691	
mixotroph	Dinoflagellata/ Noctiluca	Adl et al. 2018	https://doi.org/10.1111/jeu.12691	
mixotroph	Dinoflagellata/ Margalefidinium	Adl et al. 2018	https://doi.org/10.1111/jeu.12691	
mixotroph	Dinoflagellata/ Blastodinium	Adl et al. 2018	https://doi.org/10.1111/jeu.12691	
mixotroph	Dinoflagellata/ Dinophysis	Adl et al. 2018	https://doi.org/10.1111/jeu.12691	
mixotroph	Dinoflagellata/ Tripos	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648	
mixotroph	Dinoflagellata/ Gonyaulax polygramma	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648	
mixotroph	Dinoflagellata/ Fragilidium	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648	
mixotroph	Dinoflagellata/ Scripsiella	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648	
mixotroph	Dinoflagellata/ Goniodoma	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648	
mixotroph	Dinoflagellata/ Symbiodinium	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648	
mixotroph	Dinoflagellata/ Alexandrium_ andersonii	Lee et al. 2018	https://doi.org/10.1016/j.hal.2016.09.008	
mixotroph	Dinoflagellata/ Amphidinium_gibbosum	Murray et al. 2004	https://doi.org/10.1046/j.1529-8817.2004.03132.x	lifestyle not described, but contains plastids
mixotroph	Acantharea/ Acanthochiasma	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648	
mixotroph	Acantharea/ Acanthostaurus	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648	
mixotroph	Acantharea/ Amphilonche_elongata	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648	
mixotroph	Acantharea/ Symphyacanthida (Clade F)	Decelle et al. 2012	https://doi.org/10.1073/pnas.1212303109	
mixotroph	Ciliates/ Laboea	Lynn, Denis 2008. The ciliated protozoa: characterization, classification, and guide to the literature. Springer Science & Business Media		
mixotroph	Ciliates/ Pseudotontonia	Lynn 2008		
mixotroph	Ciliates/ Spirotontonia	Lynn 2008		
mixotroph	Ciliata/ Strombidium spp.	Haraguchi et al. 2018	https://doi.org/10.3389/fmars.2018.00272	
eukaryotic autotrophs	Bacillariophyta	Round et al. 1990	ISBN 0521363187	
eukaryotic autotrophs	Bolidophyceae	Kuwata et al. 2018	https://doi.org/10.3389/fmars.2018.00370	
eukaryotic autotrophs	Phaeothamniophyceae	Fehling et al. 2007	https://doi.org/10.1016/B978-012370518-1/50007-2	
eukaryotic autotrophs	Pelagophyceae	Andersen et al. 1993	https://doi.org/10.1111/j.0022-3646.1993.00701.x	
eukaryotic autotrophs	Prymniales	Edwardsen et al. 2011	https://doi.org/10.1080/09670262.2011.594095	some might be mixotroph
eukaryotic autotrophs	Cocco (order level)	Paasche, E.J. "Biology and physiology of coccolithophorids," Annual Reviews in Microbiology 22, no. 1 (1968): 71-86.		
eukaryotic autotrophs	Calcihaptophycidae	Paasche, E.J. "Biology and physiology of coccolithophorids," Annual Reviews in Microbiology 22, no. 1 (1968): 71-86.		
eukaryotic autotrophs	Phaeocystales	Paasche, E.J. "Biology and physiology of coccolithophorids," Annual Reviews in Microbiology 22, no. 1 (1968): 71-86.		
eukaryotic autotrophs	Cryptophyceae	Dag Klaveness (1989) Biology and Ecology of the Cryptophyceae: Status and Challenges, Biological Oceanography, 6:3-4, 257-270 https://doi.org/10.1080/01965581.1988.10749530		
eukaryotic autotrophs	Dinoflagellata/ Pyrocystis	Adl et al. 2018	https://doi.org/10.1111/jeu.12691	
eukaryotic autotrophs	Dinoflagellata/ Alexandrium_ insuetum	Lim et al. 2019	https://doi.org/10.4490/algae.2019.34.11.21	

eukaryotic autotrophs	Dinoflagellata/ Dissodinium_pseudolunula	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648
eukaryotic autotrophs	Dinoflagellata/ Protoceratium_reticulatum	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648
eukaryotic autotrophs	Chlorophyta	Not et al. 2012	https://doi.org/10.1016/B978-0-12-391499-6.00001-3
prokaryotic autotrophs	Cyanobacteria	Campbell et al. 1998	https://doi.org/10.1128/MMBR.62.3.667-683.1998
eukaryotic heterotrophs	Pseudofungi	Richards et al. 2011	https://doi.org/10.1146/annurev-marine-120710-100802
eukaryotic heterotrophs	Opalozoa	Cavalier-Smith 1993	https://doi.org/10.1111/j.1550-7408.1993.tb06117.x
eukaryotic heterotrophs	Sagenista	Cavalier-Smith & Chao 2006	https://doi.org/10.1007/s00239-004-0353-8
eukaryotic heterotrophs	Telonemia	Shalchian-Tabrizi et al. 2006	https://doi.org/10.1098/rspb.2006.3515
eukaryotic heterotrophs	Katablepharidophyta	Okamoto & Inouye 2005	https://doi.org/10.1016/j.protis.2004.12.003
eukaryotic heterotrophs	Apicomplexa	Morrison 2009	https://doi.org/10.1016/j.pt.2009.05.010
eukaryotic heterotrophs	Choanoflagellida	Fenchel 1986	https://doi.org/10.1007/978-1-4757-0611-6_2
eukaryotic heterotrophs	Fungi	Richards et al. 2011	https://doi.org/10.1146/annurev-marine-120710-100802
eukaryotic heterotrophs	Mesomycetozoa	Mendoza et al. 2002	https://doi.org/10.1146/annurev-micro.56.012302.160950
eukaryotic heterotrophs	Endomyxa	Dumack et al. 2019	https://doi.org/10.1111/1755-0998.13112
eukaryotic heterotrophs	Polycystinea	Anderson 1988	https://doi.org/10.1007/978-3-662-11340-0
eukaryotic heterotrophs	Apusozoa	Glücksman, E. (2011). Taxonomy, biodiversity, and ecology of Apusozoa (Protozoa) (Doctoral dissertation, Oxford University, UK).	
eukaryotic heterotrophs	Dinoflagellata/ Kofoidinium	Adl et al. 2018	https://doi.org/10.1111/jeu.12691
eukaryotic heterotrophs	Dinoflagellata/ Protoperidinium	Adl et al. 2018	https://doi.org/10.1111/jeu.12691
eukaryotic heterotrophs	Dinoflagellata/ Cucumeridinium	Adl et al. 2018	https://doi.org/10.1111/jeu.12691
eukaryotic heterotrophs	Dinoflagellata/ Gyrodinium	Adl et al. 2018	https://doi.org/10.1111/jeu.12691
eukaryotic heterotrophs	Dinoflagellata/ Abedinium_dasyus	Cooney et al. 2020	https://doi.org/10.1093/gbe/eva196
eukaryotic heterotrophs	Dinoflagellata/ Amoebophrya	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648
eukaryotic heterotrophs	Dinoflagellata/ Euduboscquella	Coats et al. 2012	https://doi.org/10.1111/j.1550-7408.2011.00588.x
eukaryotic heterotrophs	Dinoflagellata/ Hematodinium	Li et al. 2010	https://doi.org/10.1016/j.hal.2010.04.001
eukaryotic heterotrophs	Dinoflagellata/ Syndinium_turbo	Skovgaard et al. 2005	https://doi.org/10.1016/j.protis.2005.08.002
eukaryotic heterotrophs	Acantharea/ Chaunacanthid	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648
eukaryotic heterotrophs	Acantharea/ Gigartacon	Schneider et al. 2020	https://doi.org/10.3897/BDJ.8.e56648
eukaryotic heterotrophs	Dinoflagellata/ Dino Group I	Guillou et al. 2008	https://doi.org/10.1111/j.1462-2920.2008.01731.x
eukaryotic heterotrophs	Dinoflagellata/ Dino Group II	Guillou et al. 2008	https://doi.org/10.1111/j.1462-2920.2008.01731.x
eukaryotic heterotrophs	Ciliata/ Pelagostrobilidium	Haraguchi et al. 2018	https://doi.org/10.3389/fmars.2018.00272
eukaryotic heterotrophs	Ciliata/ Akenasia	Haraguchi et al. 2018	https://doi.org/10.3389/fmars.2018.00272
eukaryotic heterotrophs	Ciliata/ Tintinnida	Haraguchi et al. 2018	https://doi.org/10.3389/fmars.2018.00272
prokaryotic heterotrophs	all bacterial phyla except Cynobacteria were considered as heterotrophs, sensu Pomeroy et al. 2007 https://www.jstor.org/stable/24860040	Herndl et al. 2008 https://doi.org/10.3354/ame01225 ; Presence of chemoautotrophy in e.g. nitrifying organisms was not considered in this study due to 1) low abundance of chemoautotrophy in pelagic, open ocean ecosystems and 2) all samples were well oxygenated. see Middelburg 2011; https://doi.org/10.1029/2011GL049725 Presence of proteorhodopsin is widely distributed among prokaryotes (see Ferrera et al. 2015; http://dx.doi.org/10.1016/j.mib.2015.03.007) but is not considered in this study. Likewise, prokaryotes were not categorized according to their energy source	

Supplementary 3

Supplementary to Chapter 3

Supplementary A: Additional information Materials and Methods

HE492 and HE533:

Seawater samples were collected during the on the HE492 expedition between 03.08.2017 and 16.08.2017, and the HE533 expedition between 20.05.2019 and 06.06.2019. Water samples were collected in triplicates (A–C) with Niskin bottles mounted on a Seabird' SBE911+' CTD probe with additional turbidity, oxygen, and fluorescence sensors.

A total of 20 L of seawater at three different depths (3m, DCM which varied between 8–28m, and 40m depth) was pooled and gravity-filtered through 200 μm and 20 μm mesh size sieves, and subsequently filtered through 3 μm and 0.2 μm polycarbonate filters (147 mm diameter, Millipore) using a Millipore Tripod filterholder and a peristaltic pump within a time window of max. 30 min. Warm lysis buffer was added to filters, followed by snap freezing and storage at -80°C until further processing. Picoplankton DNA was extracted using the NucleoSpin® Soil kit (Macherey-Nagel, Germany) following the manufacturer's protocol. For the sample lysis step, a bead beater was used to break up the cells (MagNA Lyser, Roche). As other datasets in this manuscript did not have replicates, we subsampled only the first replicate (A) of the dataset for further analysis.

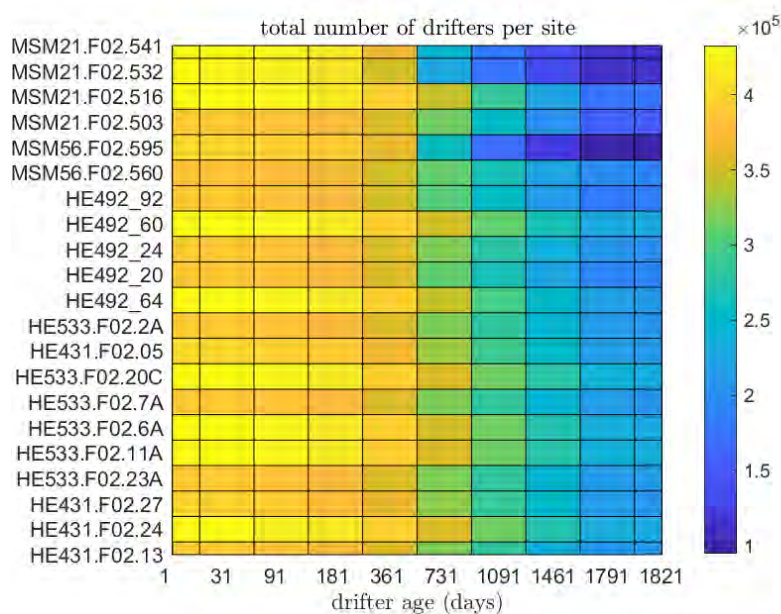


Figure S1. Total number of numerical synthetic drifters per site and for individual temporal bins.

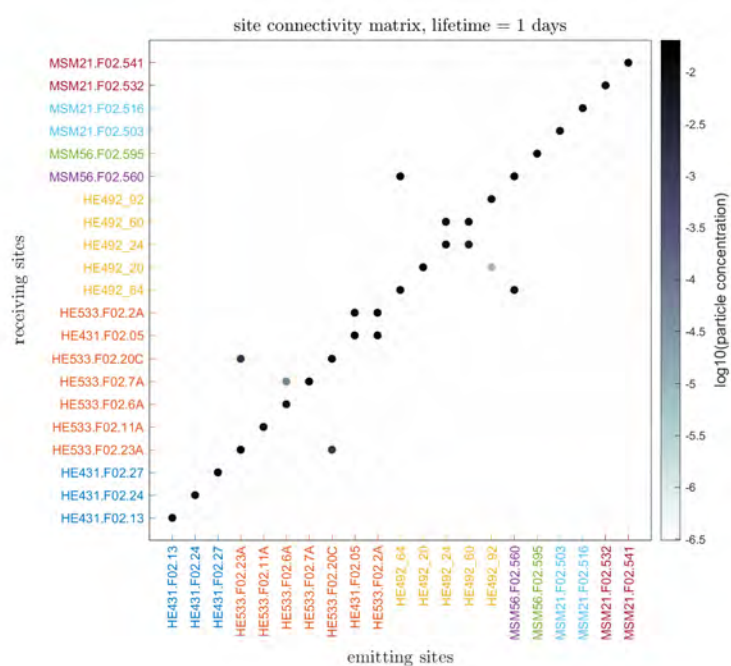


Figure S2. Connectivity matrix of synthetic numerical drifters that are 1 day old.

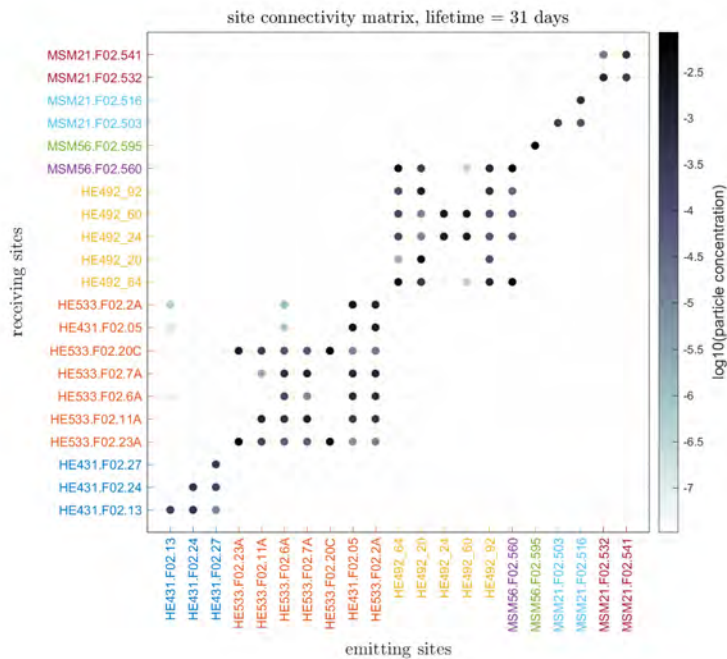


Figure S3. Connectivity matrix of synthetic numerical drifter concentration after 1 month (31 days)

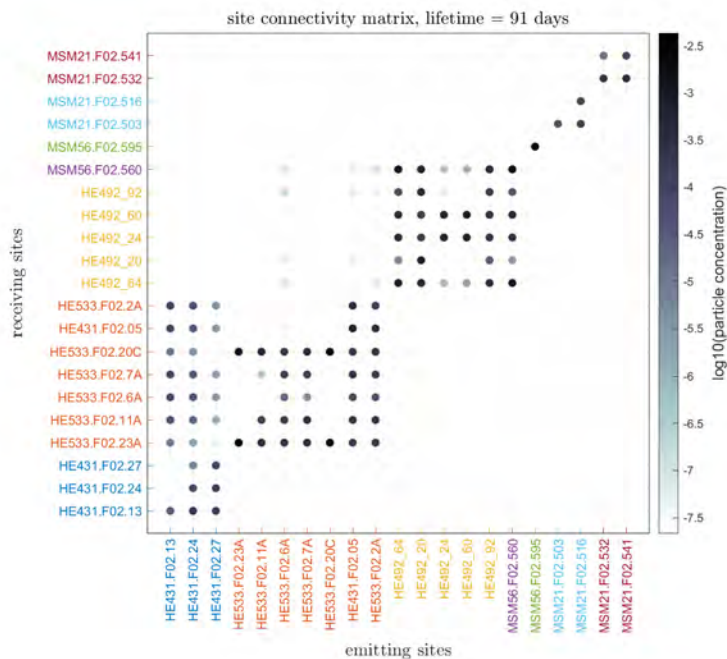


Figure S4. Connectivity matrix of synthetic numerical drifter concentration after 3 months (91 days)

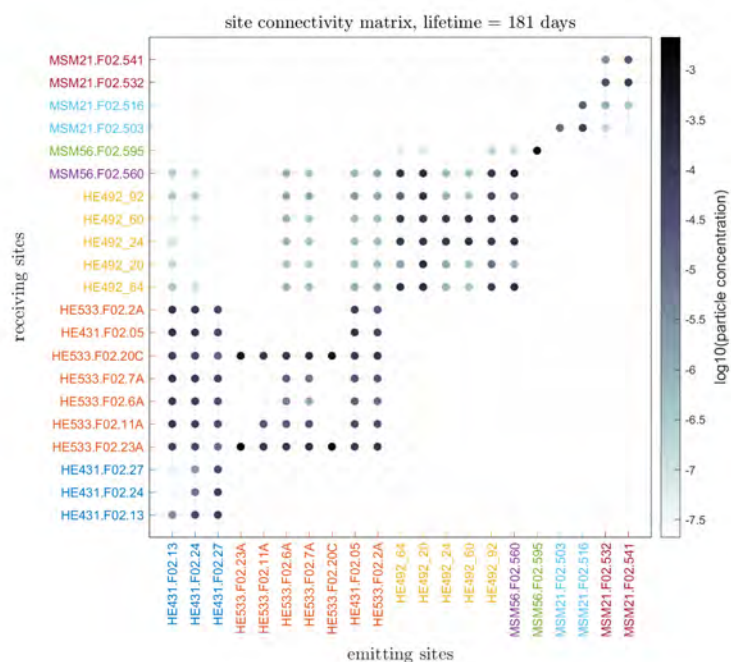


Figure S5. Connectivity matrix of synthetic numerical drifter concentration after 6 months (181 days)

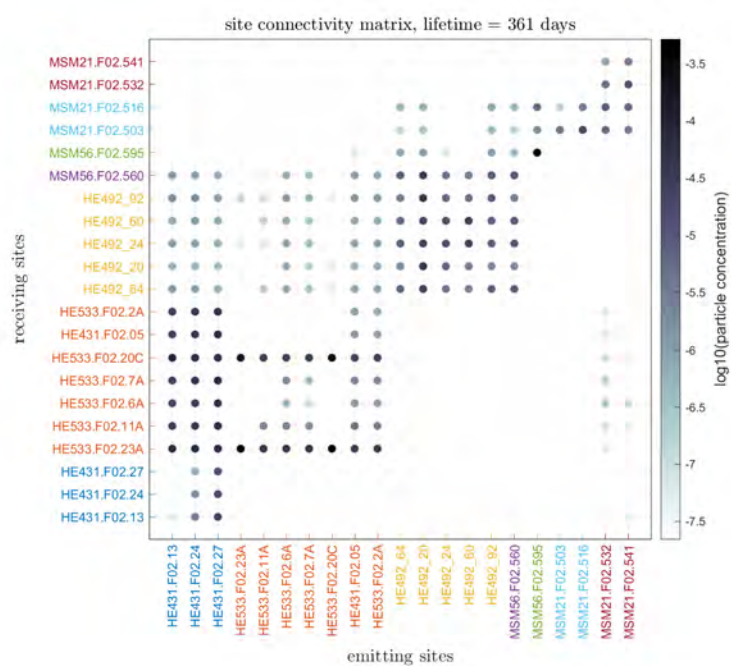


Figure S6. Connectivity matrix of synthetic numerical drifter concentration after 1 year (361 days).

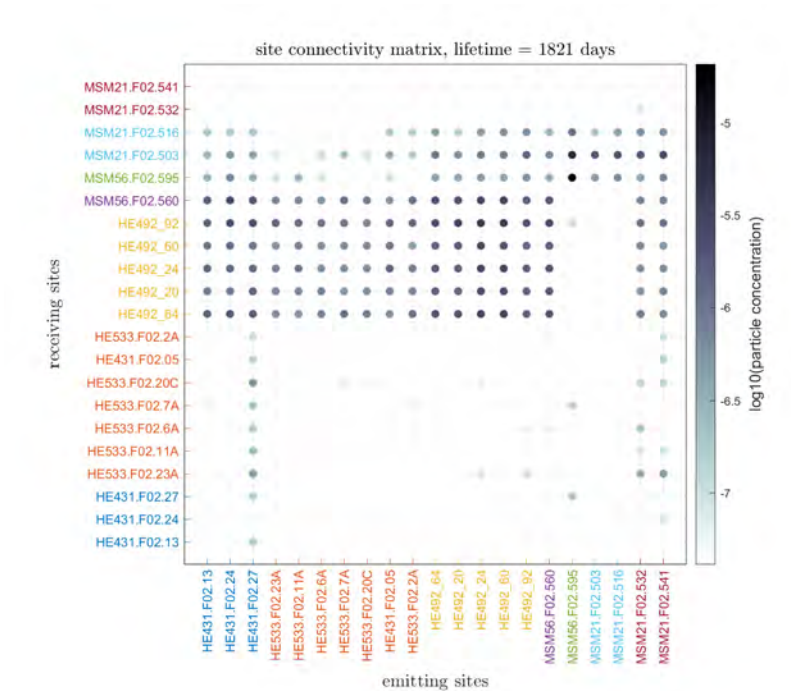


Figure S7. Connectivity matrix of synthetic numerical drifter concentration after 5 years (1821 days).

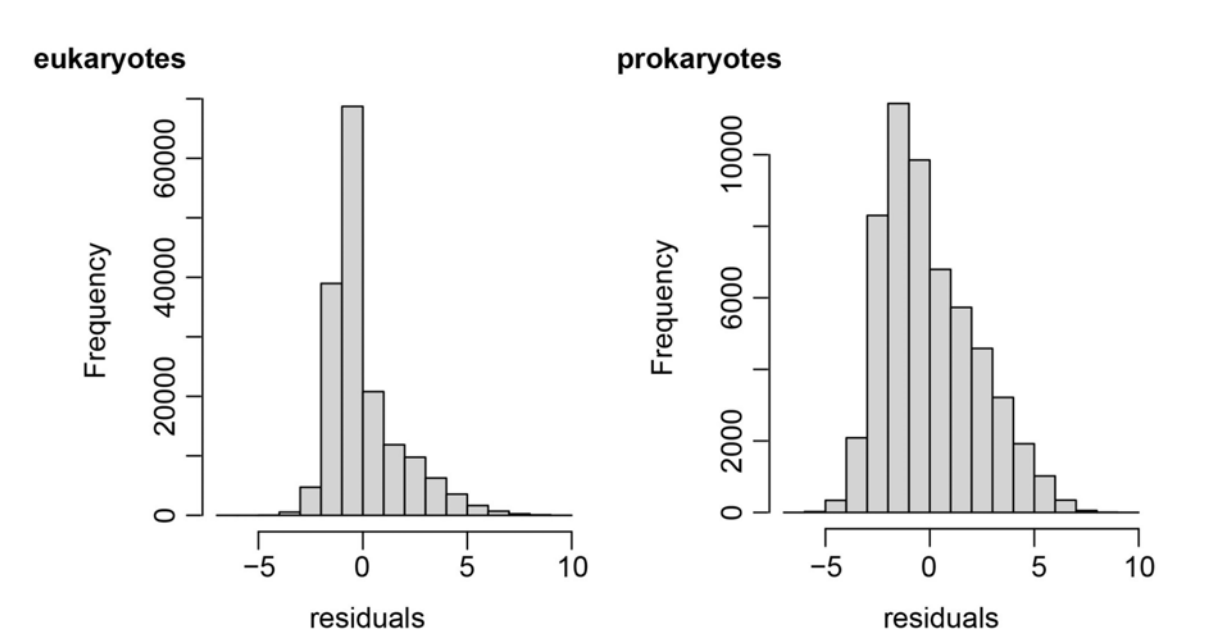


Figure S8. residuals of both eukaryotes and prokaryotic RDA are normally distributed.

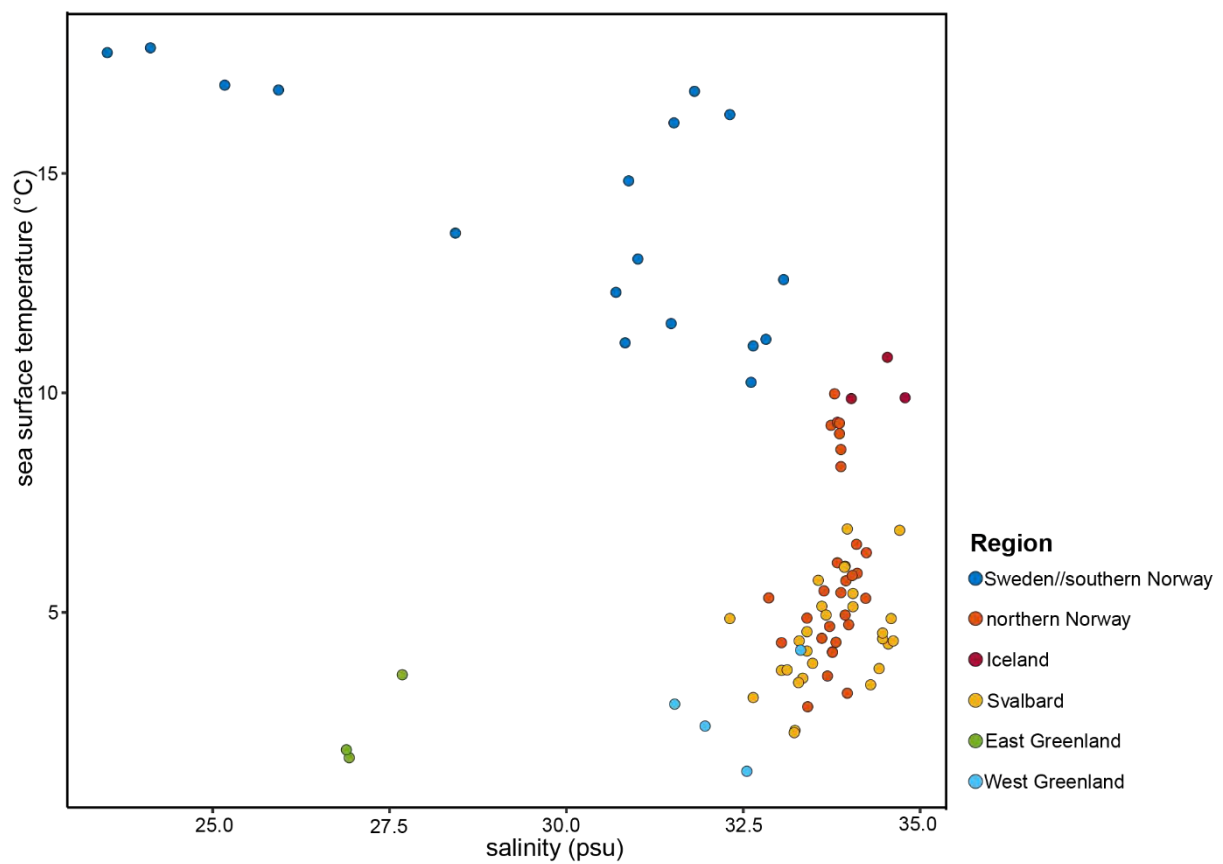


Figure S9. Temperature–salinity plot of all sites from surface water samples. Sites are color-coded according to geographic region.

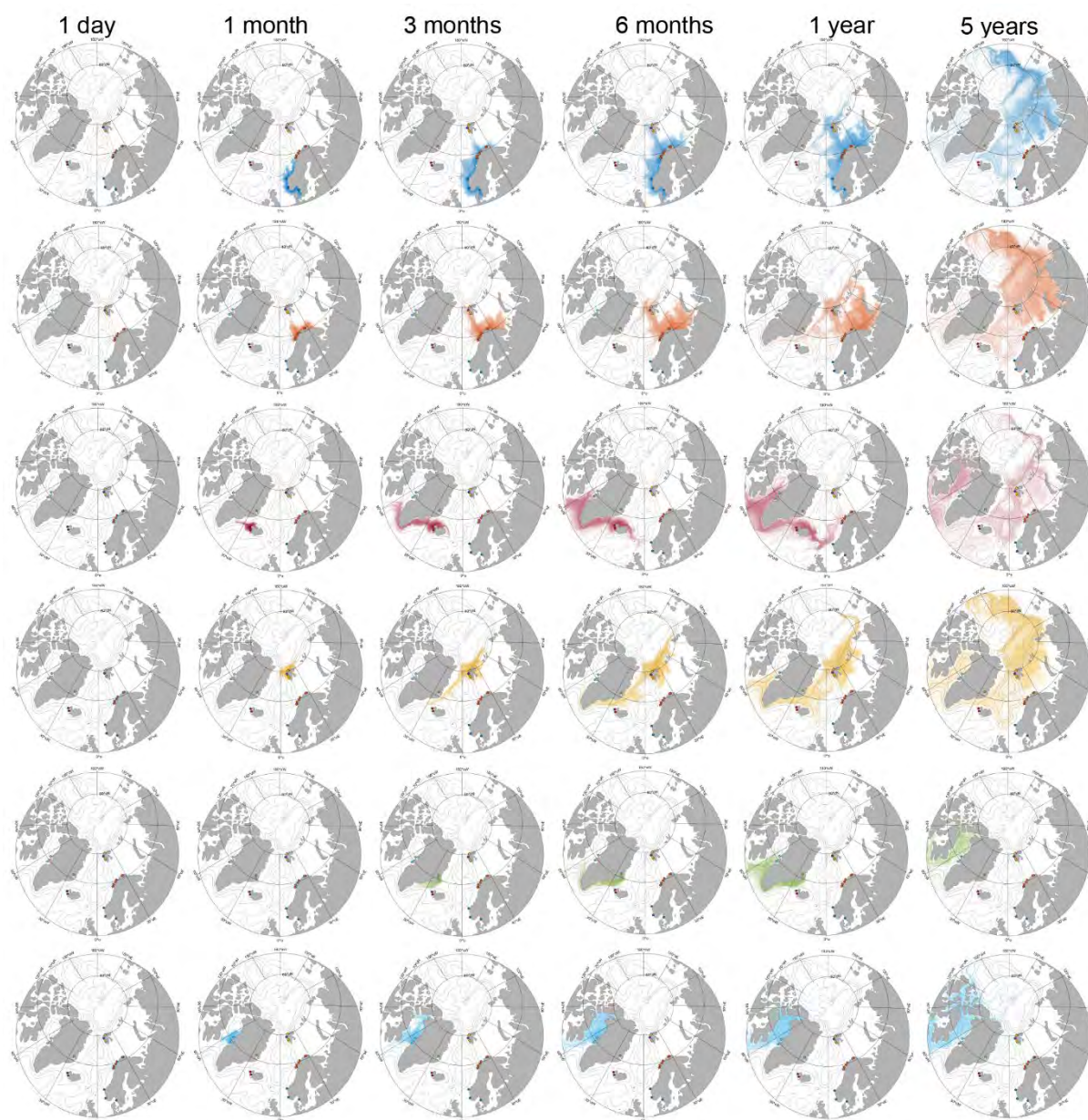


Figure S10. Synthetic particle dispersal of individual regions after 1 day, 1 month, 3 months, six months, 1 year, and 5 years.

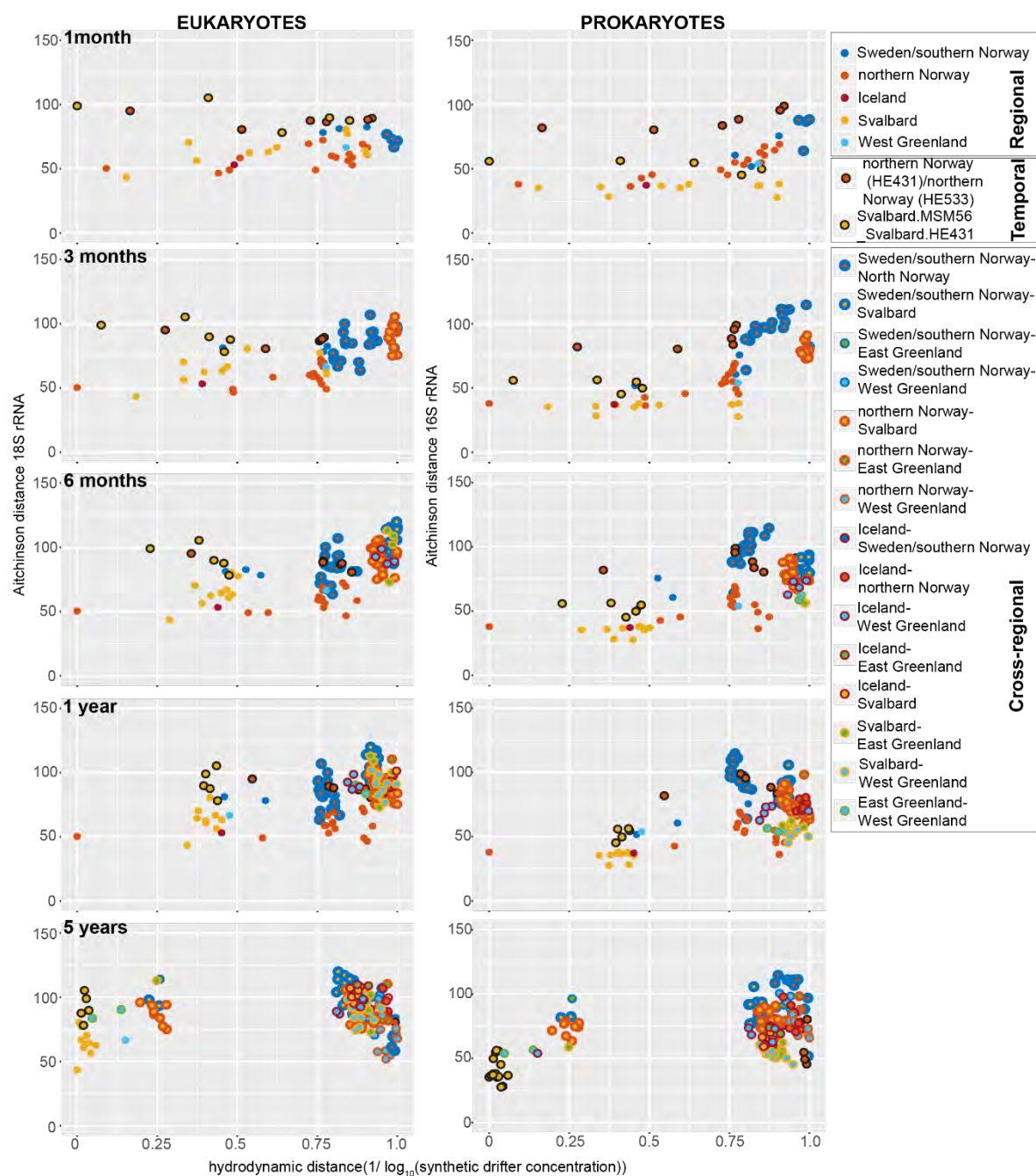


Figure S11. Microbial beta diversity distance (Aitchinson distance) based on 16S rRNA (prokaryotes) and 18S rRNA (eukaryotes) sequences analyses against hydrodynamic distance based on the inverse of the normalized \log_{10} synthetic particle concentration.

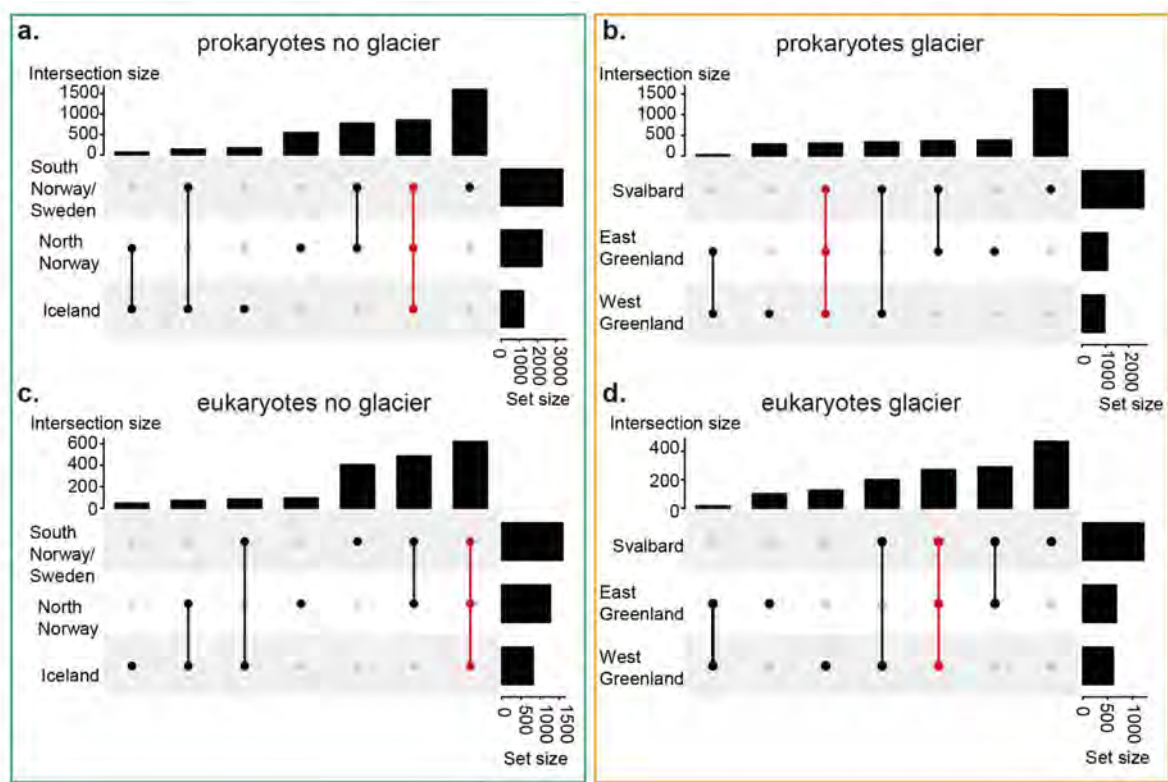


Figure S12. upsetR plots representing unique prokaryotic and eukaryotic ASV only occurring within one geographic region and the overlaps of ASVs between geographic regions. a. prokaryotic ASVs between regions without marine-terminating glaciers, b. prokaryotic ASVs between regions with marine terminating glaciers, c. eukaryotic ASVs between regions without marine-terminating glaciers, d. eukaryotic ASVs between regions with marine terminating glaciers.

Table S1. Sampling stations including time, location, dissolved inorganic nutrient concentrations as well as grouping into geographic location, individual fjords and according to the presence of marine-terminating glaciers.

Station	Region	Fjord	Date	temperature (°C)	salinity	Fluorometer	PO ₄ μmol l ⁻¹	Si μmol l ⁻¹	NO ₃ μmol l ⁻¹	Glacier	n
HE533.F02.25A	North,Norway	Tanaafjord	03.06.2019–04.06.2019	4.3–5.3	33.04–34.23	0.5–1.6	0.2–0.3	2.1–5.2	3.2–3.3	No	4
HE533.F02.21A	North,Norway	Laksefjord	01.06.2019–02.06.2019	3.2–4.1	33.76–33.97	0.3–1.0	0.2–0.3	0.6–1.2	1.3–2.6	No	3
HE533.F02.17A	North,Norway	Porsangerfjord	30.05.2019–31.05.2019	2.9–4.7	33.41–33.99	0.6–3.7	0–2	1.1–8.5	1.7–2.3	No	4
HE533.F02.11A	North,Norway	Lyngefjord	27.05.2019–28.05.2019	4.9–5.9	32.86–34.28	0.6–2.1	0.3–0.4	1.6–4.1	2.0–4.6	No	6
HE533.F02.10A	North,Norway	Balsfjord	25.05.2019–26.05.2019	5.0–5.8	32.05–34.04	1.4–7.8	0.2–0.4	2.5–3.4	0.8–4.3	No	6
HE431.F02.03	North,Norway	Lofoten.HE431	24.08.2014–25.08.2014	8.3–10.0	33.74–33.88	0.9–1.4	0.8–2.1	0.1–0.2	1.4–4.3	No	7
HE533.F02.2A	North,Norway	Lofoten.HE33	23.05.2019–24.05.2019	5.5–6.5	33.83–34.24	0.6–1.5	0.2–0.3	1.4–1.7	0.9–3.7	No	5
HE431.F02.13	South,Norway	Nordfjord	28.08.2014–29.08.2014	10.2–12.6	32.61–33.15	0.6–0.9	0.9–2.5	0.001–0.08	1.5–3.3	No	5
HE431.F02.19	South,Norway	Sognefjord	30.08.2014–31.08.2014	11.1–13.1	30.70–31.48	0.7–1.5	1.0–3.2	0.05–0.2	2.0–2.4	No	4
HE431.F02.23	South,Norway	Boknafjord	02.09.2014–03.09.2014	14.8–16.9	30.9–32.3	1.3–2.7	0.001–0.6	0.001–0.2	1.4–2.0	No	4
HE431.F02.27	South,Norway	Orust-Tjörn.Fjord	05.09.2014–06.09.2014	13.6–17.9	23.51–28.43	0.9–3.3	0.001–4.0	0.001–1.2	2.9–18.7	No	5
HE492_12	Svalbard	Van.Mijen.Fjord	03.08.2017–04.08.2017	3.1–5.1	32.31–34.59	0.25–1.89	0.1–0.2	1.3–3.7	0.5–1.6	Yes	5
HE492_100	Svalbard	Isfjord	15.08.2017–16.08.2017	5.0–5.8	23.32–33.67	1.0–2.1	0.1–0.2	0.4–1.3	0.1–1.0	Yes	6
HE492_64	Svalbard	Kongsfjord	11.08.2017–12.08.2017	5.4–6.9	33.93–34.71	0.3–1.8	0.1–0.2	0.8–1.9	0.4–1.0	Yes	4
MSM56.F02.553	Svalbard	Kongsfjord	03.07.2016–04.07.2016	3.4–4.5	34.21–34.62	0.08–0.09	0.2–0.3	0.6–1.2	0.3–1.7	Yes	8
HE492_44	Svalbard	Woodfjord	09.08.2017–10.08.2017	2.3–4.1	33.04–33.4	0.3–2.0	0.1–0.3	1.3–3.5	0.3–3.0	Yes	5
HE492_24	Svalbard	Wijdefjord	06.08.2017–07.08.2017	3.4–4.6	33.28–33.48	0.6–2.2	0.2–0.4	1.1–3.4	0.9–3.8	Yes	5
MSM21.F02.527	Iceland	Iceland	05.08.2012–08.08.2012	9.9–10.8	34.06–34.82	0.48–0.94	0.15–0.9	0.96–5.56	0.1–9.8	No	4
MSM56.F02.580	East.Greenland	Nordvestfjord.Scoresby.Sund	12.07.2016–17.07.2016	1.7–9.0	16.18–27.68	0.02	0.04–10.2	1.2–2.6	0.02–0.5	Yes	4
MSM21.F02.503	West.Greenland	Disco Bay	27.07.2012–28.07.2012	1.4–6.0	31.44–33.31	0.1–2.9	0.2–0.8	1.7–12	2.9–11	Yes	6

Table S2. Global network properties of the whole network components.

	Non-glacial influence	glacial influence
Number of components	70	64
Clustering coefficient	0.27	0
Modularity	0.74	0.83
Positive edge percentage	74.42	75.55
Edge density	0.007	0.008
Natural connectivity	0.01	0.01

Table S3. Ten most abundant prokaryotic and eukaryotic ASVs of fjords. Rows are color coded according to presence or absence of marine-terminating glaciers. Dark colors indicate those ASVs which also appeared as central nodes in the co-occurrence network analysis.

Top 10 eukaryotic ASVs										
Fjord	1	2	3	4	5	6	7	8	9	10
Tanaifjord	ASV4/ Cryptomonadales_X	ASV6/ Mamiellaceae	ASV7/ Phaeocystaceae	ASV12/Pyramimonadales_XX	ASV14/ Mamiellaceae	ASV17/Dinophyceae_XXX_sp.	ASV18/Dinophyceae_XXX_sp.	ASV54/Dinophyceae_XXX_sp.	ASV122/Chrysophyceae_Clade-F	ASV147/ Dinoflagellata
Laksefjord	ASV1/ Cryptomonadales_X	ASV4/ Cryptomonadales_X	ASV6/ Mamiellaceae	ASV7/Phaeocystaceae	ASV12/Pyramimonadales_XX	ASV14/ Mamiellaceae	ASV17/Dinophyceae_XX	ASV31/Picozoa_XX	ASV39/ Stramenopiles	ASV122/Chrysophyceae_Clade-F
Porsanger fjord	ASV1/ Cryptomonadales_X	ASV2/ Mamiellaceae	ASV7/ Phaeocystaceae	ASV12/Pyramimonadales_XX	ASV14/ Mamiellaceae	ASV17/Dinophyceae_XX	ASV18/Dinophyceae_XX	ASV31/Picozoa_XXX	ASV39/ Stramenopiles	ASV63/ Strombidida
Lyngen fjord	ASV1/ Cryptomonadales_X	ASV4/ Cryptomonadales_X	ASV6/ Mamiellaceae	ASV7/ Phaeocystaceae	ASV12/ Pyramimonadales_XX	ASV14/ Mamiellaceae	ASV16/Mamiellales	ASV28/Chrysochromulinales	ASV28/ Dinophyceae_XX	ASV31/ Picozoa_XXX
Balsfjord	ASV1/ Cryptomonadales_X	ASV6/Mamiellaceae	ASV7/ Phaeocystaceae	ASV12/Pyramimonadales_XX	ASV14/Mamiellaceae	ASV22/Chrysochromulinales	ASV31/Picozoa_XX	ASV38/NPK2-lineage	ASV154/ Pyrmesiacae	ASV155/Dino-Group-II-Clade-16
Lofoten HE431	ASV6/Mamiellaceae	ASV7/ Phaeocystaceae	ASV14/ Mamiellaceae	ASV16/ Mamiellales	ASV30/Dino-Group-I-Clade-1	ASV99/Dino-Group-II-Clade-10-and-11	ASV170/Dino-Group-I-Clade-4	ASV212/Bathycocaceae	ASV227/Dino-Group-II-Clade-10-and-11	ASV457/ Group-II_X
Lofoten HE533	ASV1/ Cryptomonadales_X	ASV4/ Cryptomonadales_X	ASV6/ Mamiellaceae	ASV7/ Phaeocystaceae	ASV14/Mamiellaceae	ASV16/Mamiellales	ASV23/Parmales_X	ASV25/Telonemniacae	ASV27/ Leegardellidae_B	ASV39/ Stramenopiles
Nordfjord	ASV4/ Cryptomonadales_X	ASV6/Mamiellaceae	ASV7/ Phaeocystaceae	ASV12/Pyramimonadales_XX	ASV14/Mamiellaceae	ASV16/Mamiellales	ASV47/Dino-Group-II-Clade-14	ASV50/Chrysochromulinales	ASV74/Dino-Group-II	ASV168/ Chytridiaceae
Sognefjord	ASV6/Mamiellaceae	ASV7/ Phaeocystaceae	ASV16/ Mamiellales	ASV74/Dino-Group-II	ASV86/Dino-Group-II-Clade-6	ASV99/Dino-Group-II-Clade-10-and-11	ASV136/Dino-Group-II-Clade-6	ASV157/Dino-Group-I-Clade-4	ASV168/ Chytridiaceae	ASV304/Dino-Group-II-Clade-20
Boknafjord	ASV6/Mamiellaceae	ASV7/ Phaeocystaceae	ASV16/ Mamiellales	ASV30/Dino-Group-I-Clade-1	ASV31/Picozoa_XXX	ASV86/Dino-Group-II-Clade-6	ASV99/Dino-Group-II-Clade-10-and-11	ASV164/Phaeocystaceae	ASV187/ Mamiellaceae	ASV435/Dino-Group-II-Clade-10-and-11
Onust-Tjøn Fjord	ASV4/ Cryptomonadales_X	ASV6/Mamiellaceae	ASV7/ Phaeocystaceae	ASV12/Pyramimonadales_XX	ASV14/Mamiellaceae	ASV18/ Dinophyceae_XX	ASV31/Picozoa_XXX	ASV66/Dino-Group-I-Clade-5	ASV187/ Mamiellaceae	ASV331/ Bathycocaceae
Iceland	ASV6/Mamiellaceae	ASV14/ Mamiellaceae	ASV16/ Mamiellales	ASV20/ Dinophyceae_XX	ASV21/ Dinophyceae_XX	ASV30/Dino-Group-I-Clade-1	ASV50/Chrysochromulinales	ASV80/Dino-Group-I-Clade-1	ASV89/ Dinophyceae_XX	ASV317/Dino-Group-II-Clade-10-and-11
Van Mijen Fjord	ASV2/Mamiellaceae	ASV6/Mamiellaceae	ASV14/ Mamiellaceae	ASV16/ Mamiellales	ASV30/Dino-Group-I-Clade-1	ASV87/ Mamiellaceae	ASV175/unknown	ASV176/unknown	ASV188/ MAST-9D_X	ASV250/unknown

Isfjord	ASV2 Mamiellaceae	ASV6 Mamiellaceae	ASV14 Mamiellaceae	ASV16 Mamiellales	ASV30 Dino-Group-I-Clade-1	ASV50 Chrysoschro mulinaceae	ASV65/Radial-centric-basal-Coscinodiscophyceae	ASV67 Phaeocystaceae	ASV158 Radial-centric-basal-Coscinodiscophyceae	ASV223 Radial-centric-basal-Coscinodiscophyceae
Kongsfjord	ASV2 Mamiellaceae	ASV6 Mamiellaceae	ASV7 Phaeocystaceae	ASV14 Mamiellaceae	ASV15 Mamiellales	ASV16 Mamiellales	ASV30/DinoGroup-I-Clade-1	ASV44 Mamiellaceae	ASV47 Dino-Group-II-Clade-14	ASV50 Chrysoschro mulinaceae
Woodfjord	ASV2 Mamiellaceae	ASV6 Mamiellaceae	ASV7 Phaeocystaceae	ASV14 Mamiellaceae	ASV16 Mamiellales	ASV38 Chlorarachniophyceae_X	ASV39 Stramenopiles	ASV51/Protaspalineage	ASV67 Phaeocystaceae	ASV84 Phaeocystaceae
Wijdefjord	ASV2 Mamiellaceae	ASV6 Mamiellaceae	ASV7 Phaeocystaceae	ASV14 Mamiellaceae	ASV16 Mamiellales	ASV38 Chlorarachniophyceae_X	ASV39 Stramenopiles	ASV51 Protaspalineage	ASV84 Stramenopiles	ASV203 Mamiellaceae
Disco_Bay	ASV2 Mamiellaceae	ASV6 Mamiellaceae	ASV7 Phaeocystaceae	ASV12/Pyraminonadales_XX	ASV14 Mamiellaceae	ASV20 Dinophyceae_XX	ASV22 Chrysoschro mulinaceae	ASV37/Polarcentric-Mediphyceae	ASV77/Dicthyophyceae_XX	ASV139/DinoGroup-II-Clade-10-and-11
Nordvest fjord	ASV2 Mamiellaceae	ASV6 Mamiellaceae	ASV15 Mamiellaceae	ASV16 Mamiellales	ASV30 Dino-Group-I-Clade-1	ASV44 Mamiellaceae	ASV105 Mamiellaceae	ASV111 Mamiellophyceae	ASV231 Synurales_X	ASV347/Chrysophyceae_Clade-G

	Prokaryotes									
Fjord	1	2	3	4	5	6	7	8	9	10
Tanaifjord	ASV11/SAR92 clade	ASV14 NS4 marine group	ASV15 Cryomorphaceae	ASV2 Nitrincolaceae	ASV19 Nitrincolaceae	ASV28 Aurantivirga	ASV3 Amylibacter	ASV5 SUP05 cluster	ASV7 Sulfitobacter	ASV76 Alteromonas
Lakse fjord	ASV11/SAR92 clade	ASV13 Aquibacter	ASV14 NS4 marine group	ASV2 Nitrincolaceae	ASV19 Nitrincolaceae	ASV29 Polaribacter 1	ASV3 Amylibacter	ASV5 SUP05 cluster	ASV6 Polaribacter 3	ASV7 Sulfitobacter
Porsanger fjord	ASV1/ Pseudoalteromonas	ASV14 NS4 marine group	ASV15 Cryomorphaceae	ASV2 Nitrincolaceae	ASV19 Nitrincolaceae	ASV29 Polaribacter 1	ASV3 Amylibacter	ASV5 SUP05 cluster	ASV6 Polaribacter 3	ASV7 Sulfitobacter
Lyngen fjord	ASV1/ Pseudoalteromonas	ASV11 SAR92 clade	ASV13 Aquibacter	ASV15 Cryomorphaceae	ASV17 Flavobacteriaceae	ASV2 Nitrincolaceae	ASV3 Amylibacter	ASV36 Colwellia	ASV6 Polaribacter 3	ASV9/SAR86 clade
Balsfjord	ASV1/ Pseudoalteromonas	ASV11 SAR92 clade	ASV16 Methylophagaceae	ASV2 Nitrincolaceae	ASV20 Ascidiacehabitans	ASV22 Ulvibacter	ASV23/Pseudoalteromonas	ASV3 Amylibacter	ASV41 Vibrio	ASV5 SUP05 cluster
Lofoten HE431	ASV12/ SAR86 clade	ASV24 NS5 marine group	ASV25 Arenicellaceae	ASV30 Clade IV	ASV3 Amylibacter	ASV35/Synechococcus CC9902	ASV39/AEGEAN -169 marine group	ASV40/Synechococcus CC9902	ASV73/Synechococcus CC9902	ASV9/SAR86 clade
Lofoten HE533	ASV1 Pseudoalteromonas	ASV17 Flavobacteriaceae	ASV2 Nitrincolaceae	ASV20 Ascidiacehabitans	ASV23 Pseudoalteromonas	ASV3 Amylibacter	ASV37 Pseudofuivibacter	ASV41 Vibrio	ASV5 SUP05 cluster	ASV9/SAR86 clade

Nordfjord	ASV12 clade	SAR86	ASV25 Arenicellaceae	ASV3 Amylibacter	ASV35/Synechococcus CC9902	ASV39 AEGEAN-169 marine group	ASV40/Synechococcus CC9902	ASV8 Planktomarina	ASV70/OM60(N OR5) clade	ASV94/AEGEAN-169 marine group
Sogne fjord	ASV12 clade	SAR86	ASV25 Arenicellaceae	ASV30 Clade IV	ASV3 Amylibacter	ASV35/Synechococcus CC9902	ASV39/AEGEAN-169 marine group	ASV8 Planktomarina	ASV82 Marine Group II	ASV77 SAR116 clade
Bokna fjord	ASV12 clade	SAR86	ASV137 Clade IV	ASV20 Ascidiaceihabians	ASV30 Ascidiaceihabians	ASV35/Synechococcus CC9902	ASV39/AEGEAN-169 marine group	ASV43Candidatus Actinomarina	ASV8 Planktomarina	ASV94/AEGEAN-169 marine group
Orust-Tjörn fjord	ASV132 clade	OM43	ASV40/Synechococcus CC9902	ASV43Candidatus Actinomarina	ASV49 SAR86 clade	ASV54 Pseudohongiella	ASV68 NS4 marine group	ASV80 Lentibacter	ASV70/OM60(N OR5) clade	ASV77 SAR116 clade
Iceland	ASV1 Pseudoalteromonas	SAR86	ASV12 SAR86 clade	ASV13 Aquibacter	ASV23 Pseudoalteromonas	ASV253 Psychrobacter	ASV30 Ascidiaceihabians	ASV34 Clade Ia	ASV35/Synechococcus CC9902	ASV9 SAR86 clade
Van.Mijken Fjord	ASV12 clade	SAR86	ASV2 Nitriocolaceae	ASV24 NS5 marine group	ASV23 Pseudoalteromonas	ASV3 Amylibacter	ASV34 Clade Ia	ASV5 SUP05 cluster	ASV7 Sulfotobacter	ASV8 Planktomarina
Isfjord	ASV12 clade	SAR86	ASV13 Aquibacter	ASV2 Nitriocolaceae	ASV209 Flavobacterium	ASV39/AEGEAN-169 marine group	ASV5 SUP05 cluster	ASV8 Planktomarina	ASV81 Luminiphilus	ASV9 SAR86 clade
Kongsfjord	ASV11 clade	SAR92	ASV13 Aquibacter	ASV16 Methylophagaceae	ASV2 Nitriocolaceae	ASV24 NS5 marine group	ASV34 Clade Ia	ASV6 Polaribacter	ASV7 Sulfotobacter	ASV8 Planktomarina
Woodfjord	ASV11 clade	SAR92	ASV119 Polaribacter 1	ASV13 Aquibacter	ASV2 Nitriocolaceae	ASV209 Flavobacterium	ASV24 NS5 marine group	ASV6 Polaribacter	ASV8 Planktomarina	ASV9 SAR86 clade
Wijdefjord	ASV11 clade	SAR92	ASV119 Polaribacter 1	ASV13 Aquibacter	ASV2 Nitriocolaceae	ASV24 NS5 marine group	ASV211 Cryomorphaceae	ASV6 Polaribacter	ASV7 Sulfotobacter	ASV8 Planktomarina
Disco_Bay	ASV11 clade	SAR92	ASV13 Aquibacter	ASV15 Cryomorphaceae	ASV2 Nitriocolaceae	ASV22 Ulvibacter	ASV213 Colwelliaceae	ASV3 Amylibacter	ASV6 Polaribacter 3	ASV7 Sulfotobacter
Nordvest fjord	ASV113 Colwellia		ASV1/Pseudoalteromonas	ASV11 SAR92 clade	ASV12 SAR86 clade	ASV156 marine group	ASV2 Nitriocolaceae	ASV8 Planktomarina	ASV87 Colwellia	ASV99Candidatus Aquiluna

Supplementary 4

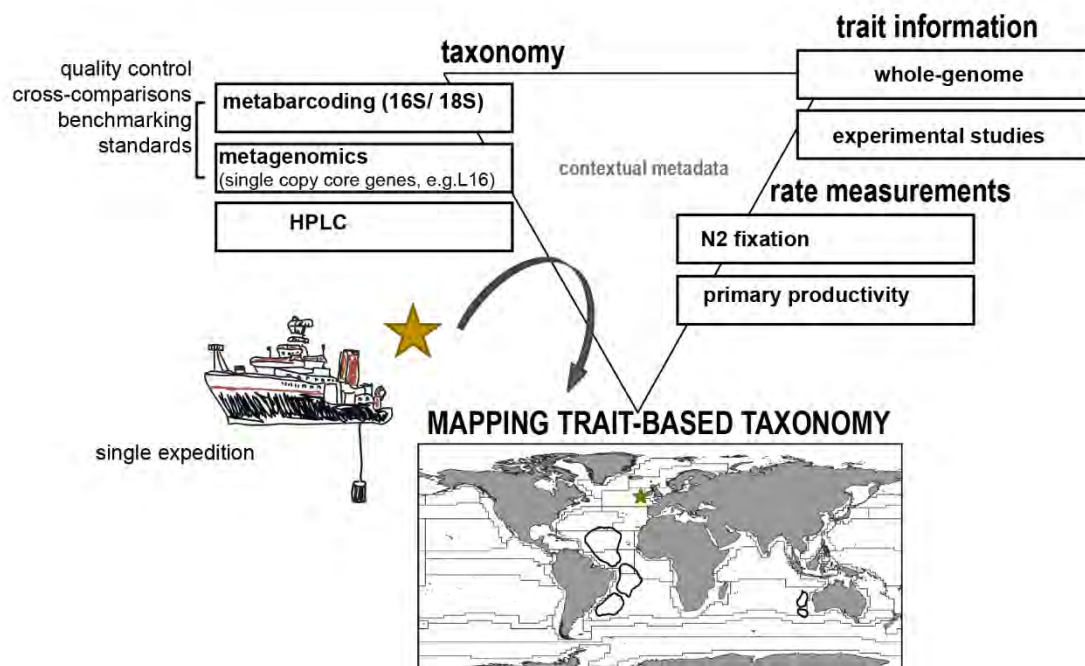


Figure S1. Illustration of microbial rate measurements, phylogenetic and functional microbial data that are interlinked and cross-compared that any single observation can be nested taxonomically and functionally in a global-scale analysis.

Statutory Declaration

Family Name, Given/First Name	Hörstmann, cora
Matriculationnumber	20332220
What kind of thesis are you submitting: Bachelor-, Master- or PhD-Thesis	PhD thesis

English: Declaration of Authorship

I hereby declare that the thesis submitted was created and written solely by myself without any external support. Any sources, direct or indirect, are marked as such. I am aware of the fact that the contents of the thesis in digital form may be revised with regard to usage of unauthorized aid as well as whether the whole or parts of it may be identified as plagiarism. I do agree my work to be entered into a database for it to be compared with existing sources, where it will remain in order to enable further comparisons with future theses. This does not grant any rights of reproduction and usage, however.

The Thesis has been written independently and has not been submitted at any other university for the conferral of a PhD degree; neither has the thesis been previously published in full.

German: Erklärung der Autorenschaft (Urheberschaft)

Ich erkläre hiermit, dass die vorliegende Arbeit ohne fremde Hilfe ausschließlich von mir erstellt und geschrieben worden ist. Jedwede verwendeten Quellen, direkter oder indirekter Art, sind als solche kenntlich gemacht worden. Mir ist die Tatsache bewusst, dass der Inhalt der Thesis in digitaler Form geprüft werden kann im Hinblick darauf, ob es sich ganz oder in Teilen um ein Plagiat handelt. Ich bin damit einverstanden, dass meine Arbeit in einer Datenbank eingegeben werden kann, um mit bereits bestehenden Quellen verglichen zu werden und dort auch verbleibt, um mit zukünftigen Arbeiten verglichen werden zu können. Dies berechtigt jedoch nicht zur Verwendung oder Vervielfältigung.

Diese Arbeit wurde in der vorliegenden Form weder einer anderen Prüfungsbehörde vorgelegt noch wurde das Gesamtdokument bisher veröffentlicht.

19.05.2022 
.....
Date, Signature